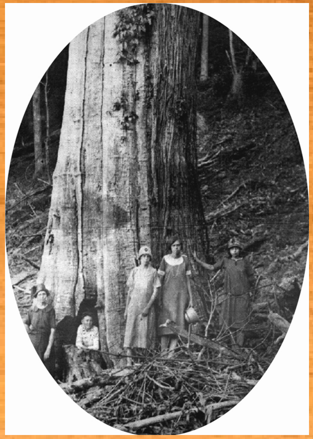
Fagaceae Genomics and American Chestnut

Ronald Sederoff
Forest Biotechnology
North Carolina State University



NSF Genomic Tool Development for the Fagaceae

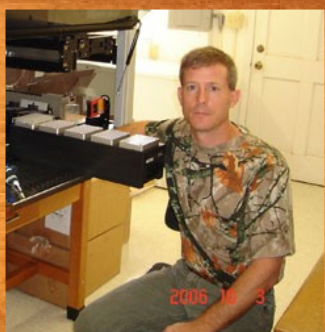
Who we are

Ron Sederoff, NCSU

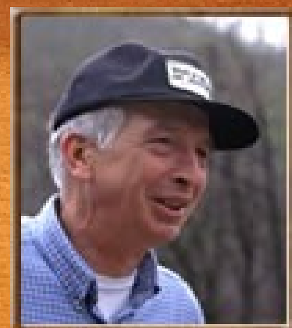
John Carlson, Penn State

Jeff Tomkins, Clemson

Bill Powell, SUNY-ESF



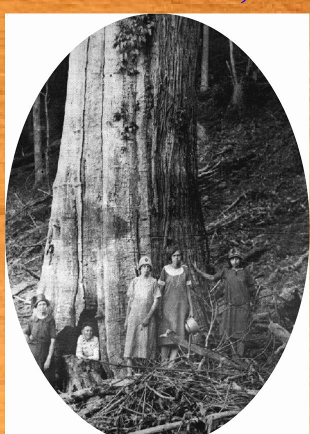
Tom Kubisiak, SIFG



Fred Hebard, TACF

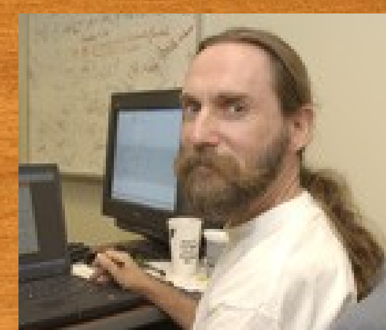
Sandra Anagnostakis, CAES

Nick Wheeler, NCSU



Dahlia Nielsen NCSU

Paul Sisco, TACF



Chris Smith, NCSU

Comparative Genomics.
The Fagaceae (beech family) contains 7 genera and ~1000 species.
Main interest: Beech, chestnut and oak

The species we study are:

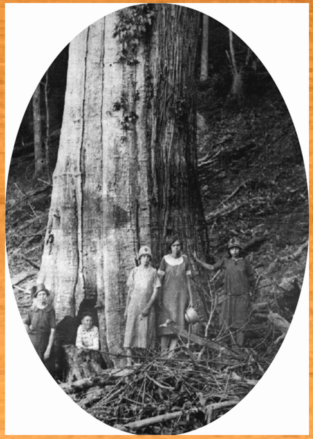
Chinese chestnut: *Castanea mollisima*

American chestnut: *Castanea dentata*

Northern Red oak: *Quercus rubra*

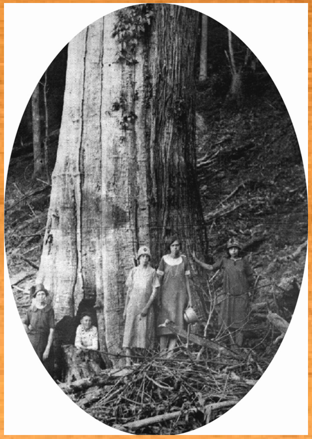
White oak: *Quercus alba*

American beech: *Fagus grandifolia*



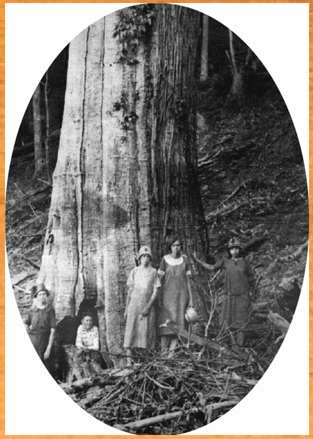
Some research objectives

- Gene discovery
- Genetic marker development
- Large new mapping populations
- New genetic maps
- Physical maps
- Comparative genomics
- Web site (Clemson).



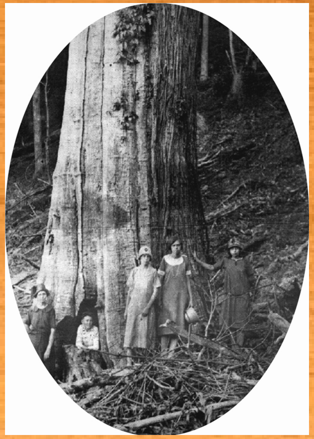
DNA Sequencing

- All EST based: cDNA libraries from different species and tissues. (Powell)
- Traditional ABI sequencing (Tomkins).
- 454 sequencing. 2 million reads total on 5 species (Carlson).
- Annotation (Smith).



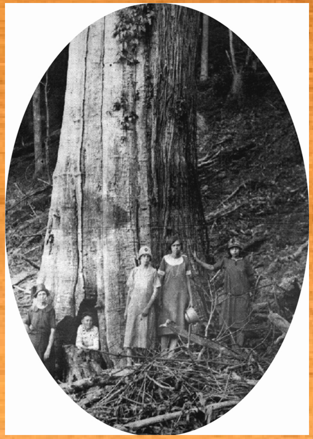
Gene discovery
Predicted proteins

- There are 10,562 of hits to Populus
  - (at e-15).
- Of these, 8266 also hit Arabidopsis.
- Many hits are known only by their "gene model" predictions.
- So our data verify their predictions as real genes.

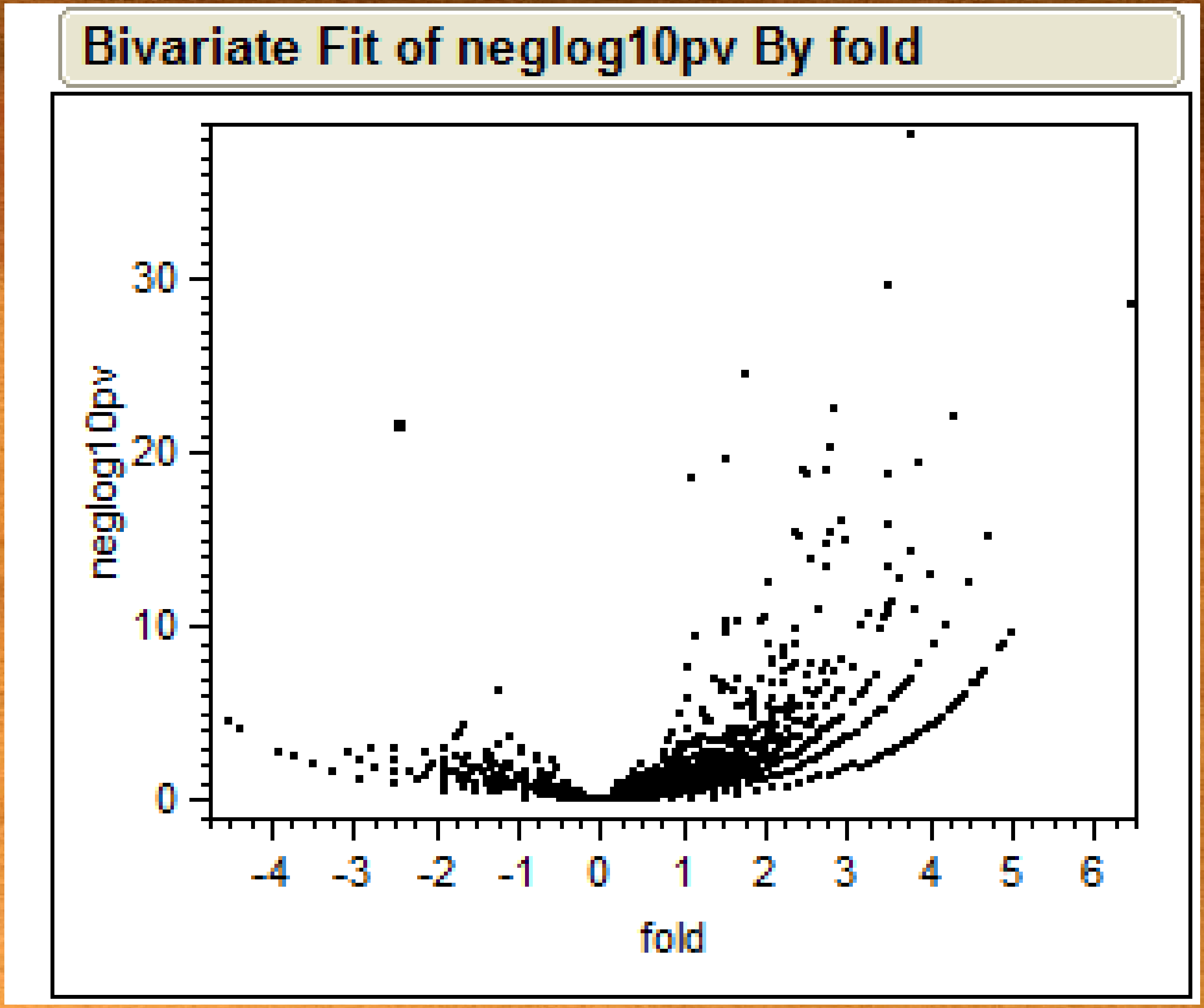


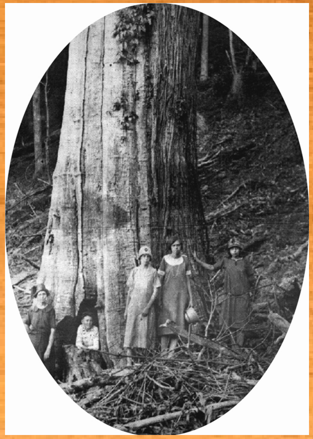
Genes Identified by species

Species:	Predicted Proteins
Chinese chestnut	7719
American chestnut	1067
Red oak	4985
White oak	3525
Total Fagaceae	10,562



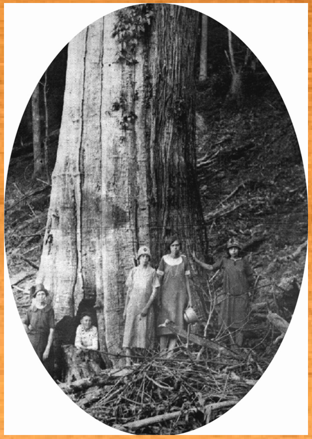
Expression data: electronic northerns





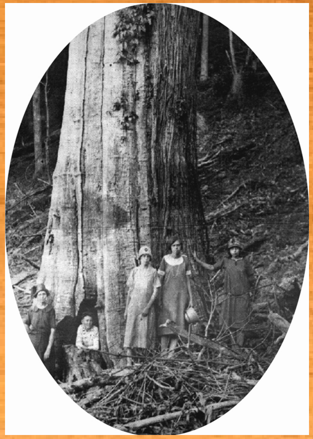
Marker Development
(Kubisiak and Smith)

- Many more flexible informative markers.
- Genotyping
- Solving mapping problems
- Comparative mapping
- Using 454 EST sequence data to find SSRs and SNPs.
- Anchoring ESTs to genetic and physical maps.



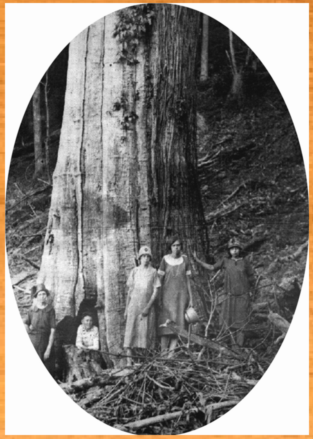
Comparative marker development

- Survey sequence data for SSRs and SNPs for polymorphisms that vary within species, and between species.
- Find Chinese vs American species specific markers to aid mapping in hybrid backcrosses.
- Explore variation within Fagaceae, such as comparative mapping.



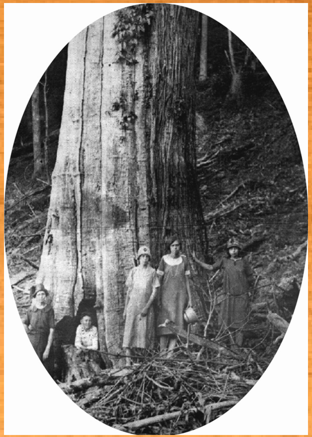
Looking for SSRs

- Looking within species for SSRs in 454 contigs.
- Ran scripts to search for repeated patterns and variation in the repeats within the contig.
- Found 104 in the first pass and have successfully tested a third (so far).
- SSR genotyping has turned up some misidentified parents.



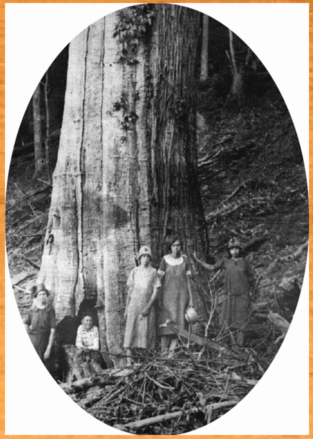
Looking for SNPs

- BLAST against Populus transcripts (N).
- Group all reads that match the same gene.(e-10 cutoff).
- Toss out genes with less than 4 reads to get 12,158 genes.
- Assemble the reads for each gene (PHRAP).
- Run PolyBayes on each assembly to find SNPs.
- Parse output for SNPs with >0.95 probability.



SNP frequency

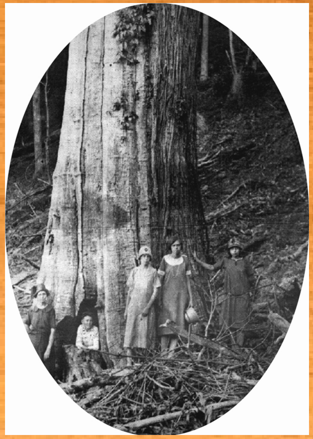
- Total SNPs found by PolyBayes: 17,748.
- Substitutions: 15,563.
- Deletions: 2,185.
- Number of aligned bases screened: 10,153,440.
- Potential SNP frequency: 1/ 572.



Genetic Mapping Objectives

1. Genetic mapping in parent Chinese and American chestnut and derived hybrids.
- Looking for "species specific" markers to track genome segments in hybrids.
- Solve some problems with distortion and linkage (missing LG-B).

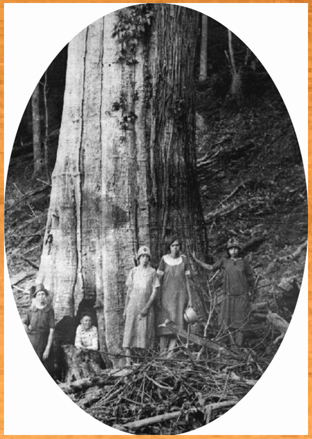
4. QTL mapping for resistance to chestnut blight to confirm a three gene model.



Problems with Current Genetic Maps (Sisco, Kubisiak)

Comparison of European and AC/CC hybrid maps found one linkage group missing: LG-B

LGs "B" and "E" could easily form a single linkage group in the BC_1 map.

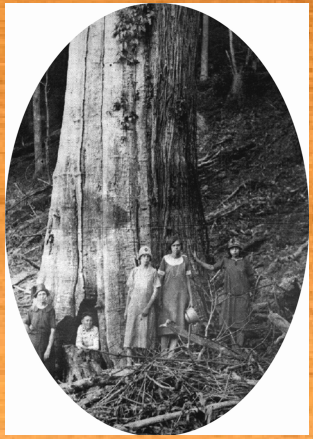


High segregation distortion in both the F_2 and BC_1 maps. (20 to 31%).

Translocations between American and Chinese hybrids could affect breeding and map based cloning.

Identification of resistance genes

- Confirm or extend the three QTL model.
- Use the BC3 populations to “fine map” the QTLs by association in very large populations.
- Identify genomic regions associated with resistance.
- Need resistance to Phytophthera too.



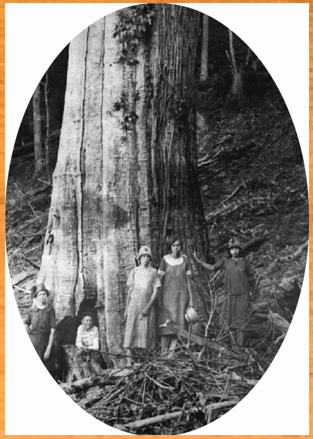
Mapping focus is in Chinese chestnut because that is where the resistance genes are.

To facilitate map based cloning of resistance genes, we need:

A high resolution genetic map.

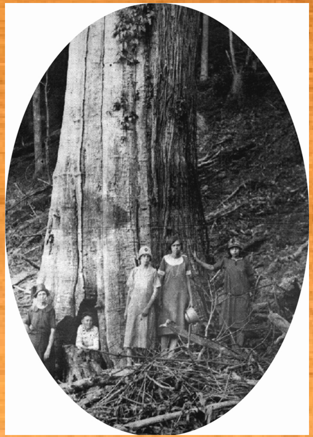
A high quality physical map (BACs)

And to integrate them.



Physical Mapping (Tomkins)

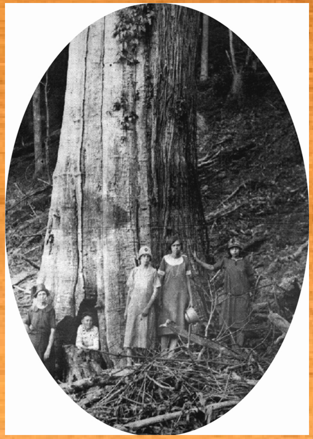
- Construction of 20X BAC libraries of Chinese chestnut.
- Construction of a physical map for Chinese chestnut by fingerprinting and BAC end sequencing.
- Integration of the genetic and physical maps.
- Investigate micro and macrosynteny in the chestnut and oak.



5. Platform for map based cloning.

The Fagaceae Website and Database

- Physical mapping data.
- Genetic mapping data.
- Sequence data.
- Marker development data
- Integration of physical and genetic maps.
- All reports and presentations by project members.



A different kind of forest biotechnology

- This project shows how genome technology could contribute to rescue endangered tree species.
- Substantial social benefits are possible.
- Model for future efforts to conserve and protect endangered tree species through genetics and biotechnology.