Seed morphology and emergence variability in a conservation collection of *Pinus torreyana*

Lionel N Di Santo[1,4], Monica Polgar[1], Storm Nies[1], Paul Hodgkiss[2], Courtney A Canning,[3] Jessica W Wright[2], and Jill A Hamilton[1]

[1] North Dakota State University, Department of Biological Sciences, Fargo, ND, USA

[2] USDA Forest Service, Pacific Southwest Research Station, Davis, CA, USA

[3] USDA Forest Service, Pacific Southwest Research Station, Placerville CA, USA

[4] Correspondence: lionel.disanto@ndsu.edu

ABSTRACT

PREMISE: Optimizing the amount of genetic diversity captured in seed collections is a long-standing objective of ex situ conservation. Particularly for rare species where limited genetic information is available, it poses a significant challenge. However, understanding the within and among population distribution of trait variation within seed collections may provide a means to approximate standing genetic variation and inform conservation efforts.

METHODS: We quantified seed morphological variation and seedling emergence both within and among populations for existing seed collections of Torrey pine (*Pinus torreyana*), a critically endangered pine endemic to California. Using a simulation-based approach, we used estimates of within-population variance to estimate the number of maternal families required to captured 95% of trait variation within each existing seed collection.

KEY RESULTS: On average, 21% and 22% of seed morphological variation in Torrey pine was explained by population origin and maternal family within populations respectively. Despite significant morphological differences, seedling emergence was similar across populations. Simulations revealed that 80% and 68% of all maternal families within island and mainland seed collection respectively needed to be resampled to ensure 95% of seed trait variation within existing collections was captured.

CONCLUSIONS: Overall, these results suggest that substantial structure exists for seed morphological traits both within and between populations. From a conservation perspective, this indicates that to optimize genetic diversity captured in Torrey pine *ex situ* conservation collections, maximizing the number of maternal families sampled within each population will be necessary.

KEY WORDS: ex situ conservation; island-mainland phenotypic differentiation; seed trait variation; Torrey pine; variance partitioning.

INTRODUCTION

Ex situ seed collections preserve species genetic diversity outside of their native range, providing the raw material for species reintroductions and germplasm to augment restoration (Guerrant Jr et al., 2013; Potter et al., 2017). An invaluable conservation resource, particularly for rare species, ex situ collections protect against biodiversity loss in the wild while preserving species evolutionary potential. Ensuring ex situ collections represent genetic variation found in natural populations is critical to both contemporary conservation and potential future restoration efforts (Schaal and Leverich, 2004; Basey et al., 2015). However, the cost and logistical constraints associated with seed collection pose a significant challenge. Given this challenge, identifying means to optimize ex situ sampling efforts are needed (Hoban and Schlarbaum, 2014; Di Santo and Hamilton, 2020). One approach may be to use the distribution of trait variation existing within contemporary ex situ collections as a proxy for quantifying standing genetic variation within and among populations. These data may then be used to optimize and target supplemental ex situ efforts. In this way, quantifying within and between population variation for traits largely considered heritable within existing ex situ population collections may be invaluable to optimizing future collections.

The distribution of heritable genetic variation estimated via common garden experiments (Weber and Kolb, 2014; Hamilton et al., 2017; Yoko et al., 2020) or genetic data (Zhang and Zhou, 2013; Hausman et al., 2014; Tamaki et al., 2018) can be used to quantify the distribution of

standing genetic variation. However, when common garden experiments or genetic data are unavailable, quantifying trait variation within and among ex situ seed population collections may provide a reasonable proxy for the distribution of genetic variation. Seed collections established to preserve genetic variation, as well as to provide restoration or research material are numerous (Guerrant Jr et al., 2013; Westengen et al., 2013; Etterson et al., 2016). Millions of seed accessions of known provenance have been stored in gene banks internationally (FAO, 2010), representing a large conservation and research resource. Although common garden experiments are preferred when available, heritability of seed morphological traits and ease of access to seeds through existing ex situ collections suggests assessment of morphological trait variation may provide a reasonable potential proxy to quantifying the distribution of standing genetic variation.

Seed traits, such as seed size (H^2 = 0.47-0.99, h^2 = 0.19-0.56), relative embryo size (H^2 = 0.89-0.99), seed shape (h^2 = 0.59-0.79), seed coat thickness (H^2 = 0.59), and embryo length (H^2 = 0.92-0.96) are often highly heritable (Pandey et al., 1994; Cober et al., 1997; Mera et al., 2004; Hakim and Suyamto, 2017). Given these observations, variation in seed morphological traits likely has a genetic basis, and may reflect standing genetic variation responding to natural selection (Barrett and Schluter, 2008; Carlson et al., 2014). Particularly for rare species where existing genetic data is often limited, the distribution of trait variation within existing seed collections may provide a valuable resource, even if estimates of variation may not necessarily be conservative. For these species, where the goal is to limit the loss of biodiversity while preserving evolutionary potential, quantifying the distribution of morphological variation within and among seed populations provides a timely approach to evaluating variation preserved in collections. Leveraging existing seed morphological data from conservation collections may aid

in understanding the distribution of variability currently maintained ex situ, and may be used to optimize supplemental conservation collections where needed.

Pinus torreyana Parry ex Carr. (Torrey pine), is one of the rarest pines in the world (Critchfield and Little, 1966; Dusek, 1985), endemic to two discrete natural populations in California. Torrey pine occupies one mainland population of approximately 6000 trees at the Torrey Pine State Reserve in La Jolla, CA, and an island population of approximately 3000 reproductively mature trees on Santa Rosa Island, CA one of the Channel Islands (Ledig and Conkle, 1983; Hamilton et al., 2017) (Fig. 1). Listed as critically endangered by the IUCN (2020), Torrey pine is of critical conservation concern due to multiple factors; including low population size (Franklin and Santos, 2011; Hall and Brinkman, 2015), low genetic diversity (Ledig and Conkle, 1983; Whittall et al., 2010), climate change, and environmentally and human-mediated disturbances (Franklin and Santos, 2011; Hamilton et al., 2017). While in situ conservation has preserved the whole of the species' range, with fewer than 10,000 reproductively mature individuals in native populations, there are undeniable risks to population-level extirpation events. To mitigate potential losses in the wild, conservation efforts have focused on preservation of seed ex situ. While ex situ seed collections provide an invaluable conservation resource, they may be also used to understand within and among population variation to ultimately inform future conservation efforts.

In this study, we evaluate morphological trait variation in a large ex situ conservation collection of Torrey pine seed sourced from the two native extant populations. Specifically, we quantify the distribution of variation for 14 seed morphology traits and assess differences in emergence

between island and mainland seedlings. In addition, we use existing ex situ collection data to provide supplemental population sampling guidance for future Torrey pine collections. For this latter objective, we use simulations to estimate the number of maternal families required to capture 95% of seed morphological variation existing in contemporary ex situ collections, for both island and mainland population independently. This study provides baseline understanding of population- and family-level variability for an extremely rare endemic pine, and uses these data to inform population sampling and aid in optimizing conservation efforts and preservation of evolutionary potential in this species.

MATERIAL AND METHODS

Cone collection and seed processing – Reproductively mature Torrey pine (*Pinus torreyana* Parry ex. Carr) cones were collected as part of a large ex situ conservation collection between June and July of 2017. Cones were collected from 157 trees on Santa Rosa Island (Channel Islands National park), CA (island population) and 201 trees at Torrey Pine State Reserve in La Jolla, CA (mainland population), representing the species entire native distribution (Fig. 1). Where possible, we collected between five to ten cones per maternal tree at each location. Seeds, organized by individual maternal tree, were then extracted from cones using a combination of mallet and pliers and processed for inclusion in a long-term ex situ conservation collection (see below).

Seed viability tests – A float test was used first to quickly assess seed viability within the ex situ collection. Floating seeds were presumed to lack an endosperm or embryo, while seeds that sunk were presumed filled. Seeds were dropped into water for a period of approximately 15 seconds to

differentiate presumed non-viable, floating seed from presumed viable, sinking seeds (Gribko and Jones, 1995; Morina et al., 2017). Those seeds classed as likely viable were organized by maternal tree using paper bags, and then placed in a Blue M drying oven (Thermal Product Solutions, White Deer, Pennsylvania, USA) maintained at 37°C for 24 hours to remove potential surface moisture. Following this, x-ray photographs were taken to visually assess viability and morphological variation of each seed and its contents at the Placerville Nursery, CA. Acrylic seed trays [20.3 cm x 25.4 cm x 0.48 cm], with a 9 x 11 array of wells, were used to separate and position each Torrey pine seed over the x-ray film. Kodak x-OMAT HBT film (20.3 cm x 25.4 cm) was placed in a lightproof x-ray film cassette which was positioned in the x-ray machine with the seed tray centered on top of the film, with a shelf height of 55.9 cm. The x-ray was taken using a 17 kVP exposure for a total of two minutes, based on standardized conditions established previously for *Pinus coulteri* (Sara Wilson, USDA Forest Service, personal communication). X-ray images were digitized using a Nikon D40 digital camera mounted on a tripod over a light box. Because size standards were internal to each image (see below), no additional calibrations or corrections were required.

Morphological measurement of seed traits – Using ImageJ (Abràmoff et al., 2004), eight seed morphological traits were measured across 80 mainland maternal families and 30 island maternal families, representing a haphazard subset of the complete collection (Fig. 2; Table 1). Each x-ray picture was scaled using the diameter of a seed tray well (1.87 cm) to express pixels as trait values in centimeters. Directly measured seed traits included seed length (SL, cm), seed width (SW, cm), embryo length (EL, cm), embryo width (EW, cm), seed coat width (SCW, cm), seed area (SA, cm^2), endosperm area (ESA, cm^2), and embryo area (EA, cm^2). Using these traits, six

additional traits were derived (Table 1); including seed length/width ratio (SLW), embryo length/width ratio (ELW), relative embryo size (RES), relative endosperm size (REndS), seed coat area (SCA, cm^2), and relative seed coat size (RSCS). We measured five randomly selected seeds per maternal tree, including three technical replications per seed for each trait, with the exception of seven mainland (TPSR 442: 30 seeds, TPSR 457: 20 seeds, TPSR 477: 15 seeds, TPSR 1176: 23 seeds, TPSR 3005: 20 seeds, TPSR 3202: 10 seeds, and TPSR 3204: 10 seeds) and seven island maternal trees (SRI 143: 10 seeds, SRI 243: 10 seeds, SRI 1088: 10 seeds, SRI 1136: 10 seeds, SRI 2028: 9 seeds, SRI 2048: 10 seeds, and SRI 2124: 10 seeds). Measurements were averaged across technical replicates to summarize the mean trait value per seed. In total, 677 seeds were measured from across 110 maternal trees spanning the two Torrey pine populations.

Seed emergence test – In January 2018, a seed emergence trial was conducted using a random subset of seeds from the ex situ collection, including seeds sourced from Torrey Pine State Reserve and Santa Rosa Island, CA. Following x-ray, seeds were stored at 4°C in sealed mylar bags (USA emergency supply, Beaverton, Oregon, USA) placed in plastic boxes, each box contained desiccant crystals to decrease ambient moisture and reduce likelihood of mold. Seeds from eight maternal families per population were selected for the emergence trial. Between eight to ten seeds per maternal tree were weighed and then stratified under cold, moist conditions for 30 days (placed in plastic boxes on a moist paper towel at 4°C). Seeds were sown directly into a 164 ml Ray Leach “Cone-tainer” ™ (Stuewe & Sons, Tangent, Oregon, USA) filled with Sunshine® Mix #4 (Sungro horticulture, Agawam, Massachusetts, USA), pressed halfway into the soil, and then covered with a thin layer of gravel. For approximately one month following

planting, seeds were misted for one minute at hourly intervals over a daily eight-hour period (9am – 4pm). Following emergence, seedlings were hand watered to saturation weekly to biweekly as needed. Emergence was quantified across three separate timepoints (02/06/2018, 02/16/2018 and 03/07/2018) per maternal family as the proportion of seeds that had emerged from the total initially planted.

Evaluating the distribution of seed trait variation – We conducted a principal component analysis (PCA) using all 14 measured and derived seed traits averaged by maternal family to evaluate population-specific differentiation in seed morphology. Prior to performing the PCA, to account for differences in measurement units, all seed traits were standardized using the *scale()* function in R implementing the z-score standardization: $z_{ij} = \frac{x_{ij} - \mu_j}{\sigma_j}$ (1), where x_{ij} is the non-transformed trait value, μ_j is the mean of a given seed trait across populations, and σ_j the standard deviation of the same seed trait across populations. Subsequently, to test for seed trait differences between population means, we used either Student's two-sample test or its non-parametric equivalent when normality was not met, Wilcoxon's two-sample test, within the R package "exactRankTests" (Hothorn and Hornik, 2019). Normality was assessed using Shapiro-Wilk's test of normality within each population. In total, five of the fourteen traits were distributed normally in both Torrey pine populations, including seed width (mainland: W = 0.97, P = 0.11; island: W = 0.97, P = 0.57), embryo length (mainland: W = 0.98, P = 0.28; island: W = 0.96, P = 0.37), embryo area (mainland: W = 0.98, P = 0.46; island: W = 0.97, P = 0.59), seed length/seed width ratio (mainland: W = 0.99, P = 0.5; island: W = 0.98, P = 0.75), and embryo length/embryo width ratio (mainland: W = 0.97, P = 0.05; island: W = 0.97, P = 0.67).

To evaluate the distribution of morphological trait variation within and between Torrey pine populations, we quantified the proportion of variation attributed to population and maternal tree families using measured and derived morphological traits summarized by seeds. For each trait, we fit a linear mixed model using the R package "lme4" (Bates et al., 2014) with population considered a fixed effect and maternal families within populations considered a nested random effect: $Y_{ij} = \mu + \pi_i + r_{i/j} + e_{ij}$, where Y_{ij} is the observed seed trait value, μ is the seed trait overall mean, π_i is the effect of population origin on the seed trait mean, $r_{i/j}$ is the effect of maternal family within populations on the observed seed trait value, and e_{ij} are the effects on the seed trait value of any other variables unaccounted for in the model (residual error). For each model, normality of residual errors was visually assessed and significance of fixed- and random-effect terms was tested using the functions *anova()* and *ranova()* respectively, implemented in the R package "stats" (R Core Team 2020) and "lmerTest" (Kuznetsova et al., 2016). Proportions of seed morphological variance explained by populations (marginal R^2, R^2_m), both populations and maternal families (conditional R^2, R^2_c), and maternal families alone ($R^2_c - R^2_m$) were quantified for each model independently using the function *r.squaredGLMM()* implemented in the R package "MuMIn" (Bartoń, 2020).

Assessing differences in seedling emergence across populations – To test for differences in the probability and the timing of seed emergence in Torrey pine, we evaluated the proportion of seedling emergence both within and between populations across timepoints. First, we used Friedman's rank sum test (non-parametric repeated measures ANOVA) followed by Wilcoxon' paired two-sample test, both implemented in the R package "rstatix" (Kassambara, 2020), to assess differences in the proportion of emerged seeds between timepoints within populations. We

used a non-parametric approach for both Torrey pine populations because normality could not be assumed at select timepoints due to high frequency of zero values. We accounted for multiple testing using Benjamini & Hochberg (1995) False Discovery Rate (FDR) correction implemented in the *wilcox_test()* function. Following this, we evaluated timepoint-specific population differences in emergence. We used Shapiro-Wilk's test to assess populations' deviation from normality at each timepoint and either Student's (for timepoints passing the normality test) or Wilcoxon's two-sample test (for timepoints failing the normality test) to evaluate differences in population emergence. Timepoints 02/16/18 (mainland: $W = 0.94$, $P = 0.57$; island: $W = 0.9$, $P = 0.28$) and 03/07/2018 (mainland: $W = 0.96$, $P = 0.85$; island: $W = 0.88$, $P = 0.21$) passed the normality test, while timepoint 02/06/2018 (mainland: $W = 0.52$, $P < 0.001$; island: $W = 0.73$, $P = 0.004$) failed the normality test. All statistical analyses were performed using R version 4.0.2 (R Core Team 2020).

***Simulating variation captured in* ex situ *collection using seed morphological traits* –** For each of the 14 measured and derived seed traits, we conducted a separate simulation quantifying morphological variation captured when increasing the number of maternal families sampled from contemporary Torrey pine seed collections. Simulations were conducted in R using a customized script (Appendix S3). Resampling of ex situ collections were performed for island and mainland Torrey pine populations independently, using between one and the total number of maternal families available within each ex situ population collection (mainland: 80 maternal families, island: 30 maternal families) (N_{fam}). Maternal trees were sampled randomly without replacement from the pool of available families. All seeds within each selected maternal family were sampled as part of this simulation, except those with missing values for the trait simulated.

To evaluate the number of maternal families needed to capture 95% of seed trait variation in both island and mainland populations, we estimated the number of unique seed trait values captured in a sample of N_{fam} maternal families (N_c) relative to the total number of unique seed trait values present in a seed population (N_t). Here, we define "unique seed trait values" as the number of non-redundant standardized measurements for the seed trait simulated rounded to the first digit. Seed morphological measurements were rounded to the first digits as we believe that seed trait variation estimated using additional digits is more likely to fail to capture meaningful biological variation. Standardization of the data was performed so that all seed traits share the same unit (the number of standard deviations a value is from the overall trait mean across populations, see equation (1) above) and become comparable. Sampling of maternal families and estimation of the summary statistic, defined as the proportion of total seed trait variance captured (N_c/N_t), were repeated 500 times for each seed morphological trait and Torrey pine population to account for the variation introduced through randomly sampling maternal families.

Finally, for each number of maternal trees sampled (N_{fam}), we averaged the summary statistic across all 500 replicates. This process was repeated for each of the 14 seed morphological traits and performed for each Torrey pine population separately. Following this, the summary statistic was averaged across all seed traits and separated by populations (see Results below). Proportions of total seed trait variance captured (N_c/N_t) are provided based on proportions of maternal families sampled (instead of the number of maternal families sampled) as sample sizes varied across Torrey pine populations.

RESULTS

Island-mainland differentiation in seed morphology – A principal component analysis (PCA) using all 14 measured and derived seed traits revealed substantial differences in seed morphology between island and mainland populations of Torrey pine (Fig. 3). The first PC axis explained 57.5% of variation in seed morphological traits, primarily separating the island from the mainland population. Seed length, seed width, seed area, endosperm area, and seed coat area exhibited the five highest loadings (absolute values) on PC1 (Appendix S1; see Supplemental data with this article), indicating that seed size and seed coat thickness can largely discriminate island from mainland individuals. On average, seeds collected on island trees were longer, wider, larger and thicker than seeds collected on mainland trees (Table 1). The second PC axis explained 16.5% of seed trait variation and summarizes within population variability in seed morphology (Fig. 3). Relative seed coat size, relative endosperm size, and relative embryo size had the three highest loadings (absolute values) on PC2 (Appendix S1). This suggests that once corrected for seed size, seed coat thickness, endosperm size, and embryo size are traits contributing to within-population variation.

Contribution of population origin and maternal family to seed trait variation – Consistent with our principal component analysis, linear mixed models constructed for each of the 14 measured and derived seed traits demonstrated that considerable variation in seed morphology in Torrey pine is explained by population origin (Fig. 4). On average, population origin explained 21% (range = 0.04–0.52) of variation across the species' distribution (Appendix S2). Traits associated with seed size and seed coat thickness exhibited the highest proportion of variance explained by population origin. These include seed coat area (0.52; $F = 202.5$, $df = 1$, $P < 0.001$), seed area

(0.45; F = 143.21, df = 1, P < 0.001), seed width (0.36; F = 130.71, df = 1, P < 0.001), endosperm area (0.34; F = 95.47, df = 1, P < 0.001), seed coat width (0.29; F = 87.54, df = 1, P < 0.001), and seed length (0.28; F = 76.18, df = 1, P < 0.001). Overall, this suggests seed size and seed coat thickness are major discriminants of island and mainland Torrey pine seeds.

While population origin explained substantial variation across populations, assessment of maternal seed families within populations indicated substantial family structure to seed trait variation (Fig. 4). On average, maternal seed family explained 22% (range = 0.05–0.34) of variation within populations (Appendix S2). Endosperm area (0.34; χ^2 = 272.03, df = 1, P < 0.001), seed length (0.33; χ^2 = 217.24, df = 1, P < 0.001), seed area (0.3; χ^2 = 288.62, df = 1, P < 0.001), seed coat width (0.29; χ^2 = 180.45, df = 1, P < 0.001), and embryo area (0.29; χ^2 = 142.9, df = 1, P < 0.001) exhibited the highest proportion of seed trait variation explained by within-population maternal families. This suggests that there is substantial family-level structure to seed size, endosperm size, embryo size and seed coat thickness within Torrey pine populations

Impact of population seed trait differentiation on seed emergence – The proportion of emerged seeds increased over time for both island (Q = 15.5, df = 2, P < 0.001) and mainland (Q = 15.2, df = 2, P < 0.001) populations (Fig. 5). However, we found no significant differences in the proportion of individuals emerging between populations across observed time points. On average, 7% and 9% of mainland and island seeds emerged a month after sowing (02/06/18; W = 28, P = 0.64), 63% and 53% of mainland and island seeds emerged a month and a half after sowing (02/16/18; t = 0.81, df = 14, P = 0.43), and 78% of mainland and island seeds emerged two months after sowing (03/07/18; t = −0.06, df = 14, P = 0.95). Overall, this indicates that

under controlled conditions, the timing and probability of emergence for Torrey pine is not impacted by population differences in seed morphology. Alternatively, this may also suggest that island and mainland seeds respond similarly to artificial germination protocol.

Morphological variation captured in simulated seed collections – Simulations revealed that to capture 95% of seed trait variation present in our existing ex situ collections, on average 80% (24 out of 30) and 68% (54 out of 80) of all island and mainland families would need to be resampled respectively (Fig. 6). This indicates that both island and mainland populations harbor considerable within-population structure for seed morphological traits. Interestingly, capturing equal morphological variation across seed collections always required a higher proportion of island maternal families to be collected relative to the mainland population. This suggests that the island population may exhibit higher variability in seed morphology compared to the mainland population, and thus may be worth of additional sampling efforts.

DISCUSSION

Evaluating between- and within-population morphological variation in contemporary ex situ seed collections may aid in understanding the distribution of variation needed to guide future conservation efforts. Here, we quantified the distribution of trait variation within an ex situ collection of Torrey pine seeds, with an aim to optimize future supplemental ex situ collections. Morphologically, island and mainland seeds were significantly different from each other. Island seeds were larger on average and had thicker seed coats relative to their mainland counterparts. These morphological differences may be explained by a combination of stochastic and deterministic factors associated with population origin. Interestingly, despite substantial

morphological differentiation, emergence did not vary among populations, suggesting that either the probability and timing of emergence under controlled conditions is not impacted by differences in seed morphology or that island and mainland seeds respond similarly to an artificial germination protocol. In addition to population origin, a considerable proportion of seed trait variation within each population was explained by maternal family. This suggests that there is substantial within-population variation that will be important to conserve and maintain populations' evolutionary potential. Finally, simulations demonstrated that 80% and 68% of all maternal families within island and mainland seed collection respectively were necessary to capture 95% of morphological variation within the existing collections. To ensure we maintain representative levels of trait variation in future seed collections, our results thus suggest that the number of maternal families sampled within natural populations should be maximized, with the island population potentially requiring additional sampling efforts.

Species distributed between island and mainland origins often exhibit marked among-population phenotypic differentiation, including differentiation in seed morphology (Burns et al., 2012; Lens et al., 2013; Kavanagh and Burns, 2014; Burns, 2016). Our results revealed considerable seed morphological differences between Torrey pine populations (Fig. 3), primarily in seed size and seed coat thickness (Fig. 4; Appendix S1). On average, islands seeds were larger and had thicker seed coats than seeds collected on the mainland (Table 1). These results are consistent with previous studies of island-mainland systems that noted island populations exhibited larger seeds relative to mainland populations (Kavanagh and Burns, 2014; Burns, 2016). A combination of different factors could contribute to morphological variation among seed populations, including both stochastic and deterministic forces.

On islands, seeds traits associated with long-distance dispersal may be selected against (Cody and Overton, 1996; Schenk, 2013; Kavanagh and Burns, 2014), as they could increase the probability an individual would disperse beyond the islands' limits. For potentially anemochorous species like Torrey pine, increased seed size on islands may limit potential losses via wind-dispersal and promote zoochory (Greene and Johnson, 1993; Reader, 1993; Benkman, 1995; Gómez, 2004). Seed predation may also have contributed to differences in seed size between populations. On the island, *Peromyscus maniculatus* (island deer mouse) is the only small mammal seed predator alongside fewer than a dozen bird species (https://www.nps.gov/chis/learn/nature/birds.htm; https://www.nps.gov/chis/learn/nature/terrestrial-animals.htm). This contrasts with the mainland, where already 11 mammalian seed predators exist; including squirrels, rats, mice, and gophers (https://torreypine.org/nature-center/mammals/). If large seeds are preferentially targeted by seed-eating animals (Reader, 1993; Gómez, 2004), smaller seed size on the mainland may have evolved in response to increased predation pressures from granivorous species.

While selection may contribute to population differences, differentiation in seed morphology may result from stochastic evolutionary forces. Founder effects associated with the colonization of Santa Rosa island by mainland individuals, as well as genetic drift combined with limited gene flow may have led to seed morphological differentiation between Torrey pine populations (Ledig and Conkle, 1983). While these factors may contribute to population differences in seed morphology, further experimentation is needed to adequately test these hypotheses. As seeds were collected from natural populations, measuring seeds produced by maternal plants grown in

a common environment is required to tease apart the contribution of environment and genetics to seed trait differences observed among populations. Additionally, a reciprocal transplant experiment would be the most effective test of the action of natural selection in shaping morphological differences between island and mainland seeds.

Despite significant differences in seed morphology between Torrey pine populations, timing and probability of emergence were similar across populations (Fig. 5). Emergence rates were high throughout the trial, with 78% of island and mainland seeds germinating two months after sowing. The absence of differences in emergence between populations was surprising, as seed size often negatively correlates with time to emergence (Daws et al., 2005; Tanveer et al., 2013). However, seed coat thickness can also influence rates of emergence. Seeds with thick seed coats relative to their mass often emerge later than seeds with thinner seed coats (Daws et al., 2005; Hamilton et al., 2013). Interestingly, island seeds were not only larger, but also had thicker seed coats relative to mainland seeds (Table 1). Even after correcting for differences in seed size, seed coat thickness (RSCS) remained moderately higher in island seeds. Together, these results predict that island seed could emerge at similar timepoints as mainland seed, or later than seeds collected on the mainland (Hamilton et al., 2017), which is consistent with our observations. In addition, similar emergence rates may result from our experimental design. Abe & Matsunaga (2011), in a mainland-island comparison study, observed that cold stratification attenuates differences in seed emergence between populations of *Rhaphiolepis umbellata*. Additionally, complete and rapid germination of pine seeds is generally observed when pretreated under cold and moist conditions (Bonner and Karrfalt, 2008). Overall, this suggests that cold stratification may mask population-specific differences in seedling emergence. Concurrent emergence of

seeds from both Torrey pine populations coupled with high emergence success suggests a cold stratification protocol is valuable for Torrey pine, particularly where simultaneous emergence for nursery-grown seedlings is desired.

Although population origin explained a substantial proportion of seed trait variation, linear mixed models demonstrated that maternal seed families within populations explained nearly as much variation (Fig. 4). Given generally high heritability for seed morphological traits and the half-sib design of our collection (Pandey et al., 1994; Cober et al., 1997; Mera et al., 2004; Hakim and Suyamto, 2017), family-level seed trait variation likely provides a useful proxy for assessing within-population genetic diversity. With more than a fifth of variation explained on average by maternal families (Appendix S2), this suggest there is substantial genetic structure within Torrey pine populations. These results were surprising, as previous studies using allozymes and chloroplast DNA both suggested that the species exhibits little to no within-population genetic variability (Ledig and Conkle, 1983; Waters and Schaal, 1991; Whittall et al., 2010). However, the common garden experiment conducted by Hamilton and colleagues (2017) indicated substantial family-level variation in tree height within both island and mainland populations. Overall, these results indicate that Torrey pine populations may possess within-population genetic variation necessary for natural selection to act upon. From a conservation perspective, these findings suggest that a strategy maximizing the number of maternal families sampled would optimize genetic diversity preserved in future ex situ seed collections.

Generally, ex situ seed collections aim to capture 95% of genetic diversity present throughout a species' distribution (Marshall and Brown, 1975; Brown and Marshall, 1995; Li et al., 2002;

Gapare et al., 2008). Simulations revealed that, in order to captured 95% of morphological variation currently maintained ex situ, 24 (80% of island collection) and 54 (68% of mainland collection) maternal families within each seed collection would need to be resampled (Fig. 6). These data indicate that sampling more maternal families from the island population may be necessary to achieve the same level of representation of morphological variation. If we presume that these phenotypes result from the combination of different alleles, capturing 95% of genetic variation within the island population will always require more maternal families relative to the mainland population. For these simulations, we assumed that contemporary ex situ collections captured all morphological variation both within and between populations, including seed phenotype frequencies. However, if this is not the case, these recommendations may result in suboptimal sampling of standing variation within targeted populations. This caveat is all the more important because sample sizes differed between island (30 maternal families) and mainland (80 maternal families) seed collections. An understanding of the breadth of variation within native populations captured in contemporary ex situ seed collections would be necessary to fully optimize recommendations regarding within-population sampling efforts using this approach.

CONCLUSION

Practical and cost-effective, long-term storage of seeds ex situ is widely used to capture and maintain rare species genetic diversity. These seed collections represent an invaluable resource to estimate the distribution of seed trait variation within and between populations that may be used to guide future ex situ sampling efforts. This research leverages ex situ seed collections of Torrey pine to understand existing variation within and among populations and inform sampling

for future seed collections. Incorporating existing information from ex situ collections provides a unique opportunity to optimize conservation objectives, particularly important for rare species.

ACKNOWLEDGEMENTS

The authors thank Annette Delfino Mix, Valerie Gallup, Andrew Bower, Drew Peterson, Emma Ordemann, Conner Harrington, Forest Swaciak, Jill Wulf, and Stephen Johnson for their help collecting the Torrey pine cones. The collections were funded by The USDA Forest Service, State and Private Forestry, Gene Conservation program to J.W.W. and J.A.H. Sara Wilson provided training and assistance for the use of the x-ray machine. We also thank the USDA Forest Service, Pacific Southwest Research Station for their support of this project. This work was also supported by a new faculty award from the office of the North Dakota Experimental Program to Stimulate Competitive Research (ND-EPSCoR NSF-IIA-1355466) and funding from the NDSU Environmental and Conservation Sciences Program to JAH. Any use of product names is for information purposes only and does not imply endorsement by the US Government.

DATA AVAILABILITY

The data for this article, including seed morphological measurements and the integrality of R scripts used are available from GitHub: https://github.com/LionelDiSanto/Di-Santo-et-al-2020

SUPPORTING INFORMATION

Supplementary information to this article may be found in the supporting information tab.

LITERATURE CITED

Abe, T., and M. Matsunaga. 2011. Geographic variation in germination traits in Melia azedarach and Rhaphiolepis umbellata. *American Journal of Plant Sciences* 2: 52.

Abràmoff, M. D., P. J. Magalhães, and S. J. Ram. 2004. Image processing with ImageJ. *Biophotonics international* 11: 36–42.

Barrett, R. D. H., and D. Schluter. 2008. Adaptation from standing genetic variation. *Trends in ecology & evolution* 23: 38–44.

Bartoń, K. 2020. MuMIn: Multi-Model Inference. R package version 1.43.17.

Basey, A. C., J. B. Fant, and A. T. Kramer. 2015. Producing native plant materials for restoration: 10 rules to collect and maintain genetic diversity. *Native Plants Journal* 16: 37–53.

Bates, D., M. Mächler, B. Bolker, and S. Walker. 2014. Fitting linear mixed-effects models using lme4. arXiv. org. *preprint]. doi* 10.

Benjamini, Y., and Y. Hochberg. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)* 57: 289–300.

Benkman, C. W. 1995. Wind dispersal capacity of pine seeds and the evolution of different seed dispersal modes in pines. *Oikos*: 221–224.

Bonner, F. T., and R. P. Karrfalt. 2008. The woody plant seed manual. *Agric. Handbook No. 727. Washington, DC. US Department of Agriculture, Forest Service. 1223 p.* 727.

Brown, A. H. D., and D. R. Marshall. 1995. A basic sampling strategy: theory and practice. *Collecting plant genetic diversity: technical guidelines. CAB International, Wallingford*: 75–91.

Burns, K. C. 2016. Size changes in island plants: independent trait evolution in Alyxia ruscifolia (Apocynaceae) on Lord Howe Island. *Biological Journal of the Linnean Society* 119: 847–855.

Burns, K. C., N. Herold, and B. Wallace. 2012. Evolutionary size changes in plants of the south□west Pacific. *Global Ecology and Biogeography* 21: 819–828.

Carlson, S. M., C. J. Cunningham, and P. A. H. Westley. 2014. Evolutionary rescue in a changing world. *Trends in Ecology and Evolution* 29: 521–530.

Cober, E. R., H. D. Voldeng, and J. A. Fregeau-Reid. 1997. Heritability of seed shape and seed size in soybean. *Crop Science* 37: 1767–1769.

Cody, M. L., and J. M. Overton. 1996. Short-term evolution of reduced dispersal in island plant populations. *Journal of Ecology*: 53–61.

Critchfield, W. B., and E. L. Little. 1966. Geographic distribution of the pines of the world. US Department of Agriculture, Forest Service.

Daws, M. I., N. C. Garwood, and H. W. Pritchard. 2005. Traits of recalcitrant seeds in a semi□deciduous tropical forest in Panamá: some ecological implications. *Functional Ecology* 19: 874–885.

Dusek, K. H. 1985. Update on our rarest pine. *American forests (USA)*.

Etterson, J. R., S. J. Franks, S. J. Mazer, R. G. Shaw, N. L. Soper Gorden, H. E. Schneider, J. J. Weber, et al. 2016. Project baseline: An unprecedented resource to study plant evolution across space and time. *American Journal of Botany* 103: 164–173.

FAO. 2010. The second report on the state of the world's plant genetic resources for food and agriculture. Food & Agriculture Org, Rome.

Franklin, J., and E. V Santos. 2011. A spatially explicit census reveals population structure and

recruitment patterns for narrowly endemic pine, Pinus torreyana. *Plant Ecology* 212: 293–306.

Gapare, W. J., A. D. Yanchuk, and S. N. Aitken. 2008. Optimal sampling strategies for capture of genetic diversity differ between core and peripheral populations of Picea sitchensis (Bong.) Carr. *Conservation Genetics* 9: 411–418.

Gómez, J. M. 2004. Bigger is not always better: conflicting selective pressures on seed size in Quercus ilex. *Evolution* 58: 71–80.

Greene, D. F., and E. A. Johnson. 1993. Seed mass and dispersal capacity in wind-dispersed diaspores. *Oikos*: 69–74.

Gribko, L. S., and W. E. Jones. 1995. Test of the float method of assessing northern red oak acorn condition. *Tree Planters' Notes* 46: 143–147.

Guerrant Jr, E. O., K. Havens, and P. Vitt. 2013. Sampling for effective ex situ plant conservation. *International Journal of Plant Sciences* 175: 11–20.

Hakim, L., and S. Suyamto. 2017. Gene action and heritability estimates of quantitative characters among lines derived from varietal crosses of soybean. *Indonesian Journal of Agricultural Science* 18: 25–32.

Hall, T., and A. Brinkman. 2015. Population Dynamics of the Island Torrey Pine (Pinus torreyana ssp. insularis) on Santa Rosa Island, CA. California State University Chanel Islands.

Hamilton, J. A., R. Royauté, J. W. Wright, P. Hodgskiss, and F. T. Ledig. 2017. Genetic conservation and management of the California endemic, Torrey pine (Pinus torreyana Parry): Implications of genetic rescue in a genetically depauperate species. *Ecology and Evolution*.

Hamilton, K. N., C. A. Offord, P. Cuneo, and M. A. Deseo. 2013. A comparative study of seed morphology in relation to desiccation tolerance and other physiological responses in 71 Eastern Australian rainforest species. *Plant Species Biology* 28: 51–62.

Hausman, C. E., M. M. Bertke, J. F. Jaeger, and O. J. Rocha. 2014. Genetic structure of green ash (Fraxinus pennsylvanica): implications for the establishment of ex situ conservation protocols in light of the invasion of the emerald ash borer. *Plant Genetic Resources* 12: 286–297.

Hoban, S., and S. Schlarbaum. 2014. Optimal sampling of seeds from plant populations for ex-situ conservation of genetic biodiversity, considering realistic population structure. *Biological Conservation* 177: 90–99.

Hothorn, T., and K. Hornik. 2019. exactRankTests: Exact Distributions for Rank and Permutation Tests. R package version 0.8-31.

IUCN 2020. The IUCN Red List of Threatened Species. Version 2020-2. https://www.iucnredlist.org. Downloaded on 09 July 2020.

Kassambara, A. 2020. rstatix: Pipe-Friendly Framework for Basic Statistical Tests. R package version 0.4.0.

Kavanagh, P. H., and K. C. Burns. 2014. The repeated evolution of large seeds on islands. *Proceedings of the Royal Society B: Biological Sciences* 281: 20140675.

Kuznetsova, A., P. B. Brockhoff, and R. H. B. Christensen. 2016. Tests in linear mixed effects models version. *Cran. URL https://cran. r-project. org/web/packages/lmerTest/lmerTest. pdf*.

Ledig, F. T., and M. T. Conkle. 1983. Gene diversity and genetic structure in a narrow endemic, Torrey Pine (Pinus torreyana Parry ex Carr.). *Evolution* 37: 79–85.

Lens, F., N. Davin, E. Smets, and M. del Arco. 2013. Insular woodiness on the Canary Islands: a remarkable case of convergent evolution. *International Journal of Plant Sciences* 174: 992–1013.

Li, Q., Z. Xu, and T. He. 2002. Ex situ genetic conservation of endangered Vatica guangxiensis (Dipterocarpaceae) in China. *Biological Conservation* 106: 151–156.

Marshall, D. R., and A. H. D. Brown. 1975. Optimum sampling strategies in genetic conservation. *In* O. H. Frankel, and J. G. Hawkes [eds.], Crop Genetic Resources for Today and Tomorrow, 53–80. Cambridge University Press, Cambridge.

Mera, M., R. Jerez, H. Miranda, and J. L. Rouanet. 2004. Seed coat specific weight in Lupinus angustifolius: influence of genotype and environment and relationship with seed coat proportion. *Australian journal of agricultural research* 55: 1189–1195.

Morina, D. L., M. A. Lashley, M. C. Chitwood, C. E. Moorman, and C. S. DePerno. 2017. Should we use the float test to quantify acorn viability? *Wildlife Society Bulletin* 41: 776–779.

Pandey, M. P., D. V Seshu, and M. Akbar. 1994. Genetics of embryo size and its relationship with seed and seedling vigour in rice (Oryza sativa L.). *Indian Journal of Genetics and Plant Breeding* 54: 68.

Potter, K. M., R. M. Jetton, A. Bower, D. F. Jacobs, G. Man, V. D. Hipkins, and M. Westwood. 2017. Banking on the future: progress, challenges and opportunities for the genetic conservation of forest trees.

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Reader, R. J. 1993. Control of seedling emergence by ground cover and seed predation in

relation to seed size for some old-field species. *Journal of Ecology*: 169–175.

Di Santo, L. N., and J. A. Hamilton. 2020. Using environmental and geographic data to optimize ex situ collections and preserve evolutionary potential. *Conservation Biology*.

Schaal, B., and W. Leverich. 2004. Population genetic issues in ex situ plant conservation. Island Press: Washington, DC, USA.

Schenk, J. J. 2013. Evolution of limited seed dispersal ability on gypsum islands. *American journal of botany* 100: 1811–1822.

Tamaki, S., K. Isoda, M. Takahashi, H. Yamada, and Y. Yamashita. 2018. Genetic structure and diversity in relation to the recently reduced population size of the rare conifer, Pseudotsuga japonica, endemic to Japan. *Conservation Genetics*: 1–13.

Tanveer, A., M. Tasneem, A. Khaliq, M. M. Javaid, and M. N. Chaudhry. 2013. Influence of seed size and ecological factors on the germination and emergence of field bindweed (Convolvulus arvensis). *Planta daninha* 31: 39–51.

Waters, E. R., and B. A. Schaal. 1991. No variation is detected in the chloroplast genome of Pinus torreyana. *Canadian Journal of Forest Research* 21: 1832–1835.

Weber, A., and A. Kolb. 2014. Differences in heritable trait variation among populations of varying size in the perennial herb Phyteuma spicatum. *Conservation genetics* 15: 1329–1337.

Westengen, O. T., S. Jeppson, and L. Guarino. 2013. Global ex-situ crop diversity conservation and the Svalbard Global Seed Vault: Assessing the current status. *PloS one* 8: e64146.

Whittall, J. B., J. Syring, M. Parks, J. Buenrostro, C. Dick, A. Liston, and R. Cronn. 2010. Finding a (pine) needle in a haystack: chloroplast genome sequence divergence in rare widespread pines. *Molecular Ecology* 19: 100–114.

Yoko, Z. G., K. L. Volk, N. A. Dochtermann, and J. A. Hamilton. 2020. The importance of quantitative trait differentiation in restoration: landscape heterogeneity and functional traits inform seed transfer guidelines. *AoB Plants* 12: plaa009.

Zhang, D.-Q., and N. Zhou. 2013. Genetic diversity and population structure of the endangered conifer Taxus wallichiana var. mairei (Taxaceae) revealed by Simple Sequence Repeat (SSR) markers. *Biochemical systematics and ecology* 49: 107–114.

TABLES

Table 1. Morphological measurements of Torrey pine seeds sourced from Santa Rosa Island (Island, n=30) and Torrey Pine State Reserve (Mainland, n=80), CA. Listed are measured (A) and derived (B) seed trait values averaged over maternal families given per populations (±SE). Measurable traits: seed length (SL), seed width (SW), embryo length (EL), embryo width (EW), seed coat width (SCW), seed area (SA), endosperm area (ESA), and embryo area (EA). Derived traits: seed length/width ratio (SLW), embryo length/width ratio (ELW), relative embryo size (RES), relative endosperm size (REndS), seed coat area (SCA), and relative seed coat size (RSCS). Differences in seed morphology between mainland and island population were significant (α=0.05) for all 14 seed traits.

A – Measured seed traits

Population	SL (cm)	SW (cm)	EL (cm)	EW (cm)	SCW (cm)	SA (cm^2)	ESA (cm^2)	EA (cm^2)
Mainland	1.42±0.016	0.77±0.009	1.09±0.013	0.15 ±0.003	0.09±0.002	0.89±0.023	0.55±0.014	0.17±0.004
Island	1.67±0.019	0.97±0.014	1.18 ±0.019	0.18 ±0.003	0.12±0.002	1.39 ±0.032	0.81±0.021	0.23±0.007

B – Derived seed traits

Population	SLW (SL/SW)	ELW (EL/EW)	RES (EA/ESA)	REndS (ESA/SA)	SCA (cm^2) (SA-ESA)	RSCS (SCA/SA)
Mainland	1.86±0.013	7.3 ±0.092	0.31±0.005	0.62±0.003	0.34±0.009	0.38±0.003
Island	1.74 ±0.021	6.59±0.132	0.28 ±0.005	0.58±0.004	0.59±0.013	0.42±0.004

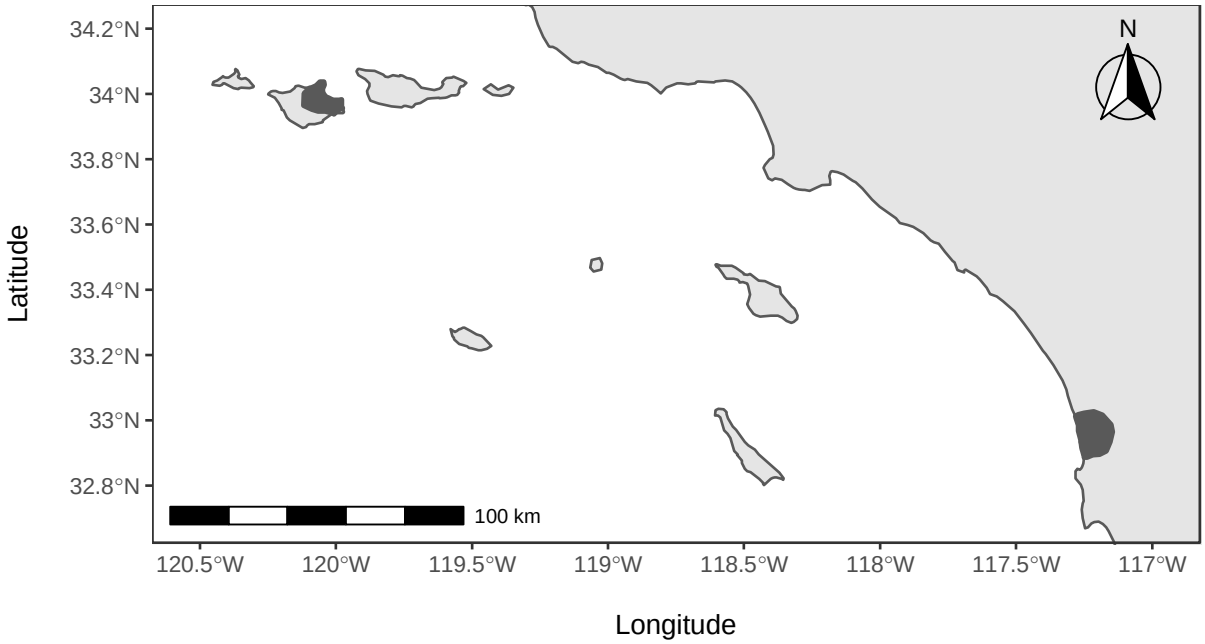

Fig. 1. *Pinus torreyana* distribution map. Top left dark grey area indicates the distribution of Torrey pine on Santa Rosa Island, CA (*Pinus torreyana* sp. *insularis*). Bottom right dark grey area indicates the distribution of Torrey pine at the Torrey Pine State Reserve, CA (*Pinus torreyana* sp. *torreyana*).

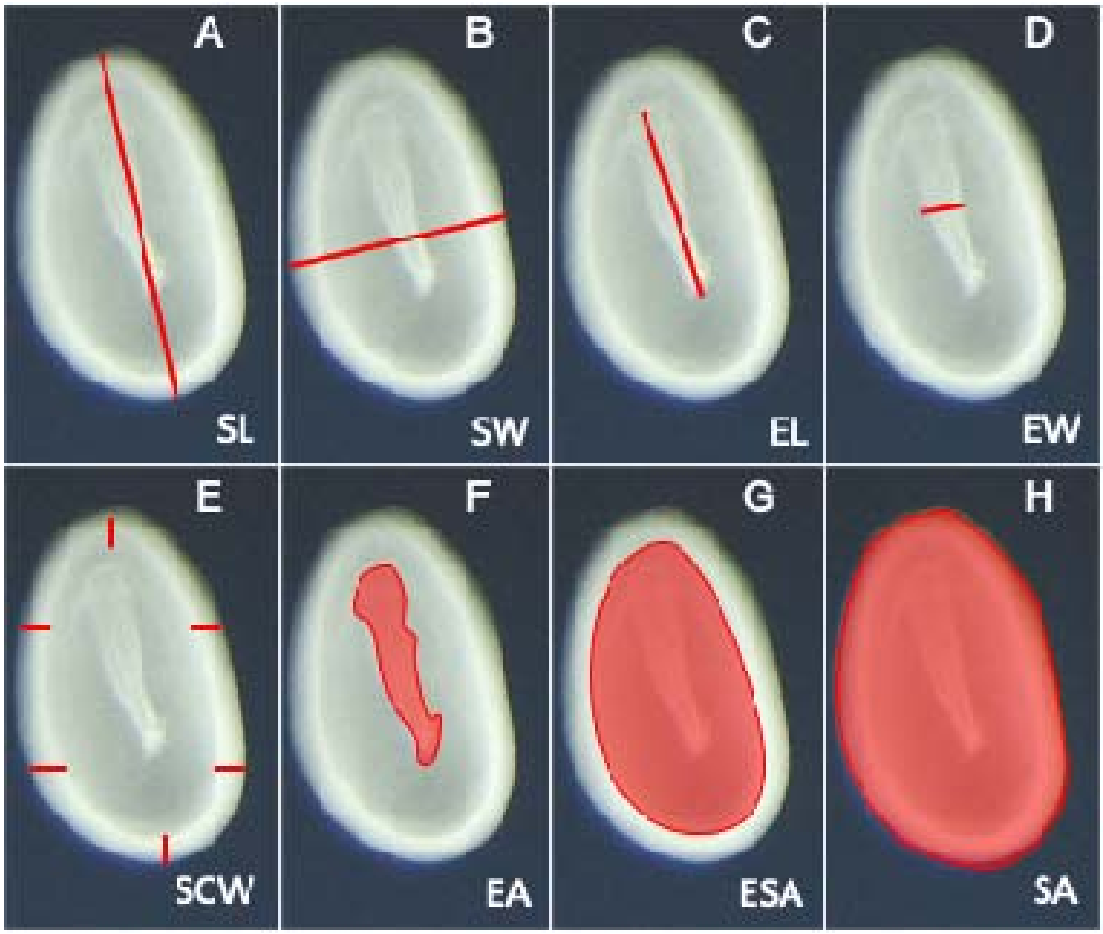

Fig. 2. Visual of morphological measurements taken using ImageJ for seeds collected on Santa Rosa Island and at the Torrey pine State Reserve. (A) Seed length [cm]. (B) Seed width [cm]. (C) Embryo length [cm]. (D) Embryo width [cm]. (E) Seed coat width [cm]. (F) Embryo area [cm^2]. (G) Endosperm area [cm^2]. (H) Seed area [cm^2].

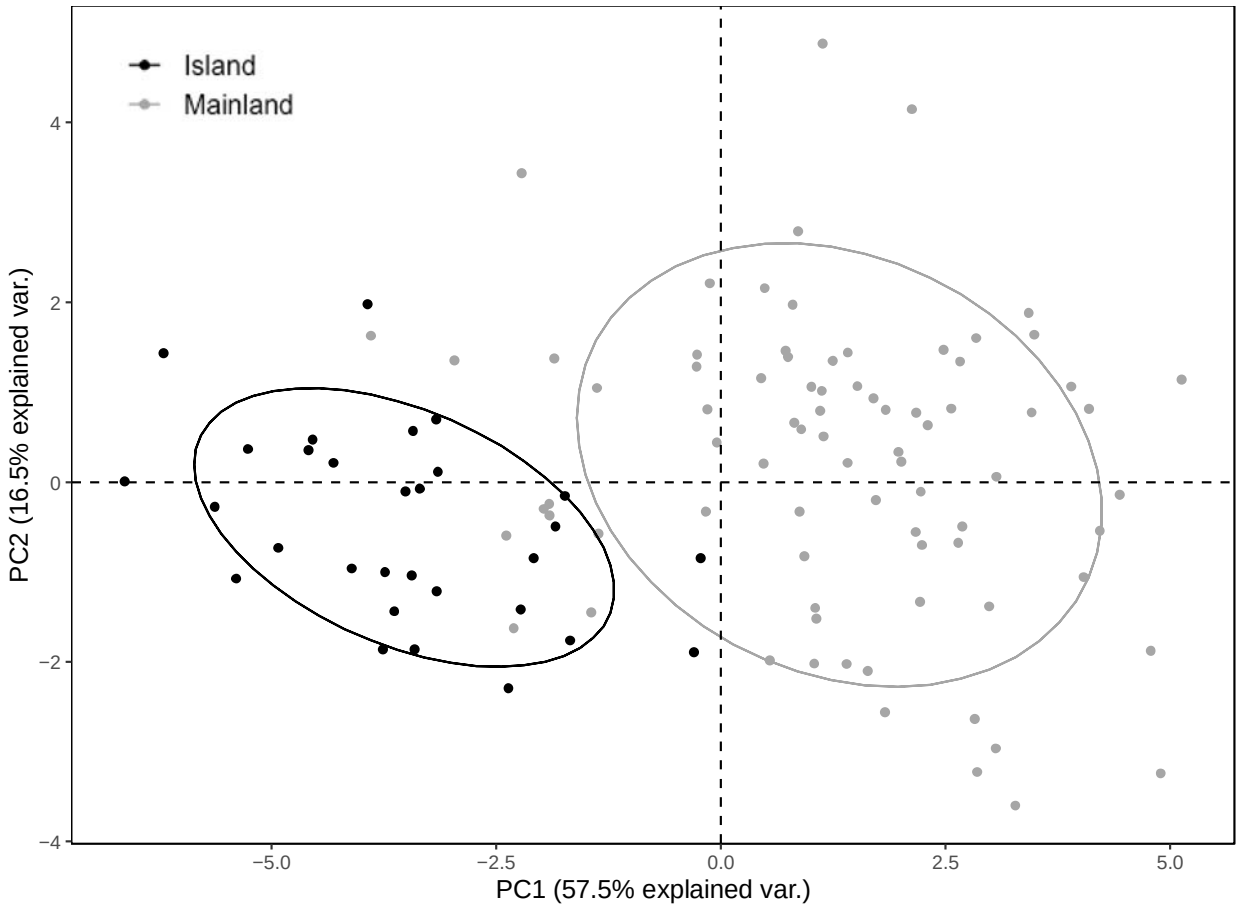

Fig. 3. Principal components analysis (PCA) using all 14 seed morphological traits measured from maternal plants collected on Santa Rosa Island (black) and at the Torrey Pine State Reserve (grey).

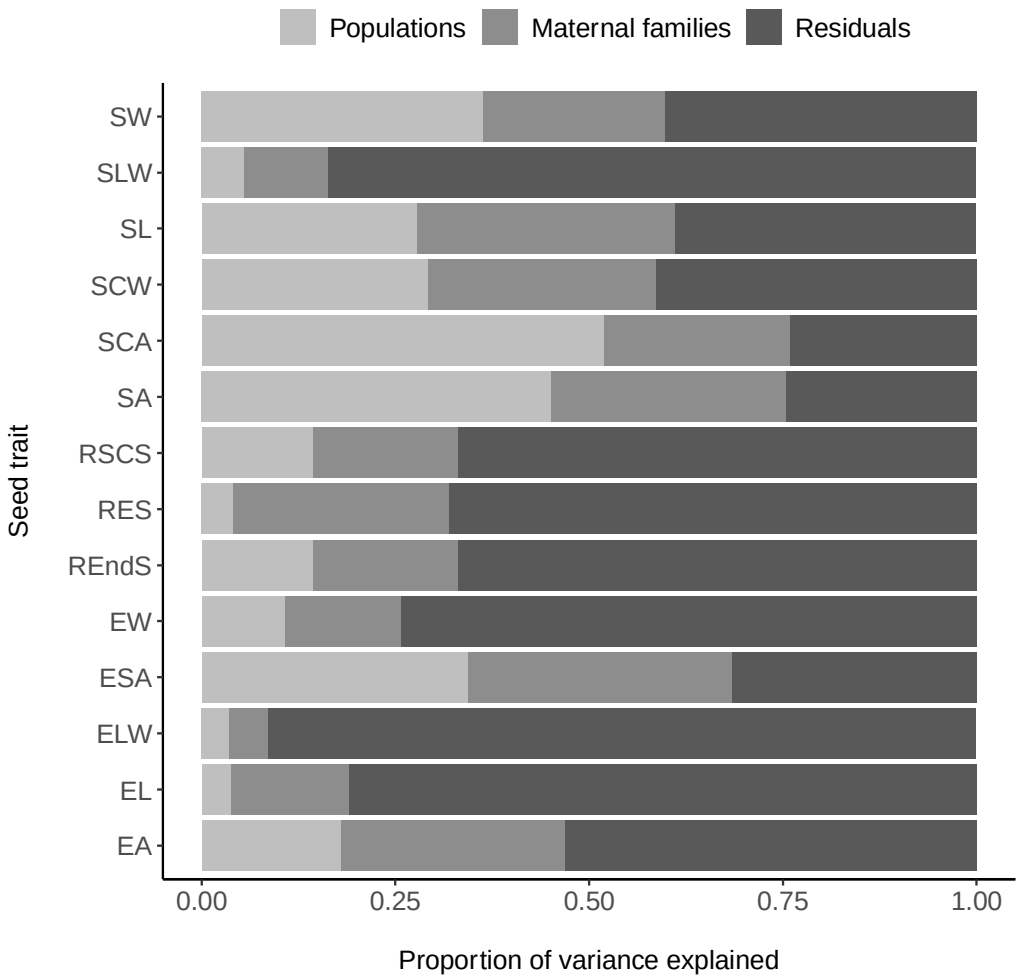

Fig. 4. Proportion of variance in seed morphology explained by populations (light grey), maternal families within populations (grey), and other variables not accounted for in the model (residuals; dark grey) for each of the 14 measured and derived seed traits. SW, seed width (cm); SLW, seed length/width ratio; SL, seed length (cm); SCW, seed coat width (cm); SCA, seed coat area (cm^2); SA, seed area (cm^2); RSCS, relative seed coat size; RES, relative embryo size; REndS, relative endosperm size; EW, embryo width (cm); ESA, endosperm area (cm^2); ELW, embryo length/width ratio; EL, embryo length (cm); EA, embryo area (cm^2). See Appendix S2 for numerical estimates.

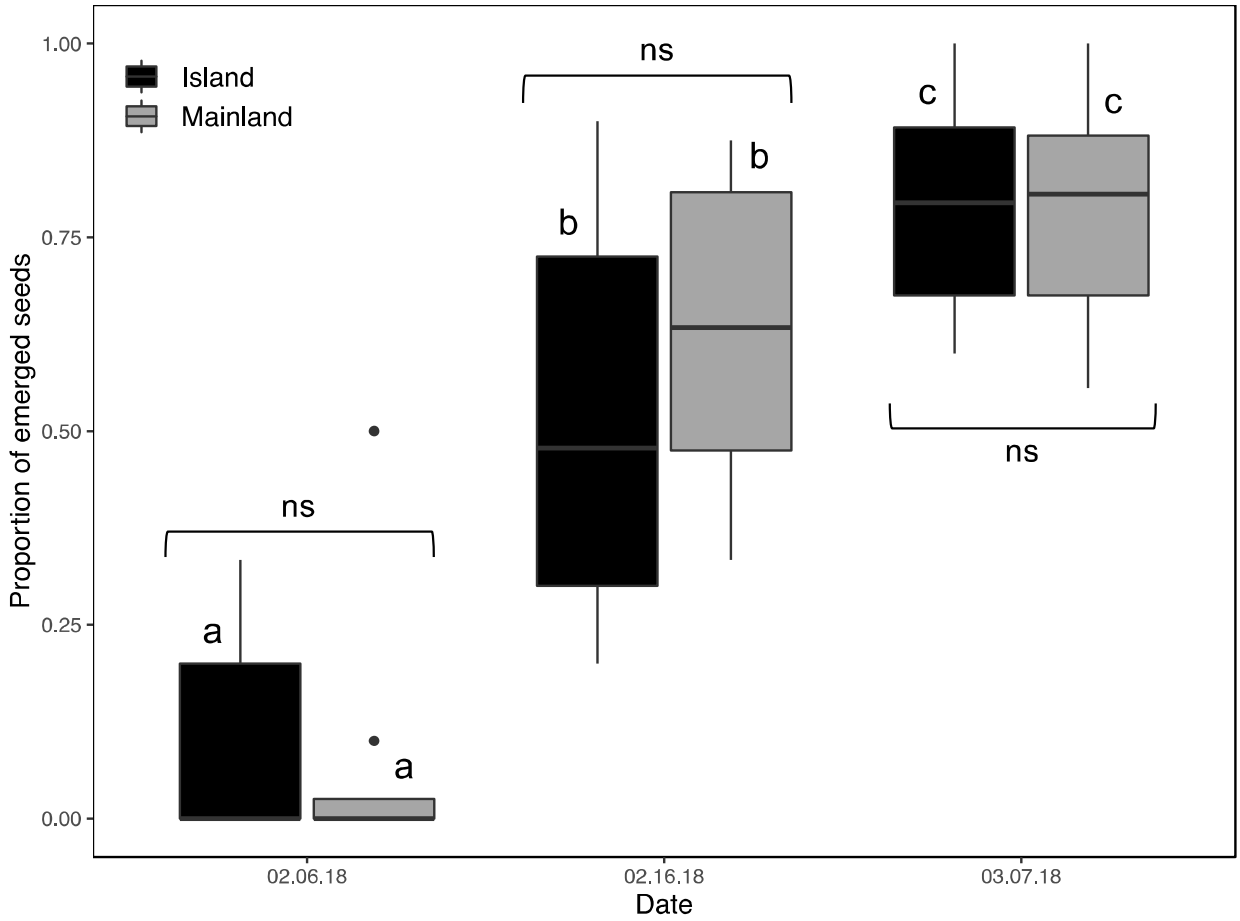

Fig. 5. Seed emergence, measured as the proportion of emerged seeds (y axis), recorded at three different timepoints (x axis) for seeds sampled on Santa Rosa Island (black) and at the Torrey Pine State Reserve (gray). Significant differences in emergence time across timepoints within populations are indicated with different letters. Comparisons between populations at each timepoint is indicated with square brackets. ns, non-significant differences (α=0.05).

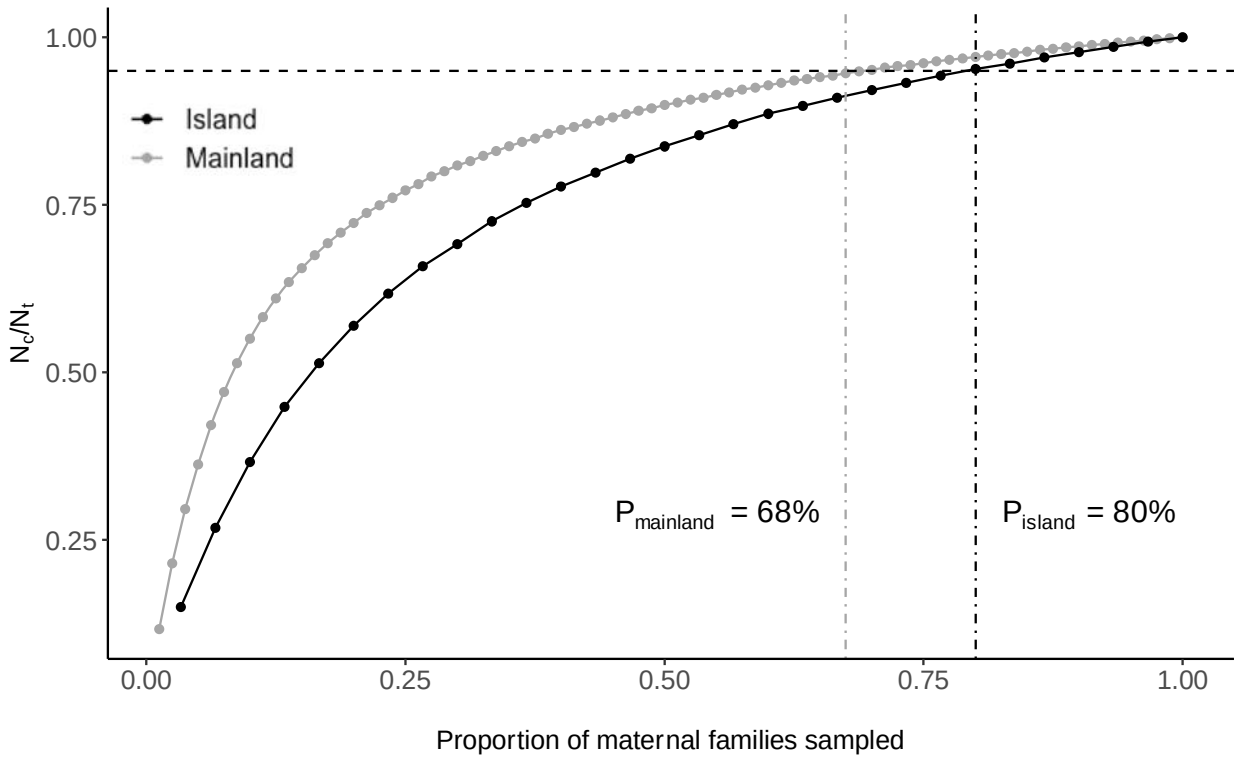

Fig. 6. Average proportion of phenotypic variation captured across seed traits estimated for various proportions of maternal families sampled. P_{island} and $P_{mainland}$ represent the proportion of maternal families required to capture 95% of morphological variation (horizontal dashed line) present in island (black) and mainland (gray) ex situ seed populations respectively.