GENETIC VARIATION IN NATURAL POPULATIONS OF AMERICAN CHESTNUT

Thomas L. Kubisiak and James H. Roberds

USDA Forest Service, Southern Research Station, Southern Institute of Forest Genetics, Saucier, MS

Prior to the blight epidemic, American chestnut (*Castanea dentata* Borkh.) was one of the most important timber and nut-producing tree species in eastern North America (U.S. Census Bureau 1908). Its native range extended from southern Maine and Ontario in the north to Georgia, Alabama and Mississippi in the south (Sargent 1905). It now exists primarily as stump sprouts across its entire native range. After nearly a century of blight, numerous living stems of American chestnut still exist (Stephenson et al. 1991). Prolific stump sprouting and the fact that the blight fungus does not infect the root system have enabled American chestnut trees to persist. However, sexual reproduction is infrequent and its gene pool will likely face serious erosion when old root systems fail to produce sprouts and perish.

In an attempt to restore the American chestnut to its former status as a dominant canopy component in eastern forests, The American Chestnut Foundation (TACF) has developed a vigorous backcross breeding program designed to introduce the resistance of Chinese chestnut (*C. mollissima* Blume) into American chestnut (Hebard 1994). TACF's initial efforts have focused on American chestnut trees in southwest Virginia, but the goal is to restore the species throughout its entire native range. Hence, separate breeding programs have been started in a number of states. One question that is of primary interest to TACF and its State Chapters is how many breeding locations or separate programs will be needed across the entire range to capture most of the genetic variation still present in the species.

Previously, little was known about how genetic variability is distributed in American chestnut. In an exploratory examination of genetic variability, Huang et al. (1998) obtained results with allozyme and random amplified polymorphic DNA (RAPD) markers that suggest as many as four

regional populations might exist. However, statistical tests were not performed to quantify the magnitude of this component and whether it was significant. Because the question of whether regional genetic structure occurs among populations had not been settled, nor patterns of genetic variability completely described, we felt compelled to embark on a more thorough examination of genetic variation using state-of-the-art microsatellite DNA and RAPD markers.

A number of people (many of whom are dedicated TACF members; Sandra Anagnostakis; Dave Armstrong, Glen Beaver, Robert Bernatzky; Mary Bunch, Peter Carson, Hill Craddock, Mark Double, Fred Hebard, Craig Hibben, E. Kenneth James, Michael Kluempke, Jeff Lewis, Paul Sisco, Bob Summersgill, Wayne Swank, Melissa Thomas-VanGundy, Wells Thurber, Cathy Townsend, Stan Webb and Eric Weisse; if I forgot anyone please forgive me!) helped to collect leaf or dormant bud samples of American chestnut. In total, samples were collected at 22 sites across the natural range (refer to Figure 1). Most samples were collected from sites in State or National Forests, but a few collection sites were located on private land holdings. Each sample was assigned a unique ID and sent to the USDA Forest Service's Southern Institute of Forest Genetics in Saucier, Mississippi for DNA extraction and analysis.

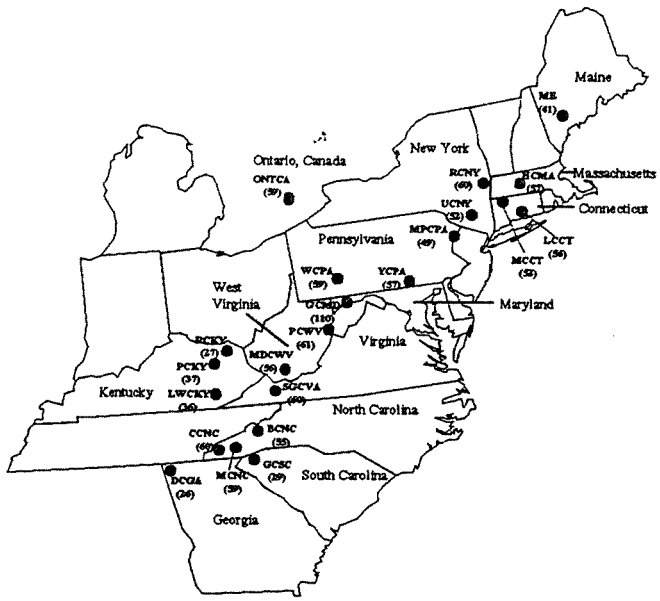

Figure 1. Map of the geographic origin of the 22 *Castanea dentata* Borkh. populations sampled in this investigation.

Prior to conducting the study, one of our main concerns regarding this investigation was the inclusion of trees that were not pure American chestnut. Inappropriate trees include hybrids or pure species other than American chestnut, especially the native relative known as chinkapin (*Castanea pumila* Mill.). Inclusion of such "contaminants" could have inflated our estimates of genetic diversity, especially in those populations containing the non-American chestnut samples, as well as potentially clouded any true patterns of genetic variability. Chloroplast DNA sequence variations have been widely used to investigate relationships among plant species (Palmer

Primers that amplified the spacer region between the *trn*T and *trn*L 5' exon of the chloroplast genome (Taberlet et al. 1991) could be used to uniquely identify American chestnut from all other *Castanea* spp.

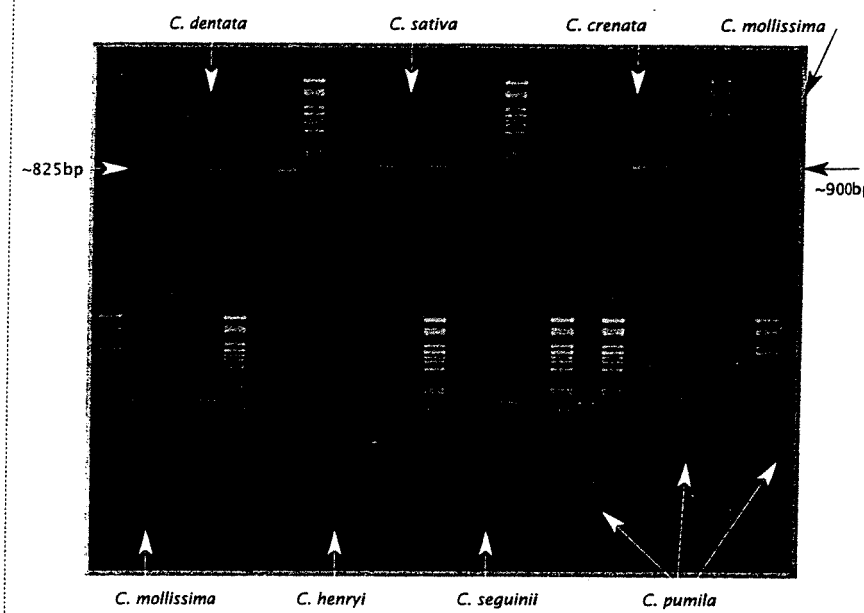

Figure 2. Chloroplast-specific marker amplified by primers a&b from Taberlet et al. (1991).

et al. 1988, Clegg et al. 1991) because they evolve slowly. To our good fortune, we identified a chloroplast-specific marker (Taberlet et al. 1991) that uniquely differentiates American chestnut from all other chestnut and chinkapin species (for example refer to Figures 2 and 3).

Unfortunately, chloroplasts are inherited only from the mother (maternally) hence this precluded our ability to distinguish hybrids of paternal origin. As a result, our sample set might still contain some hybrids, however, the number should have been small as most collections were made in either State Forests or National Forests where non-native chestnut and chinkapin species do not extensively occur. Of the 1158 trees sampled for this study, 165 trees (14.2%) from nine different sample sites were eliminated from further analysis as they were not pure American chestnut based on the size of the chloroplast marker (for example see Figure 3). In total as many as 993 trees were available for analysis of genetic variation.

The results of this study suggest that high levels of microsatellite and RAPD variability exist in American chestnut, and that most of this variation occurs within local populations (95.2% and 94.5%, respectively). These results are comparable to observations made in other long-lived, outcrossing, woody plant species with similar life history characteristics (Hamrick and Godt 1990; Hamrick et al. 1992), where as a rule, greater than 90% of the variation occurs within populations. Our results are also consistent with previous observations of allozyme variability in European chestnut (*C. sativa* Mill.) and American chestnut where 90% of the diversity was reported to exist within populations (Pigliucci et al. 1990; Huang et al. 1998). These results suggest that extensive gene flow, probably via long distance pollen movement, was possible prior to the blight. Hence, most of the genetic variation of the species is contained within any one population.

825bp=American chestnut
900bp = hybrid or species other than American chestnut

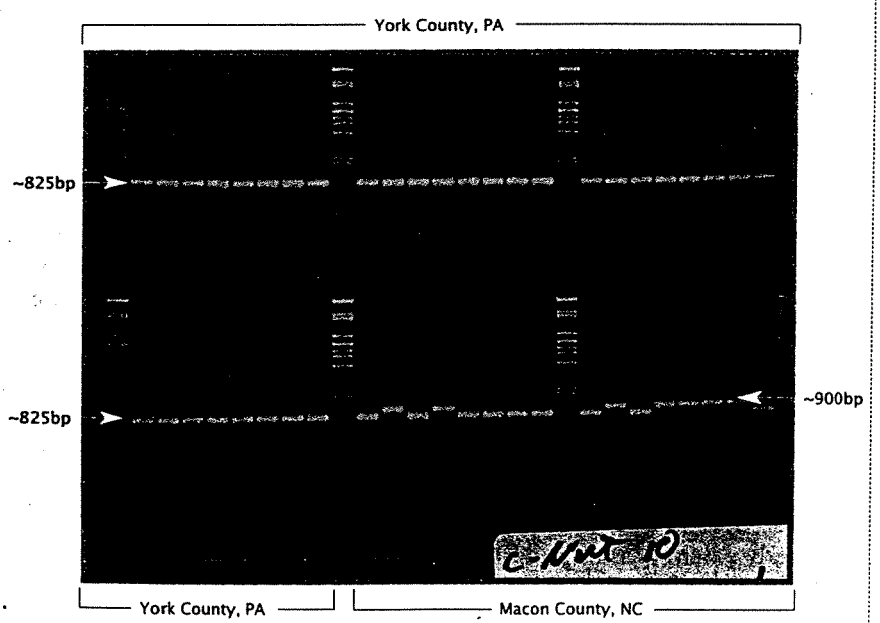

Figure 3. Example of usefulness of chloroplast-specific marker in identifying hybrids or pure species other than American chestnut

The results of this study also suggest that a cline in allele frequencies and number of rare alleles exists along the Appalachian axis. Clinal variation of allele frequencies along latitudinal and longitudinal gradients has been reported for a number of tree species (Lagercrantz and Ryman 1990; Zanetto and Kremer; Leonardi and Menozzi 1995, Tomaru et al. 1997), including European chestnut (Pigliucci et al. 1990; Villani et al. 1991; Villani et al. 1992; Villani et al. 1994). The main proposition set forth to explain this phenomenon is that geographical variation in allele frequencies resulted from post-glacial migration and founding events. Such processes are consistent with the patterns of variability we observed for American chestnut. The highest levels of gene diversity and the greatest numbers of rare alleles are found in the southwestern portion of its range. This suggests that its glacial refugium existed in the southern portions of its range, perhaps extending southward into the Gulf Coastal plain of present day Alabama and Mississippi.

Although most of the genetic variation found in American chestnut occurs within local populations, a statistically significant proportion exists among populations. Although our estimates of among population differentiation might be considered low (average 0.048), the values obtained indicate that populations significantly differ in allele frequency. Moreover, population pairwise estimates of genetic distance were shown to be significantly associated with the geographic distance between populations, suggesting that populations in close geographic proximity are slightly more genetically similar than geographically distant populations. These findings lead us to conclude that although long distance gene flow was possible in the past, it was infrequent enough to allow some genetic differentiation to take place.

Unlike the results of Huang et al. (1998), the results of this study suggest that little, if any, geographic structure exists in American chestnut. In other words, when statistical techniques such as cluster analysis or principal component analysis were performed, populations did not group or cluster together based on their geographic origin. Trees from the far northerly extent of the species range such as in Maine or Ontario were just as likely to group or cluster with trees that were sampled from North Carolina or Virginia as they were to cluster with trees from more proximal populations such as New York or Massachusetts. Prior to introduction of the blight, genetic variability in American chestnut followed a

pattern consistent with the hypothesis of a single large interbreeding 'meta'-population where genetic drift played a major evolutionary role.

Currently, roughly 95% of the neutral genetic variation of the species can be captured by sampling within any one population of American chestnut. However, we caution that the results of this study are based on neutral genetic loci and do not necessarily reflect genetic differentiation at adaptive genes or gene complexes. Such genes or gene complexes might include those that influence such traits as bud break, flowering time, cold hardiness, drought tolerance, nutrient uptake, leaf senescence, etc. Therefore, in order to assure that most of the variation produced by these types of genes or gene complexes are also captured in conservation and breeding endeavors, sampling should focus on collecting a fairly large number of individuals (50 to 100 or more) from each of several geographic areas. As proposed in Huang et al. (1998), we also suggest that a MINIMUM of at least three regions, representing northern, central, and southern portions of the species range, be considered in conservation and breeding efforts.

Acknowledgements

The authors thank Glen Johnson and Charles Burdine for their skilled technical assistance with DNA extraction, PCR amplification, microsatellite marker detection, and data acquisition; Andy David and Dave Wagner for the white oak microsatellite primer sequences; Gabriele Baccaro and Roberto Botta for the European chestnut microsatellite primer sequences; and Sandra Anagnostakis; Dave Armstrong, Glen Beaver, Robert Bernatzky; Mary Bunch, Peter Carson, Hill Craddock, Mark Double, Fred Hebard, Craig Hibben, E. Kenneth James, Michael Kluempke, Jeff Lewis, Paul Sisco, Bob Summersgill, Wayne Swank, Melissa Thomas-VanGundy, Wells Thurber, Cathy Townsend, Stan Webb and Eric Weisse for their assistance in collecting chestnut samples for this study. If we forgot anyone please forgive us, and again thank you!!

References

Clegg, M.T., Learn, G.H., and Goldberg, E.M. 1991. Molecular evolution of chloroplast DNA. In: Selander, R.K., Clark, A.G. and Whittam, T.S. (Eds.) Evolution at the molecular level. Sinauer Associates Inc.: Sunderland, Mass.

pp. 135-149.
Hamrick, J.L., and Godt, M.J. 1990. Allozyme diversity in plant species. In: Brown, H.D., Clegg, M.T., Kahler, A.L., and Weir, B. (Eds.) Plant population genetics, breeding, and genetic resources. Sinauer Associates Inc.: Sunderland, Mass. p. 43-63.
Hamrick, J.L., Godt, M.J., and Sherman-Broyles, S.L. 1992. Factors influencing levels of genetic diversity in woody plant species. In: Adams, W.T., Strauss, S.H., Copes, D.L., and Griffin, A.R. (Eds.) Population genetics of forest trees. Kluwer Academic Publishers: London. p. 95-124.
Hebard, F.V. 1994. The American Chestnut Foundation breeding plan: beginning and intermediate steps. J. Amer. Chestnut Found. 8(1):21-28.
Huang, H., Dane, F., and Kubisiak, T.L. 1998. Allozyme and RAPD analysis of the genetic diversity and geographic variation in wild populations of the American chestnut (Fagaceae). Amer. J. Bot. 85:1013-1021.
Lagercrantz, U., and Ryman, N. 1990. Genetic structure of Norway spruce (*Picea abies*): concordance of morphological and allozyme variation. Evolution 44:38-53.
Leonardi, S., and Menozzi, P. 1995. Genetic variability of *Fagus sylvatica* L. in Italy: the role of postglacial recolonization. Heredity 75:35-44.
Palmer, J.D., Jansen, R.K. Michaels, H.J., Chase, M.W. and Manbart, J.R. 1988. Chloroplast DNA variation and plant phylogeny. Ann. Missouri Bot. Garden 75:1180-1206.
Pigliucci, M., Benedettelli, S., and Villani., F. 1990. Spatial patterns of genetic variability in Italian chestnut (*Castanea sativa*). Can. J. Bot. 68:1962-1967.
Sargent, C.S. 1905. Manual of trees of North America, pp. 220-222.
Stephenson, S.L., Adams, H.S., and Lipford M.L. 1991. The present distribution of chestnut in the upland forest communities of Virginia. Bull. Torrey Bot. Club 118:24-32.
Taberlet, P., Gielly, L., Patou, G. and Bouvet, J. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol. Biol. 17:1105-1109.
Tomaru, N., Mitsutsuji, T., Takahashi, M., Tsumura, Y., Uchida, K., and Ohba, K. 1997. Genetic diversity in *Fagus crenata* (Japanese beech): influence of the distributional shift during the late-Quaternary. Heredity 78:241-251.
U.S. Census Bureau. 1908. The lumber cut of the United States, 1907. U.S. Census Bureau. Forest Products 2:1-53.
Villani, F., Pigliucci, M., Benedettelli, S., and Cherubini, M. 1991. Genetic differentiation among Turkish chestnut (*Castanea sativa* Mill.) populations. Heredity 66:131-136.
Villani, F., Pigliucci, M., and Cherubini, M. 1994. Evolution of *Castanea sativa* Mill. in Turkey and Europe. Genetic Research Cambridge 63:109-166.
Villani, F., Pigliucci, M., Lauteri, M., and and Cherubini, M. 1992. Congruence between genetic, morphometric, and physiological data on differentiation of Turkish chestnut (*Castanea sativa*). Genome 35:251-256.
Zanetto, A., and Kremer, A. 1995. Geographical structure of gene diversity in Quercus petraea (Matt.) Liebl. I. Monolocus patterns of variation. Heredity 75:506-517.