**Phylogeny Benefits Ecological Restoration: Phylogenetic Analysis of Seed Morphology Predict Seed Germination**

**Abstract:**

Seed germination is critical for species survival and colonization of new habitat. Systematically obtaining some insights of seed germination could be beneficial for seeding in wildland restoration and agricultural practice. In this study, we evaluated relationships between seed germination rate and seed morphological measurements (seed mass and seed surface area) for 19 common shortgrass prairie species. We tested the phylogenetic signal in each seed measurement using Pagel’s lambda and Blomberg’s K. We also applied phylogenetic comparative methods to explore potential patterns between seed germination rate and seed morphological traits. Our results show that seed morphology in shortgrass prairie species contains phylogenetic signal, and seed morphology is weakly correlated with seed germination. These findings support that phenotypic evolution plays a role in shortgrass prairie conservatism. To benefit the shortgrass prairie re-seeding restoration, the phylogenetic analysis of physiological seed traits is expected for future study.

**Key words:** seed germination, seed traits, Pagel’s lambda, Blomberg’s K, phylogenetic generalized least square

# Introduction

In ecological restoration, germination of a wide variety native species is the threshold. More specifically, most native seeds are dormant and need environmental triggers prior to germination. Only limited species and environmental triggers have been studied, which are far less than the species needed in ecological restoration. Instead of testing germination on each species, if we can draw some connection between seed germination and easily measured seed traits, it will help predict germination of the native seed mixture and directly benefit the ecological restoration.

Seeds contain and express genetic information to maintain specie’s traits, such as seed germination and seed morphology. Seed germination varies by species, which is largely influenced by genetic factors, dormancy, epigenetic factors, and environmental signals (Citation). Further, long persistent epic factors will drive species evolution, leading species to develop environmentally related traits and break dormancy after sense environmental signals. For instance, if a species establishes and survives in a fire dominated region, the species may develop heritable fire related traits, responding to heat, smoke or sunlight exposure (Citation). Resolving dormancy breaking traits will help identify the genetic factors of seed germination. Germination experiments are needed to understand seed dormancy as a species trait. However, if we can discover other easily measured heritable traits associate with seed dormancy, a practical measurement may develop to estimate seed dormancy and predict seed germination of each ingredient of a seed mix.

Seed morphology is related with seed germination and hazard tolerance (Westoby et al. 2002) and may be a good indicator for seed germination as well. Seed mass and seed shape are two major heritable seed indexes. Seed mass represents three fourths of the energy allocated by plants to reproduction, supporting the initial growth of a plant. Seed shape will influence seed dispersal which may lead to potential hazard by wind, water or animals and delay seed germination. Strong evidence supports (Leishman et al. 2000) that large seeds can stay dormant longer and produce stronger seedlings after germination, benefiting from nutrition stored in seed under “larger-seed-later-deployment” interpretation or “cotyledon functional morphology hypothesis”. In general, large seed can be preserved longer and germinate better than small seed under hazards. Exceptions happen in some hazardous condition may explain by larger seed size might initially evolve in one type of advantage, but also associate with other hazards later. Since seed morphology is related to seed germination, it has the potential to predict seed germination. Further, seed mass and seed shape are easily measured, it would be a good practical tool for restoration practice to understand germination variation of restoration seed mix. However, while seed morphological traits are heritable in order to maintain species’ characteristics, the heritable information also is largely shared within common ancestors and closely related species. This biological relation may lead to statistical violation for independent observation, which needs support from phylogenetic comparative analysis.

Phylogeny is the systematic analysis of living organism, which represents evolution of species over time. Phylogenetic information uses tree shape (phylogenetic tree) to express species relation and diversification. A phylogenetic tree alone may only contain structure information of branch length, tips, and nodes, but analysis of relative position of species can reveal how traits evolve across species. Typically, on phylogenetic trees, large portion of species with similar traits are positioned closely meaning under similar environmental stress their ancestor evolved similar traits and passed them to descendants. In trait evolution, the relative position of species on trees represents the trait development path. Evolutionarily closely related species are likely to inherit trait information from their common ancestor. If so, trait measurements can’t be analyzed in common statistical model, since it violates the foundation assumption of independence. In 1985, Felsenstein (1985) proposed independent contrast to solve the non-independent observation for analysis on traditional trait regression. Under the assumption of random walk process, i.e. Brownian motion, the observed individual represents independent evolution pathway, therefore differences between closely related species are independent. However, there are circumstances that phylogeny will be over-corrected by phylogenetic independent contrast, since sometimes the closely related species may not be similar. In such condition, traditional trait analysis without phylogeny may be more suitable. Some have tried to explore the relation between single seed trait and phylogenetic information, such as seed dormancy (Willis et al. 2014, Dayrell et al. 2016), seed dispersal (Razafindrastsima et al. 2016), seed bank composition (Griffiths et al. 2014), seed germination (Carta et al. 2016). Others try to combine phylogeny with other scientific models and reveal how one trait evolves and how different traits interact with each other (Gallagher et al. 2014, Paquette et al. 2015). In 2018, Dr. Barak and her colleagues (Barak et al. 2018) explored the relationship between seed traits and seed germination with phylogenetic information on tallgrass prairie restoration preferred species. Their work measured seed mass, seed shape (length, height, shape) and the ratio of embryo and seed, and concluded that seed shape best predicted seed germination response.

In this paper, we tested phylogenetic signal in seed morphological traits of shortgrass prairie species (seed mass and seed surface area). We also test the relationship between seed morphological traits and seed germination response with phylogenetic comparative method.

**Discussion**

Seed morphology of shortgrass praire species correlated with seed germination rate (Table 1), which means that seed morphology can partially explain seed germination. These results are aligned with tallgrass prairie restoration species research (Barak et al. 2018), in which they predict that seed shape is best seed germination predictor in comparison with seed mass, environmental stimulation and phylogeny. Interestingly, in Westoby and his colleague’s (2002) research summarized that across locations and species that seeds are generally filled into six categories, however, due to the narrow scope of my study, most our shortgrass prairie species are only filled into two of six size categories. Even though seed size is clustered in shortgrass prairie, it still can help to predict seed germination in general.

Seed mass and seed shape contain phylogenetic signal directly prove that phylogeny influence seed morphology (Table 2). Many of previous seed study showing the phylogenetic information in seed dormancy, seed dispersal and seed germination. Previous comprehensive seed dormancy study detected strong phylogenetic signal in seed dormancy (Dayrell et al. 2016) and morphophysiologically dormancy is most likely to be ancestral state of seed plants (Willis et al. 2014). Among evolutionary transitions between dormancy classification, the physiological dormancy is associated with high diversification rate via high speciation rate (Willis et al. 2014). In the case study of Frugivores, researchers caught nonrandom phylogenetic associations of seed dispersal, though the association fluctuated from clustered to overdispersal during the 16 months of their study (Razafindrastsima et al. 2016). The phylogenetic related of species are required similar germination condition, it also mediate by environmental factors (Carta et al. 2016). More specific in tallgrass prairie restoration species study, Barak and her colleagues proven that not only seed shape contains phylogenetic information and be the best predictor of germination, but the species position on phylogenetic tree can also catch heritable residuals beyond heritable traits.

The difference of correlations between seed morphology and seed germination rate with and without phylogeny definitely raise our awareness of analysis phylogenetic information in biological measurements (Figure 1). In general, biological trait measurements are influenced heritability and environment, which the heritable information passes through generations. When the testing species are share common ancestor recently, the biological measurements may be influenced. Further, it is violated the independent assumption of most statistical analysis. Phylogenetic comparative methods offer methods measured the evolution traits change across species under Brownian Motion assumption. Because even though the trait values of current observations may be influenced by the biological relation, evolutionary development, the differences between trait values are independent.

Phylogenetic independent contrast (PIC) was developed to peel off the heritable influence of traits (Felsenstein 1985). Later, phylogenetic generalized least squares (PGLS) were described under the same reignition of statistical non-independence of biological observations (Grafen 1989). Even though both PIC and PGLS are the same at their heart (Blomberg et al. 2012), the PGLS has advantage of excluding the assumption of closely related species are similar, which sometimes violate by convergent evolution. Further, PGLS allows incorporating phylogenetic signal automatically in packages, but the independent contrast only have the correction in theory. However, in my study of shortgrass prairie species, the similarity of closed related species makes no differences of choosing phylogenetic comparative methods, PIC and PGLS. After non-independent statistical influence by heritable information, my biological studies have been tested the effects.

In the content of ecological restoration, seed morphology including the seed mass and seed surface area can be easily measured and create an index table. Accordingly, seed morphology classification and coherent seed germination rate categories can help restoration practice to estimate species germination rate and design seed mix composition. Proven by my study and my other previous work, it’s important to take evolution into account when we do biological study. Phylogenetic comparative methods, PIC and PGLS, are able to handle no-independent statistically from phylogenetic aspects (Felsenstein 1985, Grafen 1989, Blomberg et al. 2012).

For further study, my study may contain unexpected variable from species. The germination data of this study was gathered from previous germination studies, which were selected for those purpose. It would be more comprehensive evaluation if I could have a large number of tested species designated for my purpose. For instance, most of the tested species were grass and forb, and their seed mass and surface area were closed. However, one bush species, *Prosopis glandulosa,* was an outliner. These may influence the accuracy of overall tests.

The phylogenetic information also could be more accurate. Since my phylogenetic tree is pruned from a super tree, generated from GenBank Sequencing (Zanne et al. 2014), there are two better way to generate new tree include missing species: replicate random placement of missing species; use sequencing of missing species and desire species to build my own tree. Either of these methods will improve the accuracy of phylogeny in this project.