



# Deconstructing Darwin's Naturalization Conundrum in the San Juan Islands using community phylogenetics and functional traits

Hannah E. Marx<sup>1\*</sup>, David E. Giblin<sup>2</sup>, Peter W. Dunwiddie<sup>2</sup> and David C. Tank<sup>1</sup>

<sup>1</sup>Department of Biological Sciences and Institute for Bioinformatics and Evolutionary Studies, University of Idaho, Moscow, ID 83844, USA, <sup>2</sup>Herbarium, Burke Museum of Natural History and Culture, University of Washington, Seattle, WA 98195, USA

## ABSTRACT

**Aim** Darwin posed a conundrum about species invasions, postulating the importance of functional distinctiveness from the receiving native community to avoid competition, and, at the same time, the importance of shared similarity to pass environmental filters and successfully establish. Using phylogenetic distances and functional traits, we assessed this conundrum in the flora of 80 mostly uninhabited islands, where over 30% of the species are invasive. We highlight the importance of publicly available datasets to disentangle ecological processes that may drive invasion.

**Location** San Juan Islands archipelago, Pacific Northwest of North America.

**Methods** Using a supermatrix approach, we inferred a maximum-likelihood estimate of the mega-phylogeny for the vascular plants on the San Juan Islands. We gathered measurements for five ecologically relevant functional traits – seed mass, maximum height, specific leaf area, leaf size and leaf nitrogen content. We assessed phylogenetic and functional trait similarity between invasive species and the receiving native community, and tested the significance of the observed patterns against a randomly generated invading community.

**Results** Invasive species were more closely related (phylogenetically clustered) to their nearest native than natives were on 40 of the islands and were more clustered than any random invasive in the species pool on 22 islands. Despite phylogenetic similarity, functional traits differed between the two status groups, at least for maximum height and specific leaf area. When comparing functional differences between phylogenetically close relatives, more complex patterns emerge.

**Main conclusion** Only with the combination of both evolutionary history and phenotypic traits were we able to discover support for both sides of Darwin's conundrum – although invasive species have phylogenetically close native relatives, functional traits differ between the two status groups. This implies that both environmental filtering and competitive interactions may be important for invasion success in this archipelago.

## Keywords

biological invasions, community phylogenetics, Darwin's Naturalization Conundrum, functional traits, islands, mega-phylogeny.

\*Correspondence: Hannah E. Marx, Department of Biological Sciences and Institute for Bioinformatics and Evolutionary Studies, University of Idaho, Moscow, ID 83844, USA.  
E-mail: h.marx.h@gmail.com

## INTRODUCTION

Species distributions have constantly been reshuffled on the earth's surface, implicating many species as invasive to a given community at some point. Ranges of invasive species

are currently increasing with human-mediated dispersal – especially on islands (Pysek & Richardson, 2006) – and in the face of accelerating global change, it is of economic importance to predict which species are most likely to become invasive and their potential for future spread and

impact (Bellard *et al.*, 2013). Modern species introductions are an important process to study not only for economic and conservation concerns, but also, when considered as recent colonizers in an ecological community, invaders can be used to understand the basic processes of community assembly in natural systems (Tilman, 2004).

Darwin was among the first to realize the value of invasive species in eco-evolutionary studies for understanding community assembly processes. He postulated that the more distantly related species were to each other, the greater their difference in functional traits should be, and thus, the less competition for resources they would experience due to 'diversification of structure' (Darwin, 1859). Therefore, invasive species that are distantly related to the native community would be more likely to possess 'novel' traits allowing them to successfully establish via niche differentiation and escape from competition, a prediction that is often defined as 'Darwin's Naturalization Hypothesis' (DNH, Rejmánek, 1996). However, Darwin also discussed the importance of ancestral similarity and hypothesized that pre-adapted traits would be important to pass environmental filters and survive within a particular environment. Thus, Darwin presented a conundrum, predicting both similarity and distinctiveness to be important for invasion success.

To assess 'Darwin's Naturalization Conundrum' (DNC, Diez *et al.*, 2008), the use of phylogenies is becoming standard (e.g. Strauss *et al.*, 2006b; Cadotte *et al.*, 2009b, 2010; Davies *et al.*, 2010; Schaefer *et al.*, 2011; Carboni *et al.*, 2012; Bezeng *et al.*, 2013; Park & Potter, 2013; Castro *et al.*, 2014; Lim *et al.*, 2014; Ordonez, 2014; Li *et al.*, 2015; Lososová *et al.*, 2015). The inclusion of evolutionary information using a phylogenetic approach provides an accurate quantification of relatedness by comparing phylogenetic distances (i.e. difference in branch lengths) between native and invasive species (Procheş *et al.*, 2008; Cadotte *et al.*, 2009b; Thuiller *et al.*, 2010), and relatedness has been shown experimentally to be a useful predictor of invasion success (Jiang *et al.*, 2010; (Cadotte & Strauss, 2011), but see Castro *et al.*, 2014). In addition, because related species are more likely to share ecologically relevant functional traits (Webb *et al.*, 2002), phylogenetic distance as measured on a molecular phylogeny of the species within a community is often viewed as a proxy for ecological similarity (Cadotte *et al.*, 2009a). However, because complex mechanisms like competition can produce contrasting community phylogenetic patterns (Mayfield & Levine, 2010; Jones *et al.*, 2013; Godoy *et al.*, 2014), linking an ecological process is controversial, especially at varying spatial scales (reviewed in Pavoine & Bonsall, 2010). Functional trait-based metrics combined with evolutionary distances help to distinguish between different assembly scenarios (Cadotte *et al.*, 2009a; Mouquet *et al.*, 2012), because both could provide complementary information about the ecological interactions or processes that shape communities (Cadotte *et al.*, 2013).

Due to increasing availability of molecular sequence data along with advances in phylogenetic methods (e.g. Smith *et al.*, 2009), detailed regional species inventories (e.g.

Cadotte *et al.*, 2006) and the availability of large trait databases (e.g. Kattge *et al.*, 2011), more studies are integrating phylogenies and functional traits to investigate DNC. These studies have been conducted across natural systems ranging from coastal Mediterranean dunes (Carboni *et al.*, 2012) to global communities defined by habitat types (Ordonez, 2014), within particular clades (Strauss *et al.*, 2006a; Park & Potter, 2013), or for all plants in a community (Schaefer *et al.*, 2011; Carboni *et al.*, 2012; Lim *et al.*, 2014; Ordonez, 2014; Li *et al.*, 2015; Lososová *et al.*, 2015). Some have used a mega-phylogenetic approach to estimate a community phylogeny (e.g. Schaefer *et al.*, 2011; Park & Potter, 2013; Lim *et al.*, 2014; Li *et al.*, 2015), while other studies have combined existing phylogenies in a supertree approach (e.g. Strauss *et al.*, 2006b; Carboni *et al.*, 2012; Ordonez, 2014; Lososová *et al.*, 2015). Each study has used different functional traits, different definitions for 'invasive' species, included trait information in different ways, and importantly, used different approaches to test the significance of their patterns and assess hypotheses of species invasiveness or the invasibility of a community. Not surprisingly, their results come to various conclusions when testing DNC (reviewed in Thuiller *et al.*, 2010, Table S1).

For example, in the angiosperm flora of the Azores, Schaefer *et al.* (2011) found invasive species were more phylogenetically distant from the native community than non-invasive introduced species on islands >1 km<sup>2</sup>, and although invasive species were not closely related to the native flora, their functional similarity for most traits suggested support for 'enemy release' in determining the probability of invasion. However, they did not test these observed patterns against any random expectation. Carboni *et al.* (2012) found invasive species were phylogenetically distant from the native assemblages in 4 m<sup>2</sup> plots on invaded coastal dunes, but at coarser spatial scales they were more closely related, and this pattern was robust when randomizing the invaded community. Phylogenetic signals for a few traits indicated the phylogeny was a good proxy for shared ecological similarity, so they hypothesized that their observed pattern was due to biotic resistance through competition at fine spatial resolutions, and habitat filtering at coarser spatial scales. In contrast, Park & Potter (2013) found phylogenetically close relatives in the Cardueae to be more successful invaders in California (but see Sol *et al.*, 2014; Cadotte, 2014; for a critical look at the assumptions and methods of this study), and Lososová *et al.* (2015) found invasives closely related to the native flora of the Czech Republic. A global comparison of vascular plants showed that alien taxa tended to have close native relatives in the community, yet were functionally distinct, highlighting the importance of both environmental filtering and competitive interactions in the naturalization process (Ordonez, 2014). Different methodologies make these studies difficult to compare, and hypotheses often focus on one side of Darwin's conundrum, so explanations for the ecological processes potentially driving these patterns of invasion are often weakly supported or conflicting.

Ordóñez (2014) recently shed light on DNC by defining the position of invasive species relative to the realized niche space of the native community in terms of both functional and phylogenetic diversity. When framed in terms of functional niche space, Darwin's conundrum is instead a series of alternate hypotheses. Irrespective of the ecological mechanism, such patterns could be used to predict naturalization success. Thus, invasive species could successfully naturalize under three scenarios: (1) if they are completely distinct from the recipient native community, distantly related and functionally unique (traditionally 'Darwin's Naturalization Hypothesis' (DNH), Rejmánek, 1996; Daehler, 2001), (2) if they match the niche of the recipient native community, sharing functional traits irrespective of the phylogenetic relationships ('matching hypothesis', Procheş *et al.*, 2008; Ordóñez, 2014) or (3) if they differ in functional traits to fill unoccupied niche spaces even if phylogenetically similar ('filling hypothesis', Procheş *et al.*, 2008; Ordóñez, 2014).

A comprehensive botanical survey of 80 islands in the San Juan archipelago off the north-west corner of the United States revealed that most of these native communities have been invaded by introduced plant species. While this is perhaps not surprising given increased introductions of plants on islands world-wide (Sax & Gaines, 2008) and predicted drastic range shifts globally (Sala *et al.*, 2000; Bellard *et al.*, 2013), only a few of these islands have year-round human inhabitants, and those that are inhabited are limited to caretakers and seasonal recreation. European history is relatively recent in north-western North America; therefore, these invasions have occurred within the last 200 years. Furthermore, the extensive replication of invasions across islands ranging in size from 1 m<sup>2</sup> to 0.57 km<sup>2</sup> provides an ideal system to investigate patterns of invasion in the light of both phylogenetic and functional similarity.

The definition of what constitutes an 'invasive' species is often difficult to determine (Vermeij, 1996; Richardson & Pyšek, 2006). Daehler (1998) distinguished 'invasive' introduced plants as introduced species that have self-sustaining populations in natural, undisturbed areas (in contrast to agricultural lands, urban areas, roadsides, etc.), and in an explicit effort to be consistent with other comparable studies (e.g. Schaefer *et al.*, 2011), this is the definition that we follow here. Given the hypotheses that we are investigating, this geographic criterion for calling introduced species 'invasive' is more meaningful than impact-based definitions of invasive species that may depend less on species composition of local ecosystems, and more upon climatic and edaphic conditions and levels of disturbance (Diez *et al.*, 2009; Schaefer *et al.*, 2011). Because the islands included in this study were mostly uninhabited and relatively undisturbed (besides a few islands with campsites), we categorized all non-native plants in the sampled islands as 'invasive' (as in Schaefer *et al.*, 2011), but note that to definitively classify all naturalized plants as invasive would require information on their relative abundance, exact timing of introduction and surveys monitoring the spread of each species (Richardson *et al.*, 2000).

In this study, we assessed Darwin's Naturalization Conundrum in the San Juan Islands in light of three scenarios that could explain invasion success – DNH, 'matching' or 'filling' hypotheses – using metrics and approaches that are comparable to previous studies. We used branch lengths from a time-calibrated species-level mega-phylogeny to calculate the evolutionary distances between invasive and native plants across all 80 islands in the archipelago and compared measurements for five ecologically relevant traits between native and invasive species. Specifically, we addressed (1) how invasive plants are related to their nearest native relative and to the greater native community and (2) if invasive and native species differ in measurements for phenotypic traits overall, compared to their nearest native, or to the average of the native community. We highlight the importance of biological surveys and publically available datasets to describe multiple aspects of biodiversity patterns, and the use of biologically meaningful null expectations to understand macro-ecological processes of invasion.

## METHODS

### Floristic survey and island community matrix

Floristic surveys were conducted across islands in the San Juan archipelago from 2005 to 2010 to compile a complete inventory of the vascular flora on each island visited. Therefore, surveys focused only on smaller islands – generally <25 ha – where we were confident that all habitats could be carefully searched. Permission to conduct surveys and make collections was secured in advance from appropriate public agencies and several private owners. Islands were accessed using small launches by survey teams that included multiple experts in the local flora to ensure consistent and thorough recognition of all species. To document the flora of each island as completely as possible, multiple visits were made that were timed to facilitate detection of species throughout the growing season, in late April, late May and early September. The invasive or native status of each species was determined by multiple floras (see complete reference list in Appendix S1).

### Phylogeny estimation

Recent studies testing DNC have incorporated phylogenetic distances using either a Phylomatic supertree (e.g. Ordóñez, 2014) or a mega-phylogeny approach (e.g. Schaefer *et al.*, 2011), which builds a community phylogeny from a supermatrix of DNA sequence data (Smith *et al.*, 2009) and incorporates uncertainty in topology and branch lengths using statistical phylogenetic methods. With the constant expansion of readily available sequence data, the advantages of the mega-phylogeny approach to macro-ecological studies are becoming clear (Roquet *et al.*, 2012). To assess the relationships among vascular plants in the San Juan Islands community, we took a supermatrix approach to infer a mega-phylogeny by retrieving publically available sequence data from GenBank for five gene

regions (*atpB*, *rbcl*, *matK*, *trnTLF* and ITS) using the PHLAWD pipeline (Smith *et al.*, 2009). A final by-gene partitioned maximum-likelihood (ML) estimate of species relationships for all vascular plants in the San Juan Islands species pool was inferred using the concatenated alignment, implemented in RAxML version 7.4.2 (Stamatakis, 2006).

Generating a tree with branch lengths proportional to time at this scale is a non-trivial task, so other large-scale community phylogenetic analyses have used trees with relativized branch lengths that do not incorporate real-time estimates (e.g. Webb, 2000). This is less than ideal and makes it difficult to compare metrics across studies. In contrast, time-scaled branch lengths place communities on the same axis making such comparisons possible. Because the focus (i.e. community level) and scale (i.e. vascular plants) of this phylogeny is inappropriate for direct divergence time estimation using fossil calibrations, the mega-phylogeny estimate was scaled to time using the 'congruification' approach (Eastman *et al.*, 2013), so that this community phylogeny will be comparable to others in the future (see Appendix S1 for a detailed explanation of methods for phylogeny estimation and time-scaling).

### Functional traits

Measurements for seed mass (mg), maximum height (m), specific leaf area (SLA,  $\text{cm}^2 \text{g}^{-1}$ ), leaf size (cm) and leaf nitrogen content (%) are a subset of a larger dataset assembled by Cornwell *et al.* (2014), with original data compiled from a number of sources for each trait. All measurements and citations for these functional traits are available in Appendix S1. Because the values were not normally distributed, trait measurements were log transformed (Westoby, 1998). In cases where there were multiple trait measurements for a species (e.g. for intraspecific taxa), the harmonic mean of all values was used.

In addition to being well sampled across land plants, these five traits have gained acceptance as measurable proxies to link variation in plant phenotypes to ecological strategies (Westoby *et al.*, 2002; Reich *et al.*, 2003; Cornwell *et al.*, 2014). For example, SLA, maximum height and seed mass comprise the leaf-height-seed (LHS) scheme, widely used to relate interspecific trait variation to trade-offs in ecological strategies (Grime, 1974; Westoby, 1998), in which shorter plants with thicker, narrower leaves (low SLA) fall towards the 'stress-tolerant' pole of the functional spectrum.

To test the assumption of niche conservatism, or shared similarity of traits between close relatives (Webb *et al.*, 2002; Cavender-Bares *et al.*, 2004), we used Abouheif's test (Abouheif, 1999) to calculate the phylogenetic autocorrelation between traits (see Appendix S1 for details).

### Statistical analyses and assessment of Darwin's Naturalization Conundrum

Although previous tests of DNC have used a variety of different metrics to summarize phylogenetic and functional

diversity (Table S1), most have calculated some measure of 'nearest taxon distance', capturing interactions between close relatives (i.e. 'tip effects'), and 'mean pairwise distance', describing community-wide interactions (reviewed in Thuiller *et al.*, 2010). We assessed the phylogenetic distinctiveness of invasive plants by calculating the distance to the nearest native species (DNNS) and the mean phylogenetic distance to each co-occurring native species (MDNS).

To test whether invasive plants are functionally unique, we calculated the difference in trait measurements (1) between invasive and native status groups, (2) between each species and its phylogenetically nearest native relative (the nearest native functional difference, NNFD) and (3) between each species and the average of all co-occurring natives (the mean functional difference, MFD). When there were multiple equally distant nearest native relatives, the median trait measurement was used to calculate the NNFD. Functional distinctiveness was calculated as a difference rather than a distance (e.g. Euclidean). In this way, we were able to address the magnitude of change in trait measurements between species overall, compared to their nearest native relative, or to the average of the native community in which they have naturalized.

To summarize patterns across all islands, the dated ML phylogeny of the vascular flora of the San Juan Islands was pruned to include just species co-occurring on a focal island (i.e. community). For each native (n) and invasive (i) species, phylogenetic ( $\text{DNNS}_i/\text{DNNS}_n$ ;  $\text{MDNS}_i/\text{MDNS}_n$ ) and functional distinctiveness ( $\text{NNFD}_i/\text{NNFD}_n$ ;  $\text{MFD}_i/\text{MFD}_n$ ) metrics were calculated. A t-test was used to evaluate significant differences in observed means of each metric between status groups. To assess DNC, we compared the direction of the difference in means for each significant relationship. For example, if the mean metric for invasive species (e.g. mean  $\text{DNNS}_i$ ) was greater than the mean metric for native species (e.g. mean  $\text{DNNS}_n$ ), we would conclude (phylogenetic and/or functional) distinctiveness of the invasives from the natives to be important for the successful naturalization in that community.

Finally, we verified if the observed patterns were significantly different from those given any other random invasive species that could potentially colonize each community. We randomly assembled invasive communities, drawing from the pool of all invasive species occurring in the San Juan archipelago without replacement, while preserving the observed number of invasives on each island. This null model of invasion maintains the evolutionary structure of the native community and the richness of the invasive species within each community, but randomizes the identity and the evolutionary distances of the invasive species. For each simulated community, the mean of each metric was recalculated, and the randomization was repeated 1000 times to obtain a null distribution of means. The observed means were compared to each null distribution by measuring the standardized effect size (SES) which is the difference in the observed and expected mean under the null model, divided by the

standard deviation of the distribution of the null model. This is similar to the SES described in the R package *picante* (Kembel *et al.*, 2010), and equivalent to -1 times NRI or NTI described in Webb *et al.* (2002). Positive values of the  $SES_{metric}$  (z-scores) indicate that observed invasive species are more distinct (phylogenetically or functionally) from the nearest native or native community than any other random invasive species in the species pool (i.e. phylogenetic overdispersion of invasive species compared to the natives), and negative values indicate that invasives are more similar to the nearest native and/or native community (i.e. phylogenetic clustering). Statistical significance was determined for each  $SES_{metric}$  by calculating *P*-values ( $\alpha = 0.05$ ), the proportion of simulated means that were as or more extreme than the observed mean (Kembel *et al.*, 2010). All statistical analyses were conducted in R (R Development Core Team, 2014).

## RESULTS

### Island communities

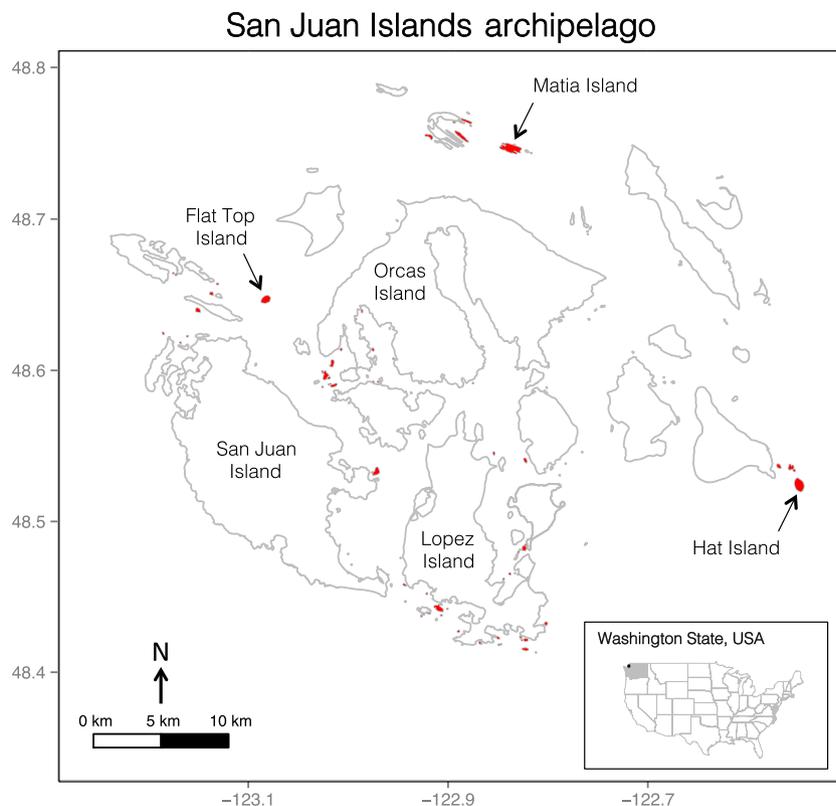
Voucher specimens were deposited in the Burke Museum of Natural History and Culture Herbarium at the University of Washington (WTU), and comprehensive lists were compiled of all species occurring on 80 islands. A total of 442 species of vascular plants were identified and are included in the Floristic Atlas of the San Juan Islands, Washington, which is publically available online (Floristic Atlas of the San Juan Islands, Washington, 2013).

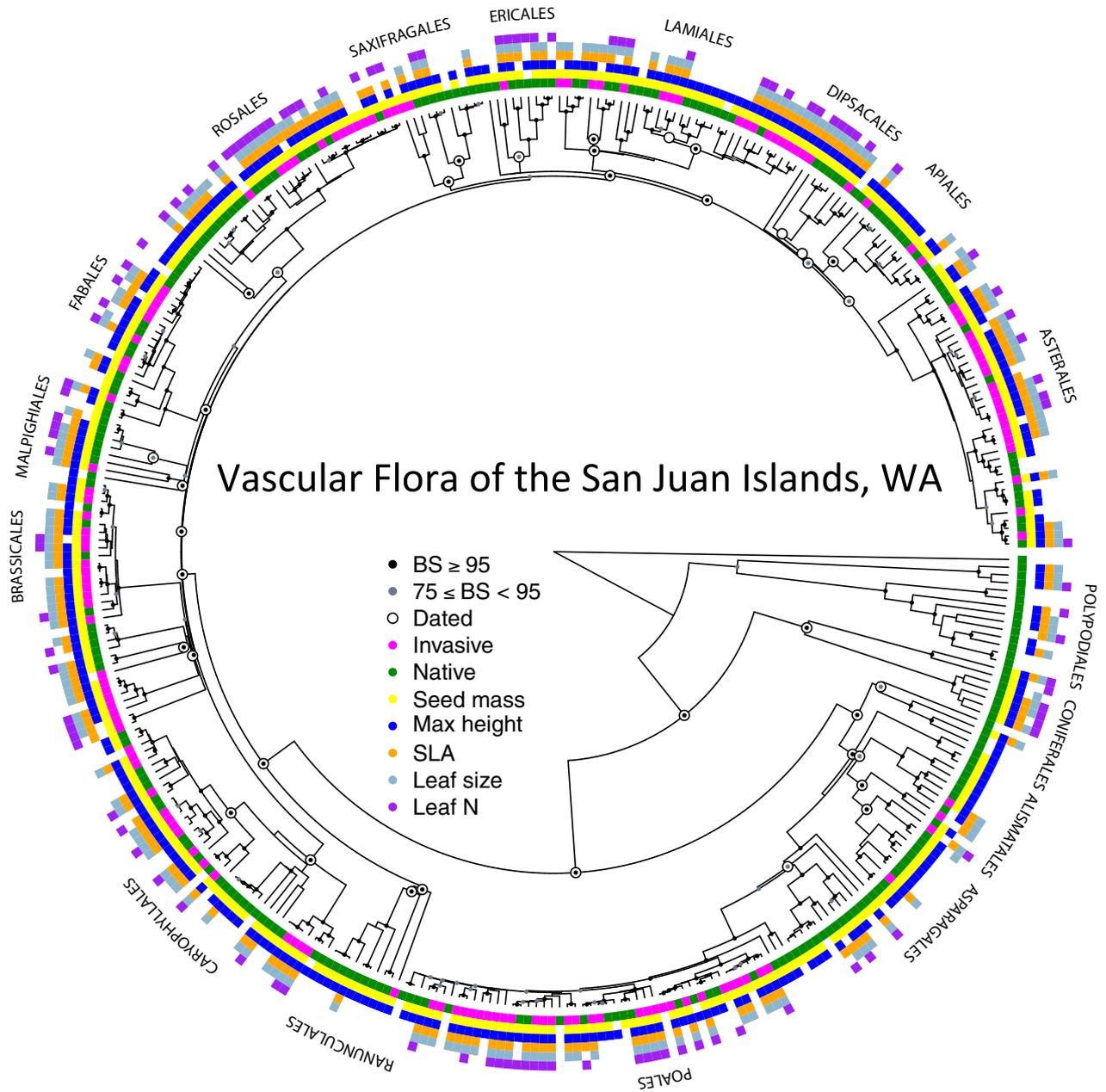
Intraspecific taxa were collapsed to species to avoid pseudoreplication, resulting in a total of 415 species used in this study, 150 (36%) of which were identified as invasive. Islands with species lacking available sequence data (Little Oak 2, Swirl Rock East), in which there was only one native (Smallest unnamed island by Long Island), only one invasive (Unnamed west of Castle Island), or no invasive species (East Sucia 5 Island, Shag Reef, Smaller Island near Charles and Swirl Rock West) were removed, resulting in a total of 72 islands used in the following analyses (Fig. 1). Islands were binned by size categories for discussion, defined by the 25% and 75% sample quartiles of island area (Appendix S2 and S3).

### Phylogeny estimation

After removal of GenBank sequences identified as submission errors, genetic data for 366 of 415 (88%) species were retained, with a total combined aligned length of 5745 base pairs (see Table S2). Because there were no major conflicts among the individual gene trees (see Fig. S1), the concatenated mega-phylogeny was used for all downstream analyses (Figs 2 and S2 for labelled tips). Highly supported nodes were distributed evenly across the tree, resolving all deep relationships with high support and very few inconsistencies at the genus level. For example, within Asteridae (bootstrap support = 96%), the Ericales are sister to the subclade Lamiidae + Campanulidae (bootstrap support = 98%), consistent with established angiosperm relationships (e.g. Soltis *et al.*, 2011, see Fig. S3 for all bootstrap support values).

**Figure 1** Map of the San Juan Islands archipelago, Washington State, USA. Islands that were sampled in this study are coloured in red. Three of the larger surrounding islands (unsampled) and three of the largest sampled islands are labelled for reference.





**Figure 2** Time-calibrated maximum-likelihood mega-phylogeny of the vascular flora of the San Juan Islands, with native species (green) and invasive species (pink) indicated by the first character state across the tips. Nodes with a light grey dot have bootstrap support (BS) between 75 and 95, and nodes with a solid black dot have bootstrap support >95. Nodes that were congruent ('congruified') between the reference timetree and the target tree (community phylogeny) are circled in a solid black line. Coloured bars next to each taxon represent the presence of data for the five functional traits: seed mass, maximum height, specific leaf area (SLA), leaf size and leaf nitrogen content (leaf N).

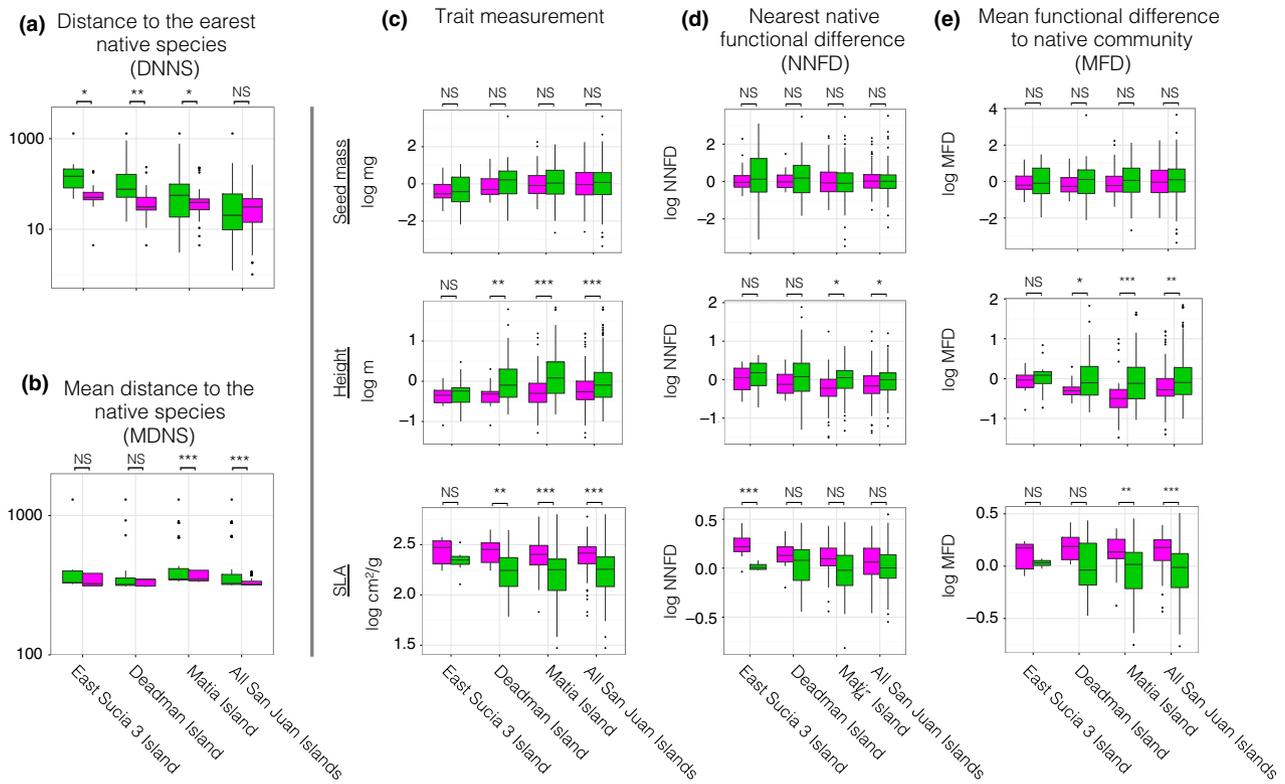
### Functional traits

More than 55% of species overall and more than 60% of the species within most islands had trait measurements for maximum height, seed mass and SLA, while leaf size and leaf nitrogen content had less complete coverage (Figs 2 and S4, Table S3 for a detailed summary). We found significant phylogenetic signal for all traits except leaf size (Table S4),

indicating that the community phylogeny should be a good proxy for ecological similarity.

### Phylogenetic relatedness patterns

Our phylogenetic analyses showed that invasive species have closer native relatives compared to the native species in each island community (Fig. 3a;  $DNNS_i < DNNS_n$ ), and



**Figure 3** Observed phylogenetic and functional trait measures for each native (green) and invasive (pink) species on a representative small (East Sucia 3 Island,  $n = 34$ ,  $60 \text{ m}^2$ ), medium (Deadman Island,  $n = 68$ ,  $13,441 \text{ m}^2$ ) and large island (Matia Island,  $n = 143$ ,  $570,866 \text{ m}^2$ ), and the entire San Juan Island archipelago ( $n = 366$ ,  $2,725,785 \text{ m}^2$ ). (a) Comparison of the distance to the nearest native species (DNNS, log million years). (b) Comparison of the mean phylogenetic distance to the native species (MDNS) between each status group for the three islands and the San Juan Island archipelago. (c) Comparison of trait measurements for each native (green) and invasive (pink) species for three traits with the most data: seed mass, specific leaf area (SLA) and maximum height. (d) Comparison of the nearest native functional difference (NNFD) for the three traits. (e) Comparison of mean functional difference to the all natives in the community (MFD). Boxes show first and third quartiles, solid horizontal lines indicate the median, vertical black lines are the range, and dots are points that lie  $1.5\times$  the interquartile range above the third quartile or below the first quartile. Asterisks indicate significant differences in observed means for each metric between status groups (see Appendix S4 and S5 for  $P$ -values of difference in means).

this difference was significant for 40 (54.05%) of the islands (Figs 4 and S5 for confidence intervals, Appendix S4 for  $P$ -values). At the scale of the whole San Juan Island archipelago the pattern switches, and invasive species are phylogenetically distinct, however, this difference was not significant ( $P = 0.7569$ ). The observed mean  $\text{DNNS}_i$  was less than expected under our null model ( $\text{SES}_{\text{DNNS}} < 0$ ) on 22 (30.56%) of the islands (Figs 5 and S5), indicating that invasive species occurring on these islands are significantly more phylogenetically similar to their nearest native relative than a random assemblage of invasives from the species pool.

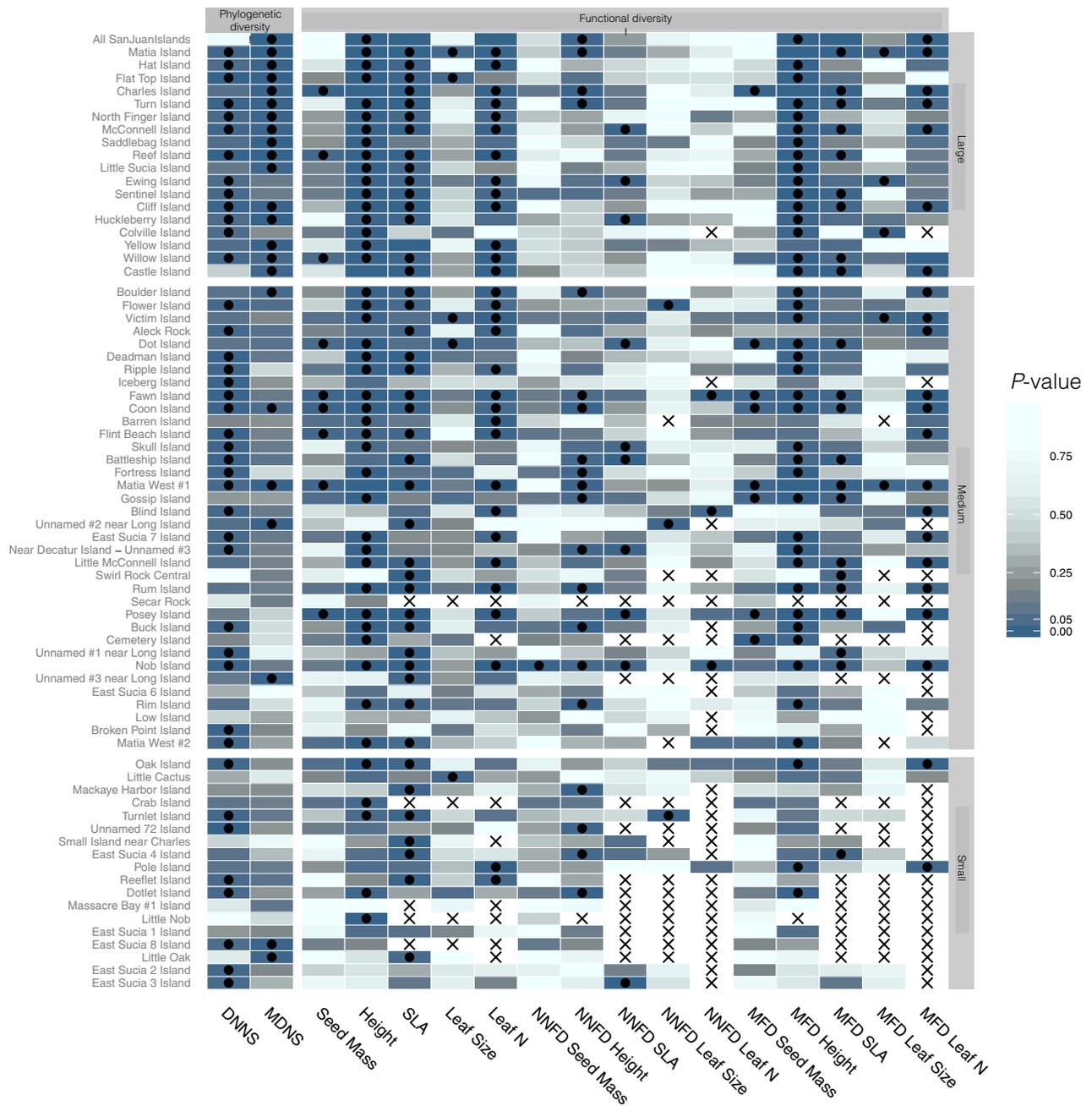
Invasive species were also closely related to the native community as a whole (observed mean  $\text{MDNS}_i < \text{MDNS}_n$  on 21 islands; Fig. 3b), and this was more significant on larger islands (Fig. 4). However, compared to the null model, invasives on 22 islands were significantly distinct from the overall native community (Fig. 5; observed mean  $\text{MDNS}_i > \text{null mean MDNS}_i$ ). On a few of the smaller

islands ( $n = 5$ ), invasives were significantly phylogenetically similar to the native community (observed mean  $\text{MDNS}_i < \text{null mean MDNS}_i$ ), but never on larger islands (Fig. 5).

### Functional trait patterns

Trait measurements showed that invasive species have a shorter maximum height (45 of the islands), a higher SLA (43 of the islands) and lower leaf nitrogen content (31 of the islands) than native species (Figs 3c and S6 for confidence intervals, Appendix S5 for  $P$ -values). Invasives had smaller seeds and leaves than the natives across all islands, but this difference was rarely significant (Fig. 4).

When accounting for phylogenetic distance, we found less significant functional differences between invasives ( $\text{NNFD}_i$ ) and natives ( $\text{NNFD}_n$ ; Figs 3d and 5). The means of  $\text{NNFD}_i$  and  $\text{NNFD}_n$  were most noticeably different for maximum height (20 of the islands) and SLA (10 of the islands).

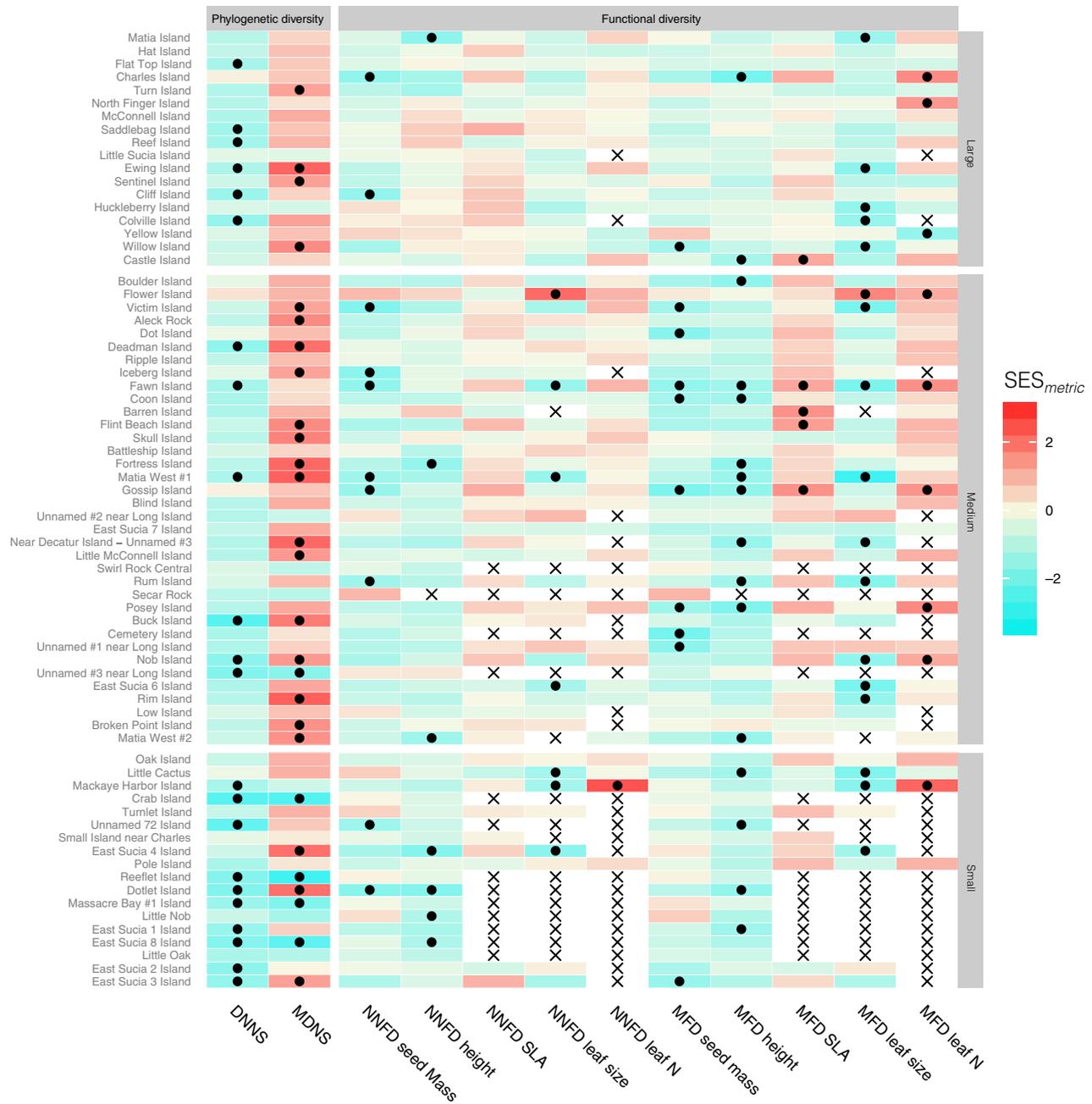


**Figure 4** Significance of difference in means between status groups for phylogenetic and functional diversity indices (x-axis). Darker blue tones indicate decreasing *P*-values from the t-test, and dots indicate significance (*P*-value < 0.05). Islands (y-axis) are ordered by decreasing size from top to bottom, and size categories are noted on the right. Blank cells marked with an 'x' represent islands lacking trait data for the occurring species.

However, compared to any random invasive in the greater San Juan Islands species pool, invasives were functionally similar to their nearest native relatives for seed mass, maximum height and leaf size (on 10, 7 and 6 of the islands, respectively; Fig. 5).

Invasives were also similar to the average trait values of the native community. For seed mass and maximum height,

the mean  $MFD_i$  was significantly less than the mean  $MFD_n$  on 8 and 41 of the islands, respectively (Figs 3e and 4). However, for SLA, mean  $MFD_i$  was significantly greater than mean  $MFD_n$  on 22 islands. Invasives were still more similar compared to the random invasive species pool for seed mass (10 islands) and maximum height (16 islands), but SLA was only significantly different on 5 islands (Fig. 5).



**Figure 5** Significance of standardized effect sizes ( $SES_{metric}$ ) of invasive species for phylogenetic and functional diversity indices ( $x$ -axis). Red hues are  $SES_{metric} > 0$  (overdispersion/distinctiveness), blue hues are  $SES_{metric} < 0$  (clustering/similarity). Darker tones indicate increasingly positive or negative  $z$ -scores, and dots indicate significance ( $P$ -value  $< 0.05$ ). Islands ( $y$ -axis) are ordered by decreasing size from top to bottom, and size categories are noted on the right. Blank cells marked with an 'x' represent islands lacking trait data for the occurring species.

## DISCUSSION

Our analysis of the flora of the San Juan Islands found statistical support that invasive species are phylogenetically similar to their nearest native relative (i.e. non-random phylogenetic clustering). Differences in observed mean  $DNNS_i$  and mean  $DNNS_n$  were significant across a majority of the islands, and compared to the null model of randomly

invaded communities, only support for phylogenetic similarity to the nearest native was found. Less of the islands showed a significant difference between observed mean  $MDNS_i$  and mean  $MDNS_n$ , and we found the opposite relationship when compared to the null model – observed invasives are mostly distinct compared to the native community as a whole (i.e. non-random phylogenetic overdispersion). Thus, the invasive flora of the San Juan Islands

exhibits a nested structure of relatedness, where successful invaders have close native relatives, yet are distributed evenly throughout the native flora, rather than clustered within one lineage.

Given the phylogenetic similarity between invasive plants and their close native relatives, and significant phylogenetic signal for all but one trait in the community phylogeny, we expected invasive species to be functionally similar to the natives. Despite this simple prediction, complex patterns emerged. Observed mean trait measurements were significantly different between the native and invasive species on many islands, at least for a few ecologically relevant traits, which is consistent with other studies (e.g. Pyšek & Richardson, 2007; van Kleunen *et al.*, 2010; Ordóñez, 2014). In general, invasive species were shorter, with a higher SLA, and higher leaf nitrogen content than co-occurring native species, but with similar leaf sizes and seed masses. Functionally, this translates to invasive plants tending to be shorter annuals with broad, nitrogen-rich leaves and may be slightly more easily dispersed than the native species in this system. Invasive species were shorter than the mean of the native species that they co-occur with, but still generally quite tall (mean height of invasives = 1.27 m, mean height of natives = 3.71 m; see Table S3), which might explain their invasion success (Moravcová *et al.*, 2015). In addition, they fall towards the 'fast-return' end (high SLA, high leaf nitrogen content) of the leaf economic spectrum (Wright *et al.*, 2004), which may indicate higher growth and photosynthetic rates compared to native species. Because the trait measurements used were summaries gathered from a variety of other studies and taxonomic treatments, they are fixed traits and do not address the importance of trait plasticity for invasion success. However, in the absence of more detailed measurements they are still useful to provide a broad perspective on functional attributes important for community assembly (Cornwell *et al.*, 2014). Further, trait measurements differed in one direction (i.e. invasive species were always shorter than natives on average) across islands with heterogeneous environments, which is more likely due to broad functional differences between invasive and native species than to plastic ecological responses.

However, when accounting for phylogenetic history along with functional trait comparisons, we saw two different trends. First, the absence of significant differences in observed trait measurements between status groups does not necessarily negate the importance of shared similarity to close native relatives. Invasive and native species had similar measurements for seed mass overall, and there were few significant differences in observed mean NNFD (Fig. 4). But when compared to the randomized invasive community, 10 islands showed significant similarity to their nearest native relative and to the community mean (Fig. 5), illustrating the importance of comparing observed relationships to biologically informed null expectations (see Thuiller *et al.*, 2010 for a discussion of null models relevant to DNC). This functional 'matching' together with overall phylogenetic

similarity to the native community could suggest an environmental filter for seed mass.

Second, when invasive and native species differed in functional trait measurements, invasive species were not necessarily functionally distinct from their nearest native relative. This is especially apparent for SLA, where we found invasives had a significantly greater observed mean SLA on 43 islands, but only 10 were significantly different compared to the nearest native (Fig. 4), and none of these were significant under the null model (Fig. 5). However, they were significantly different from the average of the native community. Therefore, our results suggest that for SLA only the disparity in traits matters for successful invasion, and the identity of the invader relative to its nearest native is irrelevant. This pattern suggests support for functional 'filling' – overall phylogenetic similarity of invasive species to the natives indicates ancestral similarity of most traits is important to pass an environmental filter and establish, but certain traits (i.e. SLA) must still differ to fill available niche spaces, escape competition and persist over time.

Although a phylogeny is often considered an appropriate proxy for ecological similarity, we have shown here that the evolutionary history reflects functional patterns for certain traits (e.g. maximum height), but contrasting patterns emerge (e.g. SLA, functional similarity to the community mean) that can only be distinguished with the addition of functional trait measurements and comparisons to null expectations. Similarly, trait comparisons without consideration of phylogenetic similarity would miss the importance of matching close relatives (e.g. seed mass). Even if overall functional distinctiveness between status groups suggests the importance of diversification of structure, phylogenetic similarity to the nearest native along with functional similarity to the natives for a few traits imply that pre-adapted traits are also important for successful invasion in this system, corroborating the findings of Ordóñez (2014) across 83 communities world-wide.

Phylogenetic similarity and functional distinctiveness could be explained by a few different mechanisms. First, we found that on smaller islands invasives were more often phylogenetically and functionally similar to their nearest native relative and the native community than on larger islands (Fig. 5). Because different ecological and evolutionary processes will likely vary in strength across different scales, the effect of spatial scale on phylogenetic patterns of invasiveness has been previously noted (Cadotte *et al.*, 2009b; Carboni *et al.*, 2012). At small spatial scales (i.e. small islands), it is possible that abiotic pressures, such as tidal flooding, sea spray or nutrient limitation, are creating a stronger effect of environmental or habitat filtering and ancestral traits are important to cope with these extreme conditions. As island size increases, the landscape is more likely to become heterogeneous, increasing the complexity of both abiotic and biotic interactions. Although we treat small and large islands as comparable communities, in actuality larger islands are

composed of many different habitats – for example mixed conifer forests, meadows and beaches. If communities were defined by environmental characteristics such as habitat type (e.g. grassland, rock outcrop or shoreline) instead of geographic boundaries, we may see more homogenous patterns emerge. The environmental heterogeneity and extreme environmental pressures we expect across the islands sampled might also explain why our results differ from some other tests of DNC (e.g. Strauss *et al.*, 2006a; Schaefer *et al.*, 2011; Carboni *et al.*, 2012). Future directions include exploring biogeographic (e.g. island size, distance from the mainland or largest nearest island) and environmental factors to disentangle the ecological processes influencing invasion dynamics in the San Juan floristic community.

Additional information on other intrinsic characters of the invasive species could help to explain the patterns of invasion we described. For example, the native range of the introduced species might better describe niche preferences. Species introduced from Europe might have evolved in a similar Holarctic setting and thus share similar broad lineages (phylogenetic clustering) but fill separate micro-niches and differ in trait space (functional dispersion). It is also possible that other traits, especially characters that were not investigated here (e.g. root structures, habit, phenology), might have different functional patterns.

Because the timing of introductions is unknown, we cannot exclude the possibility of founder effects and residence time as contributors to these patterns as well. A recent analysis of a long-term plant survey by Li *et al.* (2015) provides a mechanistic explanation to support both sides of Darwin's conundrum as naturalization progresses over time. Not only were exotic species with close native relatives more likely to naturalize and dominate the native community, but natives most closely related to exotics more likely to be replaced (Li *et al.*, 2015). Functional differences for a few traits could signify that displacement is beginning in the San Juan Islands, and additional surveys to investigate this would be an interesting development.

The era of 'big data' is contributing immensely to publicly available molecular information, and it is tempting to use phylogenetic approaches to explain patterns of community assembly and predict future invasions in the absence of additional phenotypic data for each species. However, we show that trait databases and ecological surveys provide a resource to easily unite function and phylogeny and test eco-evolutionary hypotheses of species invasion, and their complementarity can aid in deconstructing the processes of invasion in natural systems. In particular, clearly defined null models that evaluate the phylogenetic and functional structure of communities enable identification of patterns that could not be generated randomly, and branch lengths that are consistently time-scaled allows for comparison across different systems. Incorporating the growing body of diversity metrics with meaningful null hypotheses and unified methodologies into comparable frameworks has the potential to

further deconstruct the patterns of assembly we observe and predict future ecological change.

## ACKNOWLEDGEMENTS

We would like to thank Marcel Rejmánek and two anonymous referees for helpful comments that improved this study. We also thank multiple contributors of the trait data used in this study: Angela Moles, Dana Royer, Alejandro Ordóñez, Ian Wright, Peter Reich, Amy Zanne and Will Cornwell. The U.S. Fish and Wildlife Service provided financial support for fieldwork, The Nature Conservancy provided logistical support and the U.S. Bureau of Land Management and Washington State Parks and Recreation Commission allowed access to their lands for inclusion in the study. Luke Harmon and Matthew Pennell provided advice on statistics and programming, and Cody Hinchliff helped with the PHLAWD pipeline. This material is based upon work supported by a NSF Graduate Research Fellowship Grant No. DGE-1144254 and a University of Idaho Bioinformatics and Computational Biology Graduate Student Fellowship to HEM. Finally, we thank the University of Idaho Institute for Bioinformatics and Evolutionary Studies Computational Resources core, supported in part by NIH/NCRR grants P20RR16448 and P20RR016454.

## REFERENCES

- Abouheif, E. (1999) A method for testing the assumption of phylogenetic independence in comparative data. *Evolutionary Ecology Research*, **1**, 895–909.
- Bellard, C., Thuiller, W., Leroy, B., Genovesi, P., Bakkenes, M. & Courchamp, F. (2013) Will climate change promote future invasions? *Global Change Biology*, **19**, 3740–3748.
- Bezeng, B.S., Savolainen, V., Yessoufou, K., Papadopoulos, A.S., Maurin, O. & Bank, M. (2013) A phylogenetic approach towards understanding the drivers of plant invasiveness on Robben Island, South Africa. *Botanical Journal of the Linnean Society*, **172**, 142–152.
- Cadotte, M.W. (2014) Including distantly related taxa can bias phylogenetic tests. *Proceedings of the National Academy of Sciences USA*, **111**, E536–E536.
- Cadotte, M.W., Murray, B.R. & Lovett-Doust, J. (2006) Ecological patterns and biological invasions: using regional species inventories in macroecology. *Biological Invasions*, **8**, 809–821.
- Cadotte, M.W., Cavender-Bares, J., Tilman, D. & Oakley, T.H. (2009a) Using phylogenetic, functional and trait diversity to understand patterns of plant community productivity. *PLoS ONE*, **4**, e5695.
- Cadotte, M.W., Hamilton, M.A. & Murray, B.R. (2009b) Phylogenetic relatedness and plant invader success across two spatial scales. *Diversity and Distributions*, **15**, 481–488.
- Cadotte, M.W., Borer, E.T., Seabloom, E.W., Cavender-Bares, J., Harpole, W.S., Cleland, E. & Davies, K.F. (2010) Phylogenetic patterns differ for native and exotic plant

- communities across a richness gradient in Northern California. *Diversity and Distributions*, **16**, 892–901.
- Cadotte, M.W. & Strauss, S.Y. (2011) Phylogenetic Patterns of Colonization and Extinction in Experimentally Assembled Plant Communities. *PLoS ONE*, **6**, e19363.
- Cadotte, M., Albert, C.H. & Walker, S.C. (2013) The ecology of differences: assessing community assembly with trait and evolutionary distances. *Ecology Letters*, **16**, 1234–1244.
- Carboni, M., Munkemüller, T., Gallien, L., Lavergne, S., Acosta, A. & Thuiller, W. (2012) Darwin's naturalization hypothesis: scale matters in coastal plant communities. *Ecography*, **35**, 1–9.
- Castro, S.A., Escobedo, V.M., Aranda, J. & Carvallo, G.O. (2014) Evaluating darwin's naturalization hypothesis in experimental plant assemblages: phylogenetic relationships do not determine colonization success. *PLoS ONE*, **9**, e105535.
- Cavender-Bares, J., Ackerly, D.D., Baum, D.A. & Bazzaz, F.A. (2004) Phylogenetic overdispersion in Floridian oak communities. *The American Naturalist*, **163**, 823–843.
- Cornwell, W.K., Westoby, M., Falster, D.S. *et al.* (2014) Functional distinctiveness of major plant lineages. *Journal of Ecology*, **102**, 345–356.
- Daehler, C.C. (1998) The taxonomic distribution of invasive angiosperm plants: ecological insights and comparison to agricultural weeds. *Biological Conservation*, **84**, 167–180.
- Daehler, C.C. (2001) Darwin's Naturalization Hypothesis Revisited. *The American Naturalist*, **158**, 324–330.
- Darwin, C. (1859) *On the origins of species by means of natural selection*. John Murray, London.
- Davies, K.F., Cavender-Bares, J. & Deacon, N. (2010) Native communities determine the identity of exotic invaders even at scales at which communities are unsaturated. *Diversity and Distributions*, **17**, 35–42.
- Diez, J.M., Sullivan, J.J., Hulme, P.E., Edwards, G. & Duncan, R.P. (2008) Darwin's naturalization conundrum: dissecting taxonomic patterns of species invasions. *Ecology Letters*, **11**, 674–681.
- Diez, J.M., Williams, P.A., Randall, R.P., Sullivan, J.J., Hulme, P.E. & Duncan, R.P. (2009) Learning from failures: testing broad taxonomic hypotheses about plant naturalization. *Ecology Letters*, **12**, 1174–1183.
- Eastman, J.M., Harmon, L.J. & Tank, D.C. (2013) Congruification: support for time scaling large phylogenetic trees. *Methods in Ecology and Evolution*, **4**, 688–691.
- Floristic Atlas of the San Juan Islands, Washington (2013). Available at: <http://biology.burke.washington.edu/herbarium/resources/sanjuanatlas.php> (accessed 29 July 2013).
- Godoy, O., Kraft, N.J.B. & Levine, J.M. (2014) Phylogenetic relatedness and the determinants of competitive outcomes. *Ecology Letters*, **17**, 836–844.
- Grime, J.P. (1974) Vegetation classification by reference to strategies. *Nature*, **250**, 26–31.
- Jiang, L., Tan, J. & Pu, Z. (2010) An Experimental Test of Darwin's Naturalization Hypothesis. *The American Naturalist*, **175**, 415–423.
- Jones, E.I., Nuismer, S.L. & Gomulkiewicz, R. (2013) Revisiting Darwin's conundrum reveals a twist on the relationship between phylogenetic distance and invasibility. *Proceedings of the National Academy of Sciences USA*, **110**, 20627–20632.
- Kattge, J., Diaz, S., Lavorel, S., Prentice, I.C., Leadley, P., Bönsch, G., Garnier, E., Westoby, M., Reich, P.B. & Wright, I.J. (2011) TRY—a global database of plant traits. *Global Change Biology*, **17**, 2905–2935.
- Kembel, S.W., Cowan, P.D., Helmus, M.R., Cornwell, W.K., Morlon, H., Ackerly, D.D., Blomberg, S.P. & Webb, C.O. (2010) Picante: R tools for integrating phylogenies and ecology. *Bioinformatics*, **26**, 1463–1464.
- van Kleunen, M., Weber, E. & Fischer, M. (2010) A meta-analysis of trait differences between invasive and non-invasive plant species. *Ecology Letters*, **13**, 235–245.
- Li, S.-p., Cadotte, M.W., Meiners, S.J., Hua, Z.-s., Shu, H.-y., Li, J.-t. & Shu, W.-s. (2015) The effects of phylogenetic relatedness on invasion success and impact: deconstructing Darwin's naturalisation conundrum. *Ecology Letters*, **18**, 1285–1292.
- Lim, J., Crawley, M.J., De Vere, N., Rich, T. & Savolainen, V. (2014) A phylogenetic analysis of the British flora sheds light on the evolutionary and ecological factors driving plant invasions. *Ecology and Evolution*, **4**, 4258–4269.
- Lososová, Z., De Bello, F., Chytrý, M., Kühn, I., Pyšek, P., Sádlo, J., Winter, M. & Zelený, D. (2015) Alien plants invade more phylogenetically clustered community types and cause even stronger clustering. *Global Ecology and Biogeography*, **24**, 786–794.
- Mayfield, M.M. & Levine, J.M. (2010) Opposing effects of competitive exclusion on the phylogenetic structure of communities. *Ecology Letters*, **13**, 1085–1093.
- Moravcová, L., Pyšek, P., Jarošík, V. & Pergl, J. (2015) Getting the right traits: reproductive and dispersal characteristics predict the invasiveness of herbaceous plant species. *PLoS ONE*, **10**, e0123634.
- Mouquet, N., Devictor, V., Meynard, C.N. *et al.* (2012) Eco-phylogenetics: advances and perspectives. *Biological Reviews*, **87**, 769–785.
- Ordóñez, A. (2014) Functional and phylogenetic similarity of alien plants to co-occurring natives. *Ecology*, **95**, 1191–1202.
- Park, D.S. & Potter, D. (2013) A test of Darwin's naturalization hypothesis in the thistle tribe shows that close relatives make bad neighbors. *Proceedings of the National Academy of Sciences USA*, **110**, 17915–17920.
- Pavoine, S. & Bonsall, M.B. (2010) Measuring biodiversity to explain community assembly: a unified approach. *Biological Reviews*, **86**, 792–812.
- Procheş, Ş., Wilson, J.R.U., Richardson, D.M. & Rejmánek, M. (2008) Searching for phylogenetic pattern in biological invasions. *Global Ecology and Biogeography*, **17**, 5–10.
- Pyšek, P. & Richardson, D.M. (2006) The biogeography of naturalization in alien plants. *Journal of Biogeography*, **33**, 2040–2050.

- Pyšek, P. & Richardson, D.M. (2007) Traits associated with invasiveness in alien plants: where do we stand? *Biological Invasions* (ed. by W. Nentwig), pp. 97–125. Springer, Berlin.
- R Development Core Team (2014) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Available at: <http://www.R-project.org>.
- Reich, P.B., Wright, I.J., Cavender-Bares, J., Craine, J.M., Oleksyn, J., Westoby, M. & Walters, M.B. (2003) The evolution of plant functional variation: traits, spectra, and strategies. *International Journal of Plant Sciences*, **164**, S143–S164.
- Rejmánek, M. (1996) A theory of seed plant invasiveness: the first sketch. *Biological Conservation*, **78**, 171–181.
- Richardson, D.M. & Pyšek, P. (2006) Plant invasions: merging the concepts of species invasiveness and community invasibility. *Progress in Physical Geography*, **30**, 409–431.
- Richardson, D.M., Pyšek, P., Rejmánek, M., Barbour, M.G., Panetta, F.D. & West, C.J. (2000) Naturalization and invasion of alien plants: concepts and definitions. *Diversity and Distributions*, **6**, 93–107.
- Roquet, C., Thuiller, W. & Lavergne, S. (2012) Building megaphylogenies for macroecology: taking up the challenge. *Ecography*, **35**, 1–14.
- Sala, O.E., Chapin, F.S., Armesto, J.J., Berlow, E., Bloomfield, J., Dirzo, R., Huber-Sanwald, E., Huenneke, L.F., Jackson, R.B., Kinzig, A., Leemans, R., Lodge, D.M., Mooney, H.A., Oesterheld, M., Poff, N.L., Sykes, M.T., Walker, B.H., Walker, M. & Wall, D.H. (2000) Global biodiversity scenarios for the year 2100. *Science*, **287**, 1770–1774.
- Sax, D.F. & Gaines, S.D. (2008) Species invasions and extinction: the future of native biodiversity on islands. *Proceedings of the National Academy of Sciences USA*, **105**, 11490–11497.
- Schaefer, H., Hardy, O.J., Silva, L., Barraclough, T.G. & Savolainen, V. (2011) Testing Darwin's naturalization hypothesis in the Azores. *Ecology Letters*, **14**, 389–396.
- Smith, S.A., Beaulieu, J.M. & Donoghue, M.J. (2009) Mega-phylogeny approach for comparative biology: an alternative to supertree and supermatrix approaches. *BMC Evolutionary Biology*, **9**, 37.
- Sol, D., Lapidra, O. & Vila, M. (2014) Do close relatives make bad neighbors? *Proceedings of the National Academy of Sciences USA*, **111**, E534–E535.
- Soltis, D.E., Smith, S.A., Cellinese, N. *et al.* (2011) Angiosperm phylogeny: 17 genes, 640 taxa. *American Journal of Botany*, **98**, 704–730.
- Stamatakis, A. (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, **22**, 2688–2690.
- Strauss, S.Y., Lau, J.A. & Carroll, S.P. (2006a) Evolutionary responses of natives to introduced species: what do introductions tell us about natural communities? *Ecology Letters*, **9**, 357–374.
- Strauss, S.Y., Webb, C.O. & Salamin, N. (2006b) Exotic taxa less related to native species are more invasive. *Proceedings of the National Academy of Sciences USA*, **103**, 5841–5845.
- Thuiller, W., Gallien, L., Boulangéat, I., de Bello, F., Münkemüller, T., Roquet, C. & Lavergne, S. (2010) Resolving Darwin's naturalization conundrum: a quest for evidence. *Diversity and Distributions*, **16**, 461–475.
- Tilman, D. (2004) Niche tradeoffs, neutrality, and community structure: a stochastic theory of resource competition, invasion, and community assembly. *Proceedings of the National Academy of Sciences USA*, **101**, 10854–10861.
- Vermeij, G.J. (1996) An agenda for invasion biology. *Biological Conservation*, **78**, 3–9.
- Webb, C. (2000) Exploring the phylogenetic structure of ecological communities: an example for rain forest trees. *The American Naturalist*, **156**, 145–155.
- Webb, C.O., Ackerly, D.D., McPeck, M.A. & Donoghue, M.J. (2002) Phylogenies and community ecology. *Annual Review of Ecology and Systematics*, **33**, 475–505.
- Westoby, M. (1998) A leaf-height-seed (LHS) plant ecology strategy scheme. *Plant and Soil*, **199**, 213–227.
- 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.
- Wright, I.J., Reich, P.B., Westoby, M. *et al.* (2004) The worldwide leaf economics spectrum. *Nature*, **428**, 821–827.

## DATA ACCESSIBILITY

Species list with status, and trait measurements for each of the five traits, metadata for each island (total island area, assigned size category), the community matrix with presence/absence of each species across the 80 islands, sequence alignments for each gene region with GenBank accession numbers for each species, the concatenated sequence alignment, treefiles of maximum-likelihood phylograms for each gene region and the concatenated alignment, and a treefile of the maximum-likelihood mega-phylogeny (scaled to time) are available from the Dryad Digital Repository: <http://dx.doi.org/10.5061/dryad.m88g7>. Custom R scripts for all analyses and associated datasets are available on Github (<https://github.com/hmarx/San-Juan-Islands-Invasion>).

## SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

**Figure S1** Maximum likelihood phylogram for each gene tree.

**Figure S2** Maximum likelihood mega-phylogeny of the vascular flora of the San Juan Islands, with species names labelled. Nodes and tips as in Fig. 2.

**Figure S3** Maximum likelihood phylogram for the concatenated dataset.

**Figure S4** Plots of species richness and per cent of species with data for each functional trait per island.

**Figure S5** Plots of phylogenetic diversity measures (DNNS, MDNS,  $SES_{DNNS}$ ,  $SES_{MDNS}$ ) for native and invasive plant species across all islands.

**Figure S6** Plots of functional diversity measures (trait measurements, NNFD, MFD,  $SES_{NNFD}$ ,  $SES_{MFD}$ ) for native and invasive plant species across all islands.

**Table S1** Review of previous studies testing Darwin's Naturalization Conundrum with phylogenetic diversity and functional traits.

**Table S2** Table with the number of native and invasive species retrieved for each gene region, the final sequence length of each gene region, and per cent of each gene partition in the final concatenated nucleotide matrix post cleaning.

**Table S3** Table summarizing the number of species with data for each trait, and the mean, median, minimum, and maximum trait measurements for each trait.

**Table S4** Results of phylogenetic signal analysis for the five functional traits.

**Appendix S1** Detailed methods for phylogeny estimation, functional trait measurements for each species, and sources used for each trait.

**Appendix S2** Metadata for each island (total island area, assigned size category), summary of species diversity observed on each island, and the per cent of species with sequence data retrieved from GenBank.

**Appendix S3** Community matrix with presence/absence of each species (rows) across the 80 islands (columns) in the San Juan archipelago, WA.

**Appendix S4** *P*-values for difference in observed phylogenetic distance (DNNS, MDNS) between status groups for each island.

**Appendix S5** *P*-values for difference in observed functional differences (measurements, NNFD, MFD) between status groups for each island.

## BIOSKETCH

**Hannah E. Marx** is a PhD candidate in the Bioinformatics and Computational Biology graduate programme at the University of Idaho under the supervision of DCT. She is interested in phylogenetic approaches to understanding processes driving diversity dynamics of communities on islands.

**David C. Tank** is an Associate Professor in the Department of Biological Sciences at the University of Idaho, and Director of the Stillinger Herbarium. His research programme focuses on the use of molecular methods to reconstruct evolutionary histories in plants and the application of phylogenetic and phylogeographic methods to understand plant evolution across multiple levels of plant phylogeny – from the systematics and evolution of closely related groups of species, to macroevolutionary patterns of diversification in angiosperms.

**Peter W. Dunwiddie** is an ecologist focusing on rare species recovery and management of conservation lands. He works with the Center for Natural Lands Management and is an affiliate professor in the Biology Department and the School of Environmental and Forest Sciences at the University of Washington.

**David E. Giblin** is Collections Manager at the University of Washington Herbarium at the Burke Museum, where he studies the vascular plant flora of the Pacific Northwest.

Author contributions: P.W.D. and D.E.G. conducted the floristic survey of the San Juan Islands, with some assistance from H.E.M. H.E.M. and D.C.T. conceptualized the study and analysed output. H.E.M. performed phylogenetic and statistical analyses. H.E.M. wrote the first draft of the manuscript, and all authors contributed to revisions.

---

Editor: Marcel Rejmanek