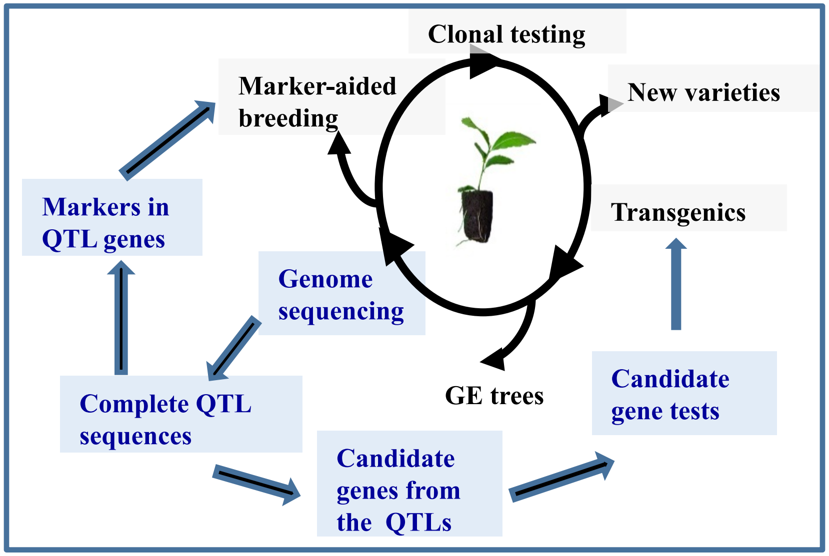
The Genome Resources and Tools Project
FHI 2012 Annual Meeting

August 9, 2012

John Carlson
The Schatz Center for Tree Molecular Genetics
Pennsylvania State University



Clonal testing

Marker-aided breeding

New varieties

Markers in QTL genes

Transgenics

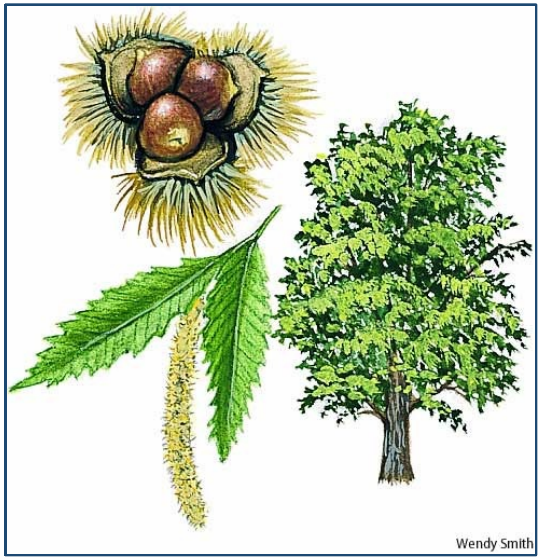
Genome sequencing

Complete QTL sequences

GE trees

Candidate gene tests

Candidate genes from the QTLs

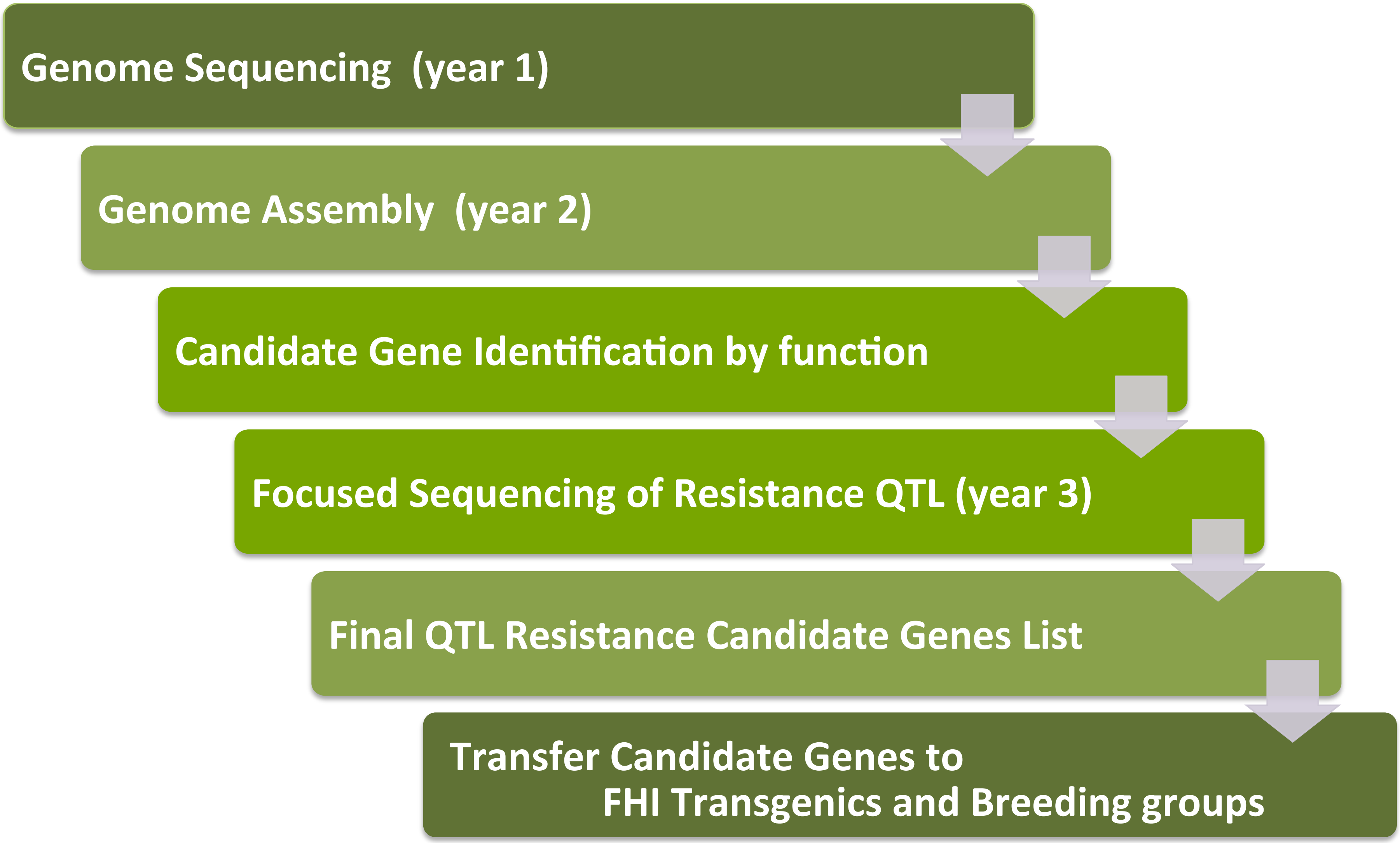

Wendy Smith

PENNSTATE College of Agricultural Sciences School of Forest Resources

CLEMSON UNIVERSITY

Forest Health Initiative
Advancing Forest Health through Biotechnology

The Genome Resources and Tools Project
General Approach



Genome Sequencing (year 1)

Genome Assembly (year 2)

Candidate Gene Identification by function

Focused Sequencing of Resistance QTL (year 3)

Final QTL Resistance Candidate Genes List

Transfer Candidate Genes to
FHI Transgenics and Breeding groups

The Genome Resources and Tools Project

Genome Sequencing (year 1)

Year One Deliverables were Achieved

1. The reference *Castanea mollissima* cv. Vanuxem genome was sequenced to over 25-fold depth.
2. Preliminary *de novo* assemblies of the reference genome sequence were conducted.

Genome Assembly (year 2)

Year Two Deliverables Achieved

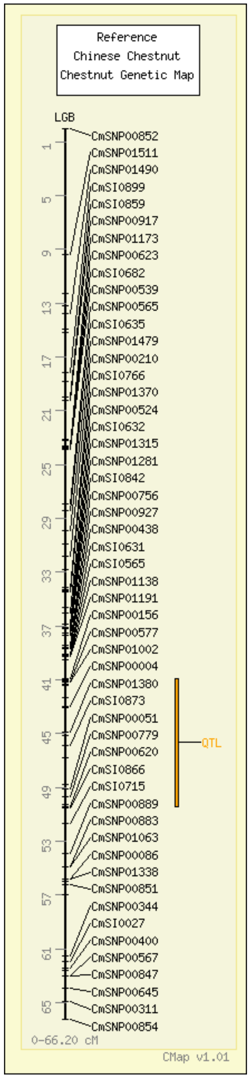
1. **Produced additional genome sequence.**
2. **Completed assembly of the genome sequence.**
3. **Identified over 60,000 genes (reflects heterozygosity)**
4. **Conducted candidate gene identification:**
   - ➢ **97% of transcripts aligned to genome assembly**
   - ➢ **66,662 gene models predicted**
   - ➢ **Candidate gene coding sequences delivered to transgenics team**
4. **Strategy for resistance gene discovery updated.**
   - ➢ **QTLs to be sequenced in depth**

Sequencing of Resistance QTL (year 3), APPROACH:

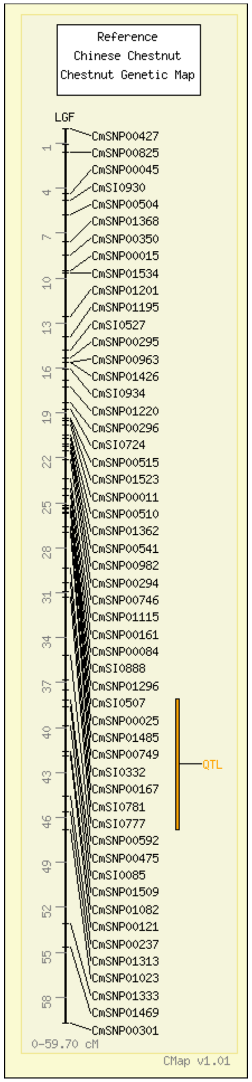
- Genetic Map (QTL markers)
- Physical Map (QTL rDNA clones)
- BAC DNA pools (covering QTLs)
- Sequencing of QTL DNA pools
- Identification of Candidate Genes for Blight Resistance in QTL

The Genome Resources and Tools Project

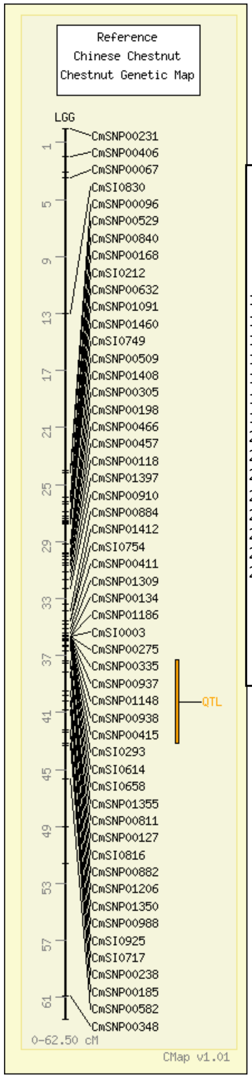
QTL cbr1
on LGB



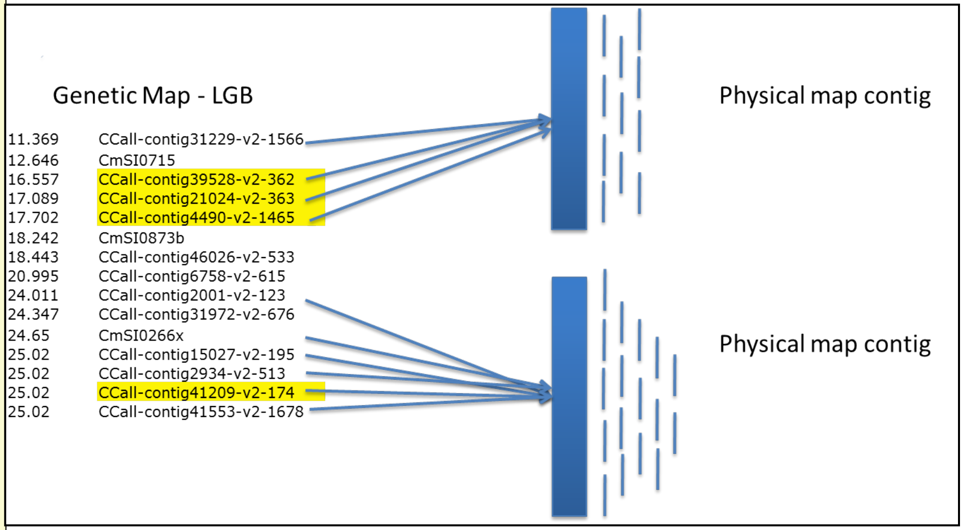
QTL cbr2
on LGF



QTL cbr3
on LGG



Individual QTL sequencing approach



Genetic Map - LGB

11.369	CCall-contig31229-v2-1566
12.646	CmSI0715
16.557	CCall-contig39528-v2-362
17.089	CCall-contig21024-v2-363
17.702	CCall-contig4490-v2-1465
18.242	CmSI0873b
18.443	CCall-contig46026-v2-533
20.995	CCall-contig6758-v2-615
24.011	CCall-contig2001-v2-123
24.347	CCall-contig31972-v2-676
24.65	CmSI0266x
25.02	CCall-contig15027-v2-195
25.02	CCall-contig2934-v2-513
25.02	CCall-contig41209-v2-174
25.02	CCall-contig41553-v2-1678

Total number of Genes Identified in the Blight Resistance QTLs

QTL	LG	# Genes	# bases / QTL	# Genes / Mbases	# Stress-response genes
cbr1	**B**	**994**	10.14 Mbases	151	98
cbr2	**F**	**548**	5.13 Mbases	137	58
cbr3	**G**	**410**	4.51 Mbases	139	38

Final QTL Resistance Candidate Genes List

QTL Cbr1 genes

l-ascorbate peroxidase	Peroxidase activity; Response to oxidative stress
peroxidase 17-like	Peroxidase activity; Response to oxidative stress
peroxidase	Peroxidase activity; Response to oxidative stress
cc-nbs-lrr resistance protein	Peroxidase activity; Response to oxidative stress
ribonuclease h-like protein	Defense response; ADP binding; DNA binding
disease resistance protein	Defense response; ADP binding; DNA binding
cc-nbs-lrr resistance protein	Defense response; ADP binding
adenine nucleotide alpha hydrolases-domain containing protein kinase	Response to stress
hypothetical protein [Vitis vinifera]	RNA binding; DNA binding; RNA-directed DNA polymerase activity; DNA replication; DNA repair

Final QTL Resistance Candidate Genes List

QTL Cbr2 genes

cc-nbs-lrr resistance protein	Defense response; ADP binding
cc-nbs-lrr resistance protein	Defense response; Protein binding; ADP binding; DNA binding
disease resistance rpp8-like protein	Defense response; ADP binding
tir-nbs-lrr resistance protein	Signal transduction; Defense response; Protein binding; ADP binding
tir-nbs-lrr resistance protein	Signal transduction; Defense response; Protein binding; ADP binding
tir-nbs-lrr resistance protein	P:defense response; F:ADP binding
tmv resistance protein –like	Signal transduction; Defense response; Protein binding; ADP binding
tmv resistance protein –like	Signal transduction; Defense response; Protein binding; ADP binding
tmv resistance protein –like	Signal transduction; Defense response; Protein binding; ADP binding
tmv resistance protein -like	Signal transduction; Defense response; Protein binding; ADP binding

Final QTL Resistance Candidate Genes List

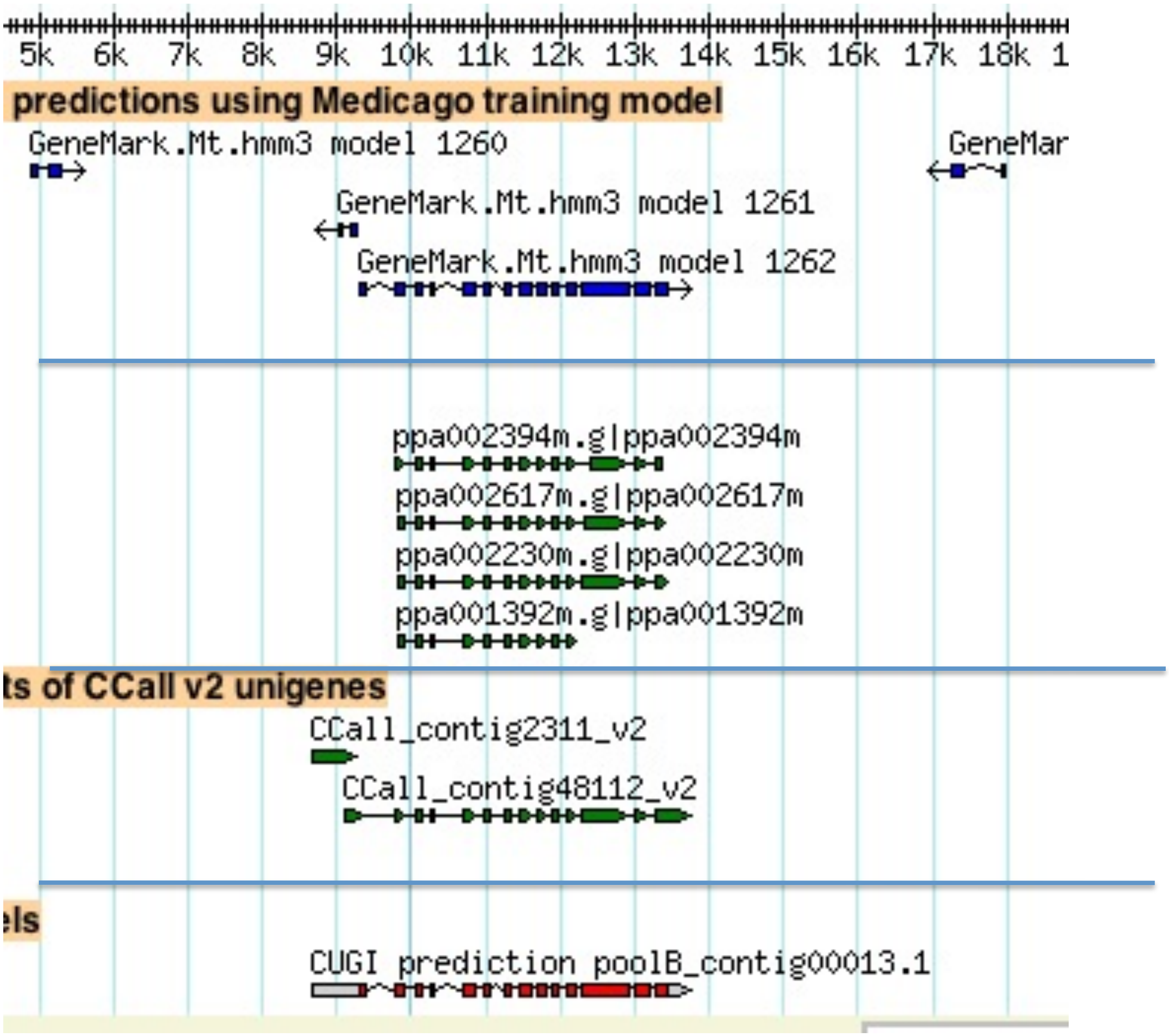
QTL Cbr3 genes

ATP binding	Response to stress; ATP binding; Protein phosphorylation; Protein kinase activity
BED finger-nbs-lrr resistance protein	Defense response; ADP binding
Cationic peroxidase 1	Peroxidase activity; Response to oxidative stress; Oxidation reduction
N-glycosylase DNA lyase	DNA repair; Damaged DNA binding

The Genome Resources and Tools Project

Transfer Candidate Genes to FHI Transgenics and Breeding groups

- Have complete genes from Chinese chestnut including regions regulating gene expression
  - This will enable cis-transgenics approach
- Also have many genes differing in expression patterns in cankers in Chinese vs American chestnut
- Developing markers in candidate genes for marker-assisted breeding by resequencing additional chestnut genotypes


5k 6k 7k 8k 9k 10k 11k 12k 13k 14k 15k 16k 17k 18k 1
predictions using Medicago training model
GeneMark.Mt.hmm3 model 1260
GeneMar
GeneMark.Mt.hmm3 model 1261
GeneMark.Mt.hmm3 model 1262
ppa002394m.g|ppa002394m
ppa002617m.g|ppa002617m
ppa002230m.g|ppa002230m
ppa001392m.g|ppa001392m
ts of CCall v2 unigenes
CCall_contig2311_v2
CCall_contig48112_v2
els
CUGI prediction poolB_contig00013.1

Continuing Research and Future Plans

- Resistance Candidate Gene development
  - Full length candidate genes from BAC clones
  - Marker discovery by resequencing
    - Markers in disease-resistance candidate genes
    - Marker genes that distinguish between the genomes of Chinese chestnut and American chestnut, for selection
  - Comparative studies with peach and poplar QTLs
- 15X more genome sequence being produced
- New, final publication-quality genome assembly soon
- 'Jamboree' of genomics experts will be called for
  - Producing best possible genome and QTL assemblies
  - Identifying the best set of blight resistance genes

The Chinese Chestnut Genome Project Team

John E. Carlson, PI, Schatz Center, Penn State

DNA Sequencing

Stephan C. Schuster, Professor of Biochemistry and Molecular Biology, Penn State

Lynn P. Tomsho, Daniela Drautz, and Lindsay Kasson - Sequencing Specialists, Penn State

Tyler Wagner and Nicole Zembower, Research Assistants, Penn State

Bioinformatics and Comparative Genomics

Webb Miller, Professor of Biology and Computer Science & Engineering, Penn State

Charles Addo-Quaye, Postdoctoral Fellow, Schatz Center, Penn State

Meg Staton, Stephen Ficklin, Christopher Saski - Bioinformatics team at CUGI

Bert Abbott, Professor Emeritus, Clemson University

Abdelali Barakat, Research Associate, Clemson University

Other FHI Collaborators: Sandra Anagnostakis, Kathleen Baier, Joe James, Nurul Faridi, Stephen Ficklin, Fred Hebard, Thomas Kubisiak, Charles Maynard, Scott Merkle, C. Joseph Nairn, William Powell, C. Dana Nelson

PENNSTATE College of Agricultural Sciences School of Forest Resources

CLEMSON UNIVERSITY

Forest Health Initiative
Advancing Forest Health through Biotechnology