TREE IMPROVEMENT

DEVELOPMENT OF PRELIMINARY SEED ZONES FOR KOA

Christina Liang, Ph.D. (USDA Forest Service, Institute of Pacific Islands Forestry), Jessica W. Wright, Ph.D. (USDA Forest Service, Pacific Southwest Research Station), Nick Dudley (Hawai'i Agriculture Research Center), and Paul F. Gugger, Ph.D. (University of Maryland Center for Environmental Science, Appalachian Laboratory)

Paper Title: Koa seed zones: What do we know? What do we need to know?

Abstract

Acacia koa A. Gray (koa) exhibits a high degree of phenotypic variation, and understanding the underlying natural genetic variation that currently exists in the species has applications for conservation, restoration, and reforestation. Understanding the genetic variation also can be helpful in identifying and refining seed zones for koa that can be used for these multiple purposes. A seed zone is a single geographical or ecological unit within the range of a species based on ecological and genetic criteria. Ideally it allows for the identification and selection of high quality seed sources for planting efforts. A general seed zone map (used for multiple species) has been developed for many regions of the United States; seed zones can be further refined for individual species based on additional genetic or ecological information. A general seed zone map does not currently exist for Hawai'i, nor does a seed zone map specifically for koa. Our goal in this study is to provide a framework to identify and guide the development of koa seed zones throughout the state of Hawai'i.

Koa is found in a wide range of environments that include subalpine, montane, wet and lowland forest eco-zones. It is found naturally on all the main Hawaiian Islands except Ni'ihau and Kaho'olawe, but contiguous and dense forest is generally only on Hawai'i, Maui, O'ahu, and Kaua'i. Because of the high variability in koa, the current recommendation is to plant locally-sourced seed to ensure the high quality of seedlings. The question arises, however, of "how local is local?" Seed or planting zones for koa are not well defined, partially because of the limited information available on koa population genetics.

To start to address this question, we sampled 311 koa trees from across the 4 main islands in order to study the naturally occurring genetic variation within the species. We sequenced the DNA and obtained 11,002 diallelic single-nucleotide polymorphisms (SNPs) using the next-generation genotyping-by-sequencing (GBS) method. SNPs are molecular genetic markers and are found throughout the genome in both coding and non-coding regions. We used the SNP allele frequency to calculate pair-wise F_{ST} values for preliminary seed zones that we first identified. F_{ST}, known as the fixation index, is a common metric used in population genetics to measure the amount of population differentiation due to genetic structure. We also used the SNP data to estimate genetic clusters using the Admixture program, which estimates individual ancestries based on maximum likelihood. While calculation of F_{ST} in tetraploid species such as koa using SNP data that are coerced to be diploid can be problematic, as assumptions about allele frequency calculations are not necessarily met, our goal is to compare the relative differences between preliminary seed zones. We believe that the assumptions about allele frequency are reasonable given the large number of SNPs, and the mechanism of inheritance in koa likely does not vary across regions.

We preliminarily defined seed zones for the state based on eco-regions. Zones first were proposed for each island, and then by aspect within each island which was generally windward and leeward. Elevational sub-zones then were defined, with low elevation from sea level to 600 m (1968 ft) elevation, mid-elevation from 600 m to 1200 m (1968 ft to 3937 ft) elevation, and high elevation from 1200 m to 1800 m (3937 ft to 5905 ft) elevation. Zones for special situations also were defined, such as for the koai'a population in north Hawai'i and the lowland population in west Maui. Once the eco-region seed zones were defined, we assigned these zones to the trees that we sampled for our GBS analysis. We identified 10 seed zones from our samples. We then calculated the pair-wise F_{ST} values for these seed zones using the SNP allele frequencies. Results from this analysis showed the highest levels of differentiation between the Hawai'i and Kaua'i populations, which also are the most geographically and chronologically distant. Lower differences within islands were found, with the greatest difference between the two Maui seed zones and smaller differences between the Hawai'i island seed zones. Overall, these results suggest that there are genetic differences between all four main islands as well as within each island.

As an exploratory analysis of genetic structure in relation to eco-regions, we modified the seed zones on Hawai'i and Maui based on our genomic data. We used the SNP data to estimate genetic clusters using the Admixture program, and assigned individual trees on Hawai'i and Maui to the cluster from which they derived their highest ancestry. We recalculated the F_{ST} values using these

modified seed zones, and found a higher level of differentiation within each island using the modified seed zones compared to the seed zones defined by eco-region alone.

The seed zones we identified are preliminary, as further analysis is needed to better define the actual zones. However, our goal with both the eco-region seed zones and the eco-region plus genomics seed zones is to show that there is genetic differentiation between and within islands, especially within Hawai'i island. We suggest that further refinement of preliminary seed zones is necessary to account for the genetic differences, as the location of origin of koa trees within islands appears to be associated with genetic differentiation. This provides the basis for further koa seed zone development based on genetic characterization and environmental variation.

Looking forward, we suggest further developing seed zones for koa. In general, there are three types of data that can be used to define seed zones: (1) environmental or ecological data, such as climate and geography, (2) quantitative genetic data that identifies heritable traits, such as the information gained from common garden experiments, and (3) molecular genetic data that identifies DNA sequences, such as SNPs or microsatellites. Research in all these areas would help to refine koa seed zones across the state.

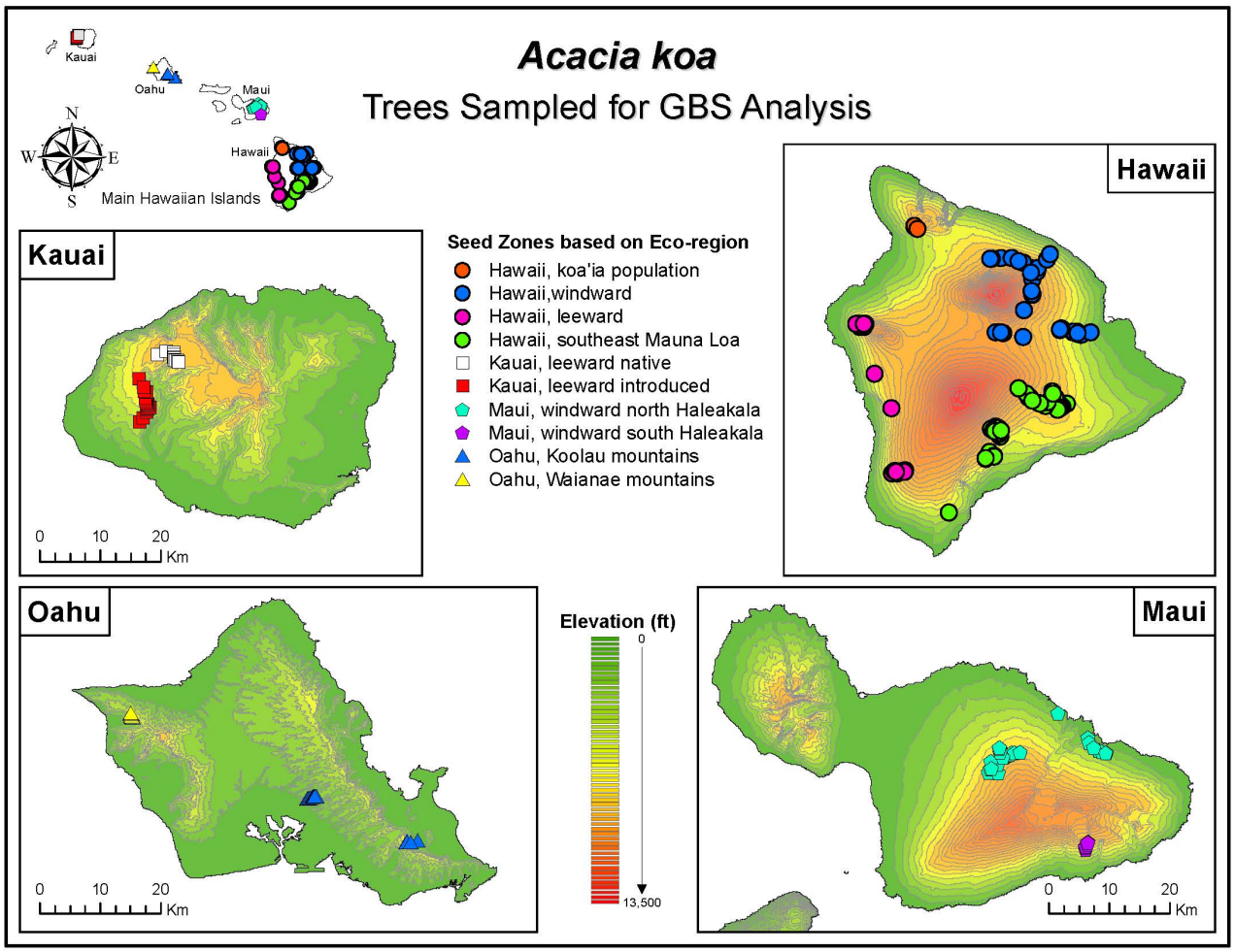

Figure 1: Preliminary *Acacia koa* Seed Zones based on Eco-region