

#### Slide 1:

Landscape genomics offers those managing landscapes forested with natural populations (undomesticated) a tool for addressing decisions regarding reforestation, habitat restoration, conservation and assisted migration. Difficult enough in the best of times, such decisions are complicated today by the anticipated changes of our global climates in the decades to come.

#### Slide 2:

Landscape genomics is a newly emerging discipline that combines the fields of population genetics and landscape ecology to study patterns of demographic and adaptive genetic variation across highly diverse landscapes. The science of landscape genomics is driven by the merger of rapidly developing technologies in both genomics and GIS (Geographic Information Systems). Wikipedia defines Landscape ecology as the science of studying and improving relationships between ecological processes in the environment and particular ecosystems. Genomics is the simultaneous study of an organism's complete genetic complement. With the ability to characterize the genetic basis of traits and map environmental variation better than ever before, the science of landscape genomics promises to bring an entirely new set of precise, diagnostic tools to the land manager. The drop-down circles seen in this photo imply that landscape genomics can inform us as to how plants have adapted to the variable landscape conditions. Before we proceed much further with the introduction of landscape genomics, it is worth taking a bit of time to reflect on the more traditional methods used to identify adapted planting stock, where tree improvement programs do not exist.

#### Slide 3:

While the science of landscape genomics is just developing, land managers have long sought to match planting stock to environment to insure their forests are thrifty. These techniques found favor because of rather bad experiences with maladapted planting stock. Today, foresters rely principally on seed zones for guiding their decisions on reforestation needs. Simply stated, stock used to replant a site in a given seed zone should have originated from that zone. Zones are roughly identified by geography, elevation, drainage, climate, and so forth. This approach is sure to provide planting stock with a generally accepted level of adaptation and performance, though does not address issues that could arise due to rapidly changing environments.

The provenance, or common garden trial, provides a more empirical means of identifying seed sources that perform well and are adapted to various environments. Though provenance trials first appeared several hundred years ago in Europe, they were only widely adopted in America by forest geneticists in the middle of the last century and remained popular for a relatively short period of time, from about 1960 to 1980, though their value and relevance remains still. A typical provenance trial may include seed collections from a few trees from each of many populations (e.g., 50 to 100 or more) across the species natural range. Seedlings are raised in a common nursery, and out-planted to replicated plantations established on many field test sites, again, distributed across the species range. In some cases, such trials are exported beyond the range, or even overseas. Species such as Sitka spruce, Douglas-fir and lodgepole pine, for instance, are now widely planted in the UK, Europe and Scandinavia as a result of such trials, having identified the best sources for the new environments.

The photos in this slide illustrate key features of provenance trials, namely scale and longevity. Taken in California, these photos feature sugar pine, one of the white pine group of trees under attack by the introduced white pine blister rust disease. The larger photo, though technically more progeny trial than provenance trial, gives us a sense of the scale or size of a provenance trial. They are large and require significant investments in establishment and maintenance costs over many years. The smaller photo was taken in a provenance trial of sugar pine that contained 124 sources, planted in row plots with 16 replicates. The trial occupied nearly 24 acres and included some 10K trees. Though the test itself was not replicated across multiple sites, many provenance trials are. One noteworthy lodgepole pine trial in British Columbia was planted on 60 sites. Provenance trials, in the absence of more highly developed tree improvement populations, are very effective at guiding seed source decisions for planting. However, the provenance trial leaves us largely clueless as to the genetic basis of the variation we observe in the traits of interest. That is, we know virtually nothing about how many genes control specific traits, what those genes do, the level of genetic diversity in the populations under study, and so forth.

#### Slide 4:

Regardless of policy or the experimental approach used to identify the most appropriate material for current and future conditions, the goal of a land manager is to insure that planting stock used for reforestation is adapted and grows well. The question then becomes, “How do we select the right material for outplanting?” The graphs shown here are useful for illustrating the potential power of landscape genomics tools. The data presented are hypothetical but reflect genetic differences commonly seen among trees in natural and select populations. In the left figure, breeding or seed zones for lodgepole pine are shown, color-coded, for the province of British Columbia in western Canada. Generally speaking, planting stock obtained from seed collected from lodgepole pine seed zones in British Columbia vary considerably in mean peak productivity, as measured by wood volume per hectare, as a function of the mean annual temperature of the planting site. Data such as these may be obtained from genetic tests or provenance trials, and are reflected in the graph at the upper right of this slide. In the absence of such resources, we rely on seed zone of origin to predict performance or adaptability, without the benefit of further genetic information. In the second (lower) graph, we learn much more about the genetic potential of our populations. The shaded areas reflect the hypothetical range of performance of the seed stocks (95% confidence intervals). From these figures we can deduce that seed stock from the TO (orange) seed zone retains a good deal more genetic variation than seed zones from further north. The pattern of declining genetic variability or diversity may reflect demographic influence – notably, the loss of variation as the species established itself following glaciation, or it could simply reflect the result of natural selection that intensifies as one moves north.

How does this knowledge help us with making planting decisions? It suggests that if we can identify the right trees from which to collect seed, we could improve productivity and diversity for both existing and potentially warmer environmental conditions. If, as predicted, mean annual temperatures (MAT) were to rise by 1 degree Celsius in the next 20 years, planting stock intended for the CP seed zone (blue) may need to be optimized for a MAT of 3.5 degrees. Our graphs suggest seed sources to meet those needs could come from either the PG (green) or TO (orange) seed zones, while not losing productivity. Random selection of seed from those other zones, however, could have undesirable results. So how do we garner such specific genetic information? Extensive field testing in provenance or genetic tests offers one alternative, landscape genomics another. Landscape genomics offers the potential to characterize

the genetic potential of thousands of trees across the landscape, within months or a few years, without the benefit of long-term field trials, though to be sure, to reach that lofty goal, the predictive power of LGs will require thorough validation. With time, it is anticipated that LG tools will improve in efficiency and utility for a wide range of tree species.

Slide 5:

To elaborate on the contrast between approaches to identifying adapted planting stock, we fall back to the classic equation used by forest geneticists to describe the relationships between phenotype, genotype and environment. The phenotype of a tree is what we can see or easily measure. It is determined by the individual tree's genotype at literally thousands of genes, and the environment in which the tree grows. Typically, phenotype has been the unit of selection. With provenance trials one seeks to identify associations between phenotypes and environments. For instance, it might tell you that trees from higher elevation, or more northern latitudes, in general, have later bud flush and are more cold hardy. Tree breeders seek to understand the association between genotype and phenotype. For traditional programs genotype is defined by pedigreed families. For intensive tree improvement programs that use genetic markers, breeders seek to reveal associations between specific genes, and their alleles, with specific phenotypes like cold hardiness. Landscape genomics, by contrast, seeks to identify associations between environments and genotypes, where genotypes are defined by knowledge of specific genes and their alleles. The premise of landscape genomics is that phenotypic traits are distributed according to environmental variation because those traits affect fitness. All approaches have utility, and may be used in complementary fashion, but come with different costs and constraints.

Tree improvement programs are costly, time and space consuming affairs that require large infrastructures and expertise. They are used for a relative few tree species and their products are deployed on highly managed plantations. Provenance trials are useful for species not under such intensive study, and provide wonderful insight into patterns of genetic variation across the landscape. Well managed trials can track changes in provenance performance across time and space, but these experiments are also very expensive to establish and maintain, and their value can slowly degrade, especially if institutional oversight is lost. For both tree improvement and provenance trials, costs are distributed out over the life of the project. Landscape genomics offers to provide basic information that should be valuable across time and space, and study of one species may lend itself to application in others. Costs for this approach are mostly front loaded. A comprehensive landscape genomics project could benefit from a small genetic field trial for validation of anticipated associations.

Slide 6:

Environmental factors like precipitation, temperature, soil chemistry and water holding capacity, insects and disease etc., determine which tree species, and individuals within species, survive and thrive in a given environment. They do so by acting on the natural genetic variation within the species. Trees with the best combination of genes will succeed and adapt, those with less favorable combinations will eventually perish.

Consider the scenario shown above, representing a timeline of 200 years. In this progression, a recently deforested landscape, scarred by fire, is naturally regenerated with a genetically diverse group of seedlings from one or more species of forest trees. Colors represent trees with various combinations of genes that confer more or less cold tolerance. In the beginning, they are relatively evenly distributed, as

one might expect from seed that flies into the valley on strong winds from nearby populations. Orange trees, with gene combinations that confer cold tolerance, eventually become dominant at higher elevations while green trees, lacking cold tolerance, become dominant at lower elevations. Landscape genomic approaches can identify those suites of genes that confer adaptation to the local environment.

Slide 7:

The translational goal of a landscape genomics research project would be to develop a set of web-based tools that land managers could use to incorporate adaptive genetic potential into reforestation, restoration, and conservation plans. Land management decisions on public lands are the responsibility of a broad spectrum of individuals with diverse training and expertise. Annually, land managers are faced with myriad decisions regarding land use. Among the most important decisions is the selection of source materials for reforestation efforts that follow timber harvests or natural disturbances such as wildfires and insect mortality. Identifying the appropriate species and sources within species to use for reforestation to insure optimal or even acceptable adaptability to site conditions and changing climates is an annual task. The genetic characterization of native and improved populations of selected conifers in a given region should provide improved precision in source assignments. Historically, foresters have relied on the simple guideline of using planting stock that originated from the area or seed zone that needs reforesting. With changing environments and climates, this practice may not confer optimal adapted stock in the long run. Genetic characterization may find even greater relevance to managers concerned with ecosystem restoration or genetic conservation efforts, where detailed information on the underlying genetics of species selected for use is critical.

Slide 8:

The scientific goals of a landscape genomics project can be generally described by noting major steps in the experimental process. The first step in the process is to characterize the genome of the species of interest. That is, identify all the genes, their putative function, and their genic or allelic diversity. This process, though seemingly impossible only a few years ago, is becoming increasingly tractable, both technically and financially. Analysis of these data will allow one to identify those genes under selection from environmental factors, both biotic and abiotic in nature. The next step is to characterize the spatial distribution of allelic diversity across the landscape and lastly, test for associations between tree genotypes and environmental traits tree genotypes and phenotypes. The results of these analyses may be used to predict or model which trees or populations are best suited for placement in areas requiring reforestation or restoration.

In the next few slides we try to describe the process in greater detail, and look at examples of where landscape genomics might serve to aid resource managers. Before embarking down this path, however, it might benefit the viewer to very briefly review the basic elements of genetics, the underlying basis of this technology.

Slide 9:

To fully understand the basis for landscape genomics, one needs to at least recognize the nature of genetic diversity. No two trees are alike. They may differ from one another in literally millions of places in their genome. If these differences occur within the part of the DNA that codes for a gene product, like a protein, there may be differences in how well that protein does its job. Genes are arrayed on

chromosomes. Chromosomes occur in pairs within an individual, one coming from the maternal parent, the other from the paternal parent. These pairs are called homologous chromosomes, since they contain DNA code for the same sets of genes. Every cell in an organism (except for germ cells) contains two copies of each gene. The term allele is used to describe the genetic nature of a gene. An individual may have two identical forms or alleles of a given gene, or they may differ at one or more points in the DNA sequence of that gene. These differences are the basis upon which natural selection operates. Different alleles may confer different adaptive states to the individuals. In a population of individuals there may be many allelic forms for a given gene, but any individual will only have two of them. A gene with multiple alleles is referred to as being polymorphic. In landscape genomics we seek to identify associations between allelic differences and environmental traits.

Slide 10:

Some traits, in trees, humans, or other species, are encoded by one or a very few genes. Sickle cell anemia in humans is a good example. However, almost all traits of interest in trees are controlled by many genes. Size is a classic quantitative trait, regardless of species. How many genes contribute to human height for instance? What might those genes control? Hormones (growth regulators), muscle and bone development, metabolic rate, etc? In forestry we are also pre-occupied with the genetics of growth, commercially because it drives profit, and biologically because it is likely the best way to measure overall fitness. We have learned through time that growth, and the component elements thereof, are controlled by numerous genes, each of which influences the trait of interest in very small increments. In landscape genomics we seek to identify which genes influence a trait or confer an environmental preference, which alleles in those genes are most desirable for the conditions observed on the ground, and how much each gene contributes to the trees adaptation to those conditions.

Slide 11:

Landscape genomic studies require considerable logistical planning and are composed of many steps, some of which are outlined here. Each step is likely to be conducted by collaborators with specific skill sets, though ultimately, the product of the effort should be usable by the parties making decisions about what genetic materials should be deployed on the landscape.

Step One: The process begins by defining the tree species and geographic regions of interest. This might involve a single species and national forest, or several species over an entire floristic region (i.e., all five needle pines in the Sierra Nevada Range).

Step Two: Once defined, the work on the ground begins with selection of study trees, collection of tissues from which DNA will be extracted, measurements of physical traits of interest, and geo-referencing of the tree location using GPS technology.

In step three, DNA is extracted from the selected trees and subjected directly to sequencing or genotyping of known genetic polymorphisms.

Step 4, which can occur concurrently with earlier steps, involves collection of environmental data such as mean annual temperature, mean annual precipitation, elevation, aspect and so forth, for the locations of all study trees. Generally speaking, this information is available or can be extrapolated from existing databases.

The next step (5) involves identifying the genes that are under natural selection pressure. This may require a bit of sophisticated analysis of genetic data obtained from the sequencing step, or, as will occur increasingly through time, as we accumulate knowledge, from previous studies. Arguably, this step could come after association testing, given that most significant statistical associations are likely to occur only with genes under selection. In practice, it is likely that all genes and allelic variants will ultimately be tested regardless of order of activities.

In step 6, we seek to identify potential associations between specific alleles and environmental characteristics.

Finally, in step 7, we take all known associations and build a model that can guide selection of seedling lots that are genetically predisposed to survival under specific environmental conditions.

Slide 12:

Land managers responsible for insuring the long-term health of our forested ecosystems face many challenges when confronted with reforestation, conservation and restoration issues. These challenges are magnified by the impacts of global climate change and a century of fire suppression on our forest vegetation. The stunning changes in vegetation types in the central Sierra Nevada region of California reflected in these side by side illustrations emphasize the scope of challenge facing land managers. The top panel illustrates vegetation types in 1920-30 and the bottom panel for 2005. These maps show how the elevation range of the ponderosa pine belt (light green) is retracting upslope, with corresponding expansion of oak woodlands (purple), likely due in part to climate change. Pine is also being replaced at upper elevations by the expansion of the Sierran mixed conifer belt likely because of effective fire suppression. So how do we decide how to reforest major disturbances that follow fire, wind-throw or beetle killed stands? Were the trees that were there to begin with optimally adapted to current conditions, or were they historical remnants of past climates or stand management choices? Landscape genomics will likely provide some guidance in addressing such questions.

Slide 13:

Increasing regional and global temperatures, in some cases coupled with extended droughts, have had dramatic effects on some forest types in the Western United States and Canada. In addition to the physical stress of drought on forest trees, rising temperatures have allowed beetle populations to explode. Coupled with over-stocked stands, the effects have been devastating, as captured in the photos shown here. In the side by side photos of a pinyon pine forest, shown at right, over 90% of the pine trees were lost as a result of an extended drought (2000 – 2003) and subsequent beetle kill. The other photos show white bark pine die-off in the Wind River Mountain and Sierra Ranges of Wyoming and California, respectively. In virtually all cases mortality is not complete. What traits and genes characterize the surviving individuals? Those are the types of questions landscape genomics attempts to address. Consider the example in the next slide.

Slide 14:

In 2010, Eckert et al. published a paper that described how their team identified an array of genes that appeared to be strongly associated with bioclimatic variables. The team gathered climate data from the WORLDCLIM 2.5-min geographical information system (GIS) layer and overlaid this on distribution maps of over 900 loblolly pine trees with known genotypes for over 3000 genes. Monthly minimum and

maximum temperatures, monthly precipitation, and 19 other bioclimatic variables were evaluated. The temperature and precipitation data were used to estimate potential evapotranspiration (PET) and an aridity index (AI) was defined as the ratio of precipitation to PET. This was done on a quarterly basis. The pattern of aridity is illustrated here for the second quarter, across the Southeastern United States.

Environmental association analysis identified five genes associated with aridity gradients. The primary functions of gene products encoded by these loci were abiotic and biotic stress responses, as determined in previous studies. The relationships between genotypes and aridity are shown here for four of the genes. The genotype is defined here by letters on the x-axis of the graphs. The letters designate the combination of alleles that an individual tree possesses at a given gene. Take for example the upper left hand graph, showing the gene identified as a photosystem II protein. Of the 900+ trees in the study, some proportion of them had two C alleles (CC), some, one A and one C allele (AC) and some, two A alleles. Collectively those groups occurred in areas of differing aridity, implying they contributed to the adaptation of those trees to those sites.

Studies such as these are very informative and can help guide decisions about what trees might do best in difficult environments or future climates. Knowledge of which genes are associated with aridity in this species can likely inform which genes should be evaluated in other species for the same trait.

Slide 15:

Landscape genomics offers the potential for land managers to fine-tune their decision making process with regards to reforestation, habitat restoration, and conservation. The knowledge gained from the study of forest tree genomes, coupled with enhanced understanding of our changing environments and landscapes, will guide decision making for seedling deployment in a flexible, sustainable manner for broad geographic regions. Common garden trials can, and probably should be employed to validate landscape genomic prescriptions, though their drawbacks must be weighed carefully. Common garden trials are expensive and time-consuming to establish and must be maintained for up to 20 years or more. In situations where planting materials are drawn exclusively from seed lots obtained from local seed zones, landscape genomic tools could significantly improve survival and productivity in a changing landscape.