NC STATE UNIVERSITY

DEPARTMENT OF Forestry and Environmental Resources

Identifying Genetic Variation in Site Adaptability in Loblolly Pine

Ross Whetten and Laura Townsend
Department of Forestry & Environmental Resources
North Carolina State University, Raleigh NC

Abstract

- The objectives of Aim 3 are
  - To develop deployment guidelines for existing germplasm,
  - To identify alleles that influence pine responses to environmental signals in pine populations, and characterize the frequency and distribution of those alleles in deployment populations
- A key requirement to accomplish the second objective is a cost-effective high-throughput genotyping method for loblolly pine
- An important deliverable for the first two years of the project is to develop such a method and provide convincing evidence of its quality, reproducibility, and cost-effectiveness.
- Four groups in Aim 3 are working to test different alternative approaches that may help to achieve this key milestone

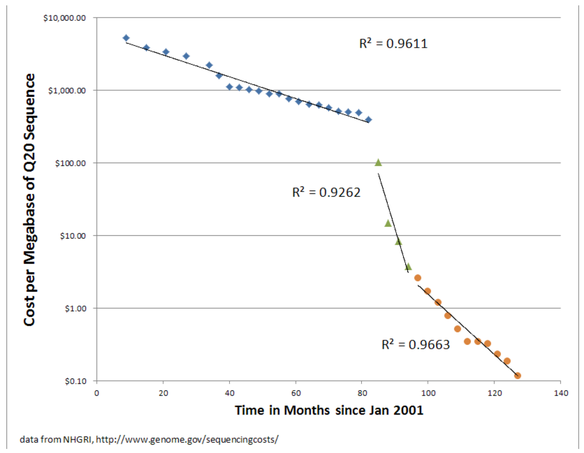
Introduction

- Existing high-throughput genotyping methods for pine are too expensive to allow routine application in breeding populations
- Incorporation of marker genotyping as a routine part of applied tree breeding programs will require a system that is robust, cost-effective, and easy to implement in terms of sample preparation and data analysis.
- The yield of data per unit cost produced by high-throughput sequencing platforms is increasing at an exponential rate, so genotyping methods based on sequencing will become more cost-effective over time rather than more expensive (Stein, 2010 – see figure below)

Research Objectives

- Develop a cost-effective high-throughput genotyping method for pine
- Genotype the available parents and a sample of progeny from different families (genetic entries) in the 7 provenances, growing at different sites
- Test for associations of individual genetic marker loci with growth or quality characteristics using climate factors as covariates

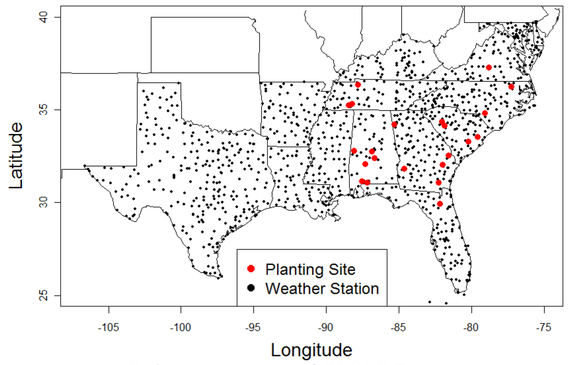
The Cost of DNA Sequencing



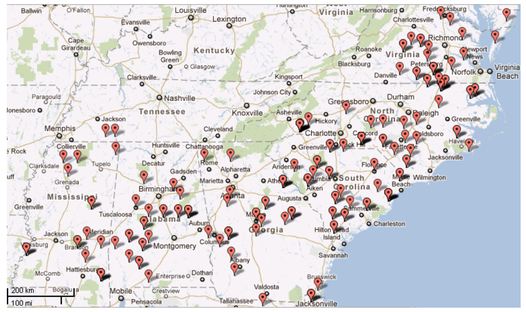
Plantation Selection Seed Source Study (PSSSS)

- Designed to determine the pattern of geographic variation in plantation selections and to assess pine genotypes by environment interactions
- Wide range of genetic entries that represent the whole pine population – 20 female parents and 40 male parents from each of seven regions; a total of 140 pollen-mix families plus checklots
- Seedlings from all families were planted at 20 test locations throughout the southeastern US; 2 additional sites have a subset.
- Growth, stem quality and fusiform rust disease were measured at tree ages 4 and 8 years

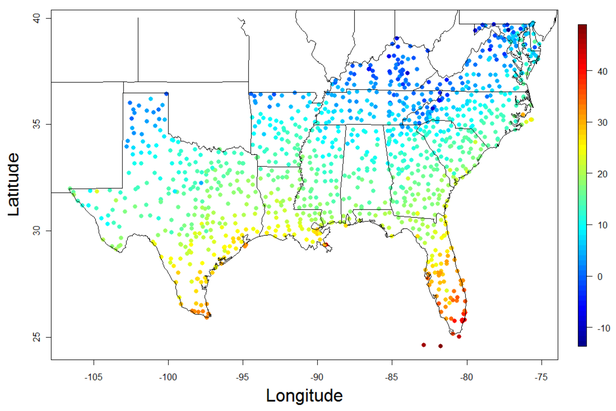
Location of PSSSS Trials



Selection sites of PSSSS Parents



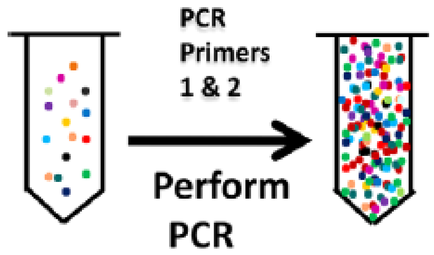
Annual Average Minimum Temperatures (1994-2005)



Methods

- Genotyping-by-sequencing was described for maize and barley by Elshire et al. (2011), and for wheat and oats by Poland et al. (2012).
- The modifications to the original protocol reported by Poland et al. (2012) include the use of two restriction enzymes rather than one, and a more diverse collection of variable-length barcodes to identify the origin of DNA sequences.
- In brief, a set of DNA samples are arrayed into wells of a 96-well plate along with oligonucleotide adapters, where each well contains an adapter bearing a different barcode (shown as different colors in the figure below, from Elshire et al (2011)).

- After digestion with restriction enzymes PstI-HF (CTGCA^G) and MspI (C^CGG), the DNAs are ligated to the adapters in their separate wells of the plate
- Following completion of the ligation step, the samples are pooled into a single tube, purified on a Qiagen PCR-quick column to remove non-ligated adapters and other small fragments, and then used as template for 8 parallel and identical PCR reactions



References

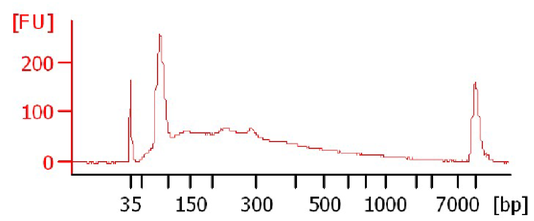
Elshire RJ, et al. 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS One 6:e19379.

Poland JA, et al. 2012. Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. PLoS One 7:e32253.

Stein LD. 2010. The case for cloud computing in genome informatics. Genome Biol 11(5):207.

Methods, continued

- After the PCR step, products are purified by adsorption to carboxylate-coated magnetic particles, and the resulting mixture is separated on a Bioanalyzer chip to see the distribution of fragment sizes



- When a high-quality library is obtained, sequencing is carried out using Illumina GAIIx or HiSeq2000 instruments. The HiSeq2000 is the current platform of choice due to the number of reads and the read quality obtained.
- The resulting sequence data are analyzed using Java-based, platform-independent software tools obtained from the Institute for Genomic Diversity at Cornell University.

Preliminary Results

- An initial experiment was conducted using the method of Elshire et al (2011), which differs from that described in the methods section in that only a single restriction enzyme (PstI) is used for the digestion of genomic DNA.
- 144 million paired-end reads were obtained from a multiplex library containing 93 different samples
- 136.5 million reads (>94%) contained barcodes and usable sequence data
- Over 17,000 putative allelic pairs of sequence tags were identified in loblolly pine samples, but only about 5% of these showed adequate reproducibility across samples and patterns of inheritance consistent with single Mendelian loci

Acknowledgments

- We thank Ed Buckler, Jeff Glaubitz, Terry Casstevens, Fei Lu, and James Harriman for making the GBS analysis pipeline software available, and for their guidance in data analysis.
- We thank the members of the NCSU Cooperative Tree Improvement Program, as well as past and current faculty and staff of the Program, for their work to establish and maintain the Plantation Selection Seed Source Study, and for making that study available for use in this project.
- We thank Will Kohlway for his capable technical assistance in getting the GBS protocol working to produce data from pine.