Clonal Fidelity in Large Colonies of *Gaylussacia brachycera* Gray (Box Huckleberry) Assessed by DNA Fingerprinting

Margaret Pooler[1,*], Rob Nicholson[2], and Andrew Vandegrift[1]

Abstract - *Gaylussacia brachycera* (box huckleberry) is a slow-growing, dwarf evergreen member of the family *Ericaceae* that is native to eight states in the eastern United States. It is a rare plant with conservation status in several states of critically imperiled (S1). Botanists have been intrigued by this enigmatic native plant since it was discovered in 1796 in Virginia. One of the mysteries of this species is whether plants in a colony arose from different genotypes or are clonal. The species reproduces primarily by means of underground runners and appears to be self-sterile, so sexual reproduction within isolated colonies could be limited. Using molecular markers, we tested samples taken from three of the best-known colonies in Pennsylvania and one in Tennessee. Based on 104 polymorphic markers, we found that one of the Pennsylvania colonies contained two genotypes among 11 samples tested; one Pennsylvania colony contained three genotypes among five samples tested; and the other two colonies exhibited no variation among the 8–10 samples tested. This study represents the first time that molecular markers have been used in a systematic assay to determine the existence of variation among individuals within a colony of box huckleberry.

Introduction

Gaylussacia brachycera (Michx.) Gray (box huckleberry) is a slow-growing, evergreen groundcover in the family *Ericaceae* that is native to Delaware, Pennsylvania, Maryland, Virginia, West Virginia, Kentucky, and Tennessee (USDA NRCS 2006), as well as North Carolina (Wilber and Bloodworth 2004). It is usually found growing in dry or well-drained acidic soils or duff in partial shade. Discovered in 1796 by Andre Michaux near Winchester, VA, this plant was originally named *Vaccinium brachycerum* (Michaux 1803). It was described and renamed *Gaylussacia brachycera* by Asa Gray (Gray 1846), based on collections made in 1845 by S.F. Baird in Pennsylvania. Awareness of the plant grew with the subsequent discovery of a large colony in Pennsylvania by H.A. Ward in 1919 (Ward 1920), as well as smaller colonies discovered in other states (Smith and Smith 1971).

Despite its distribution in eight states in the middle eastern US, box huckleberry is not a common plant. The global conservation status for box huckleberry is listed as G3 (rare), with S1 (critically imperiled) listing for the states of Delaware, Maryland, and Pennsylvania and S2

[1]USDA/ARS, US National Arboretum, Floral and Nursery Plants Research Unit, 3501 New York Avenue, NE, Washington, DC 20002. [2]Smith College Botanic Garden, 15 College Lane, Northampton, MA 01063. [*]Corresponding author - Margaret.Pooler@ars.usda.gov.

(imperiled) for Virginia and West Virginia (Center for Plant Conservation 2006). It is estimated that currently this species grows at 100 sites within its native range (Crable 1999).

Box huckleberry reproduces predominantly vegetatively by means of underground runners, with an estimated average annual spread of six inches per year (Coville 1919). Although seed set is possible, it is rare in natural populations due to self-incompatability in isolated clones where outcrossing is not possible (Coville 1919, Dirr 1998, Wherry 1934). This obligatory clonal reproduction, combined with the size of the colonies found in Pennsylvania, has led to rampant speculation on the age of the clones. Estimates ranging from 5000 years to 13,000 years have been made, which would make the former 2-kilometer-long clone in Pennsylvania the oldest documented living thing on earth (Krussmann 1977, Moldenke 1957, Wherry 1972, Willaman 1965).

The purpose of this study was to investigate the clonality of several colonies of box huckleberry in Pennsylvania that are reputed to be single-genotype stands that are thousands of years old. We used AFLP markers to fingerprint plants in order to ascertain the level of diversity among samples within each colony and to determine the level of diversity among separate colonies.

Materials and Methods

Plant materials

Leaf samples were collected from three stands in Pennsylvania that showed contiguous growth, as well as one stand in Tennessee. Approximate colony size and locations of samples within each colony are indicated in Figure 1. Leaf samples were sent to the US National Arboretum for DNA

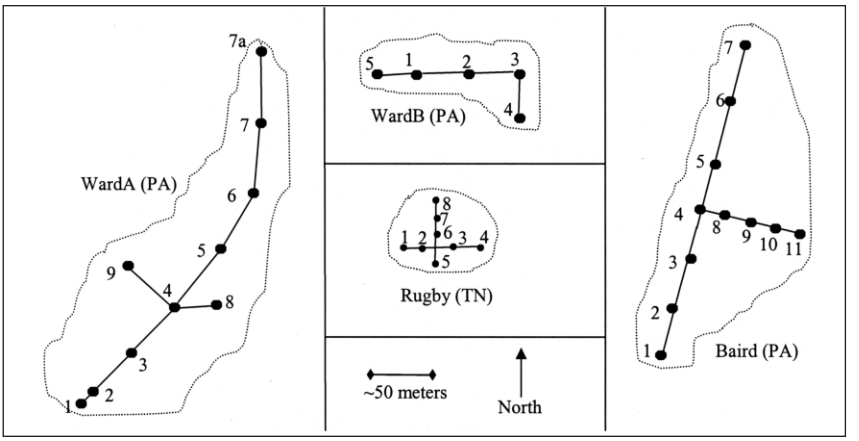

Figure 1. Sampling patterns for stands of *G. brachycera* (box huckleberry) from Perry County, PA (WardA, WardB, and Baird; sampled August 2004) and Morgan County, TN (Rugby, sampled July 2000). The WardA and WardB populations are approximately 300 m apart and 16 km east of the Baird population.

extraction and analysis. Leaves for DNA analysis were freeze-dried and stored at -80 ºC in a freezer until analysis.

DNA extraction

DNA was extracted from three freeze-dried leaves of each sample using the methods described for other species in our laboratory (Pooler et al. 2002), with the following modification: leaves were first ground to a powder in the lysing matrix (Bio101, Vista, CA) by dry vortexing for 2–3 minutes, and then 500 µl of prewarmed (65 ºC) grinding buffer (Wilson et al. 1992) were added. The mixture was processed in a FastPrep machine (Bio101), centrifuged for 1 minute at top speed (≈10,000 rpm) in a microcentrifuge, and the grinding buffer was then pipetted off. Five hundred µl of CTAB buffer (Doyle and Doyle 1987) and 1.5 µl of beta-mercaptoethanol were added to the pelleted ground plant material, and our standard DNA extraction procedure (Pooler et al. 2002) using the QIAamp tissue kit (Qiagen, Inc., Valencia, CA) was followed. Processing the leaves in the grinding buffer prior to CTAB extraction was necessary for DNA recovery from box huckleberry.

AFLP reactions

AFLP analysis was performed as described by Vos et al. (1995) and Invitrogen Corporation (2003), with slight modifications, noted below, to prepare samples for analysis on an ABI 310 Genetic Analyzer (Applied Biosystems, Foster City, CA). DNA restriction, digestion, and ligation were carried out sequentially using approximately 0.25 µg of genomic DNA. Restriction digestion was performed at 37 ºC for three hours, and ligations at 20 ºC for three hours. Preselective reactions took place in 20-µl volumes containing PCR buffer (Invitrogen Corp., Carlsbad, CA), plus 3 mM $MgCl_2$, 100 µM dNTP, 0.125 µM each preselective primer, 2.0 U of Taq DNA polymerase (Invitrogen Corp.), and 3 µl diluted restriction/ligation reaction. Completed preselective reactions were diluted 1:50 with TE, and 5 µl were used as template for all selective reactions. Selective amplification reactions were carried out in 20-µl volumes containing the same reagents as for preselective amplification, except that 0.25 µM MseI primer and 0.1 µM EcoRI primers (Table 1) were used instead of preselective primers and only 0.5 U Taq DNA polymerase was used. The EcoRI selective primers had fluorescently labeled 5' ends and were purchased from the Applied Biosystems Custom Oligonucleotide Synthesis Service (Foster City, CA). Completed selective reactions were analyzed on an ABI310 automated DNA sequencer with POP4 polymer. Samples were prepared by mixing 1.0 µl of selective reaction, 0.07 µl Genescan 500 ROX size standard (Applied Biosystems), and 10.93 µl deionized formamide. All reactions were replicated from DNA extraction through amplification to ensure repeatability.

Data analysis

Markers were visualized as peaks using Genotyper® 2.5 software (Applied Biosystems) and scored manually as present or absent for each accession.

Similarity coefficients between each accession were calculated using the SIMQUAL program in NTSYS-pc, version 1.70 (Rohlf 1992), using the Dice similarity coefficient (2a / [2a + b + c], where a = total number of bands shared by both individuals, b = bands unique to one individual, and c = bands unique to the other individual). These data were subjected to cluster analysis using the UPGMA method in the SAHN program of NTSYS to generate a phenogram. Cophenetic matrices were constructed and compared with the similarity matrices using the MXCOMP program to test the goodness of fit of a cluster (Rohlf 1992). Bootstrap analysis using 5000 replications was performed on the original raw data using WinBoot (Yap and Nelson 1996) to determine confidence limits of clusters in the UPGMA-based dendrograms (Felsenstein 1985). The bootstrap value indicates the percentage of times the group to the right of the node occurred in the bootstrap analysis.

Results and Discussion

Analysis of 13 AFLP primer pairs revealed 104 polymorphic bands out of a total of 1682 bands (average 8 polymorphic bands per primer pair), with a range of three to 17 polymorphic bands per primer (Table 1). Reproducibility between replicate samples was good, with approximately 8% of markers not used due to inconsistencies between samples or one sample that had missing data. Monomorphic bands (typically more than 100 per primer pair) were not scored, since they are not informative. This level of polymorphism is consistent with a previous study in our lab where we examined genetic distances among 24 geographically diverse box huckleberry accessions using a slightly different AFLP protocol (Pooler et al. 2006). While the present study complements our previous population genetic distance study, the results of these two studies cannot be compared directly, nor can the data be merged because different primers and a different AFLP protocol were used.

Table 1. List of AFLP selective primer extensions, total number of peaks detected, and number of polymorphic markers per primer pair among samples from four populations of *G. brachycera* (box huckleberry) accessions.

EcoRI selective primer extension with dye name	MseI selective primer extension	Total number of scorable markers	Number of polymorphic markers
AAC (NED)	CAC	118	15
AAC (NED)	CTC	104	6
AAC (NED)	CTG	139	17
ACC (FAM)	CAC	110	6
ACC (FAM)	CTC	179	10
ACC (FAM)	CTG	168	3
ACG (NED)	CAC	127	3
ACG (NED)	CAG	99	11
ACG (NED)	CTC	78	7
ACG (NED)	CTG	123	5
AGC (HEX)	CAC	138	8
AGC (HEX)	CTC	155	6
AGC (HEX)	CTG	144	7

Pairwise comparisons of accessions using the 104 polymorphic markers were computed using the DICE coefficient. In our previous work (Pooler et al. 2006), as well as that of others (Koopman et al. 2001; Mace et al. 1999a, b), we found that using UPGMA clustering, the DICE coefficient resulted in identical phenograms as the Jaccard coefficient, with similar correlation coefficients. The DICE similarity values were then used to construct a phenogram using UPGMA clustering. A phenetic approach using UPGMA clustering was used in this study because our primary objective was to determine the genetic relationships among accessions within a single colony, rather than to determine phylogenetic histories. In addition, it has been demonstrated that branches or clusters in a cladogram or phenogram that are well supported by bootstrap statistics will be similar regardless of the method used (Kardolus et al. 1998, Koopman et al. 2001). The UPGMA-based phenogram using the DICE similarity data had a cophenetic correlation coefficient (r) of 0.9985 (Fig. 2), indicating a very

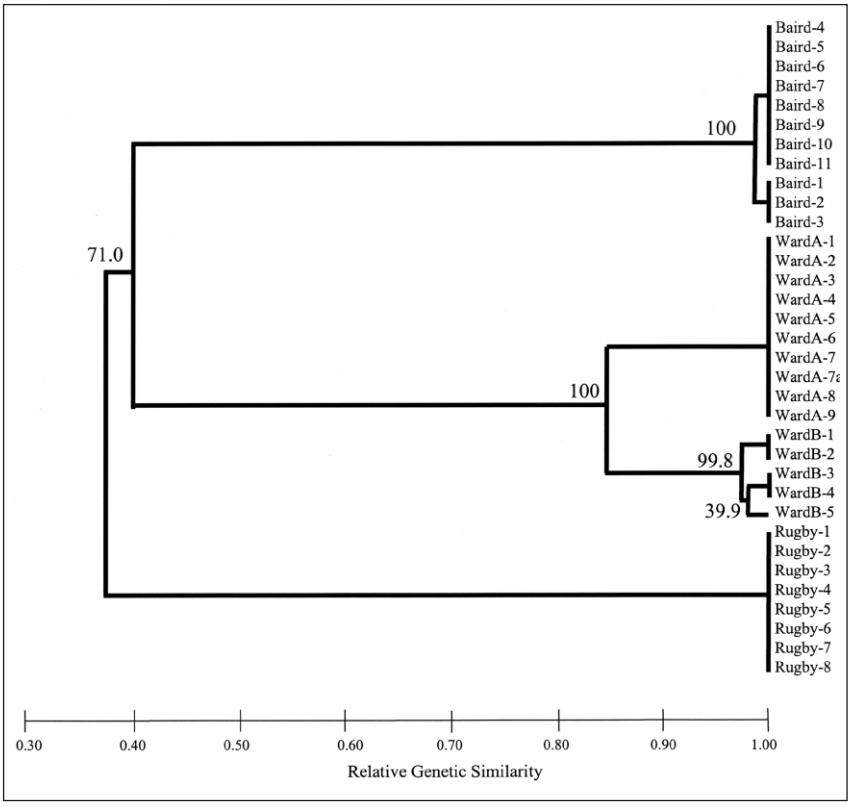

Figure 2. UPGMA-derived phenogram of genetic similarity based on the Dice similarity coefficient among *G. brachycera* (box huckleberry) accessions based on 104 polymorphic characters from 13 AFLP primer pairs. Cophenetic correlation coefficient (r) = 0.9985. Bootstrap confidence values are indicated to the left of each node and represent the percentage of times that cluster appeared in the consensus tree.

good fit of the phenogram with the original distance matrix (Rohlf 1992). In addition, the bootstrap values for most of the clusters were high, indicating that those clusters have statistical and therefore biological significance, and are not simply an artifact of the clustering technique (Yap and Nelson 1996). The low bootstrap value (39.9) in the WardB cluster indicates that these groups are not well defined, probably due to the low number of polymorphic markers between them (two out of 104). The clustering of the samples, as well as the relative genetic similarities between clusters, was directly related to physical proximity of the samples (Figs. 1 and 2). It is important to note that the genetic similarity values were based on only the polymorphic markers; therefore, like in other studies, the values reflect the genetic similarities of these individuals and colonies relative to each other, and not an absolute genetic similarity.

The AFLP fingerprints indicated that the Baird colony of box huckleberry has two genotypes represented among the 11 samples collected, with one genotype (Baird1, Baird2, and Baird3) originating from one corner of the collection site (Fig. 1). These genotypes differ by only one polymorphism out of 104 scored, however, which suggests that the different genotypes may have arisen by somatic mutation rather than sexual recombination. This colony, located near New Bloomfield, PA, was the first discovered and is probably the best-known population of box huckleberry in Pennsylvania. The existence of at least two genotypes in this population is also supported by results from our previous study, which also showed molecular variation within an albeit limited sample of two clones (Pooler et al. 2006).

The WardA population represents the second-known stand in Pennsylvania, discovered by Ward in 1919 and approximately 16 kilometers east of the Baird colony. WardA is what remains of the 2-kilometer-long patch described in 1920 (Ward 1920). Much of this stand has been lost due to a forest fire in 1963 and highway construction. The longest sampled transect of this population was 307 meters, and all 10 samples from this site had identical genotypes. The WardB stand was approximately 300 m from the WardA stand, on a separate ridge with a stream in between. The five samples from the WardB stand exhibited three different genotypes, defined by two AFLP markers unique to the WardB-5 sample and two markers unique to the WardB-1 and -2 samples. All three genotypes appeared along the longest transect of 110 meters (Fig. 1).

It is not clear why there was more variability seen in the samples from the smaller WardB colony than the larger WardA stand. Perhaps the WardB colony was originally established by seed representing several genotypes, and this diversity has been maintained through relatively noncompetitive clonal reproduction. Prevailing winds or other environmental factors affecting insect behavior could have contributed to outcrossing and subsequent seed production in this population.

The box huckleberry population that was sampled in Rugby, TN was used as a test population to obtain preliminary data on clonal fidelity in

other populations of box huckleberry. The eight samples collected from two transects had identical genotypes. The relative genetic similarity between this population and the cluster containing the Pennsylvania populations was only 0.38, which is not surprising based on their physical and geographic distance.

This study clearly illustrates the existence of genetic diversity within and between large stands of box huckleberry in Pennsylvania. It also lends support to, but cannot prove unequivocally, the hypothesis that within the former 2-kilometer-long WardA stand, all plants are clonal. Based on the large size of the WardA colony and the presence of at least three genotypes in the nearby WardB colony, more sampling from the WardA colony is necessary to provide strong evidence of the clonal nature of this colony. However, the assertions that the box huckleberry clones in Pennsylvania are the oldest living things on earth cannot be disproven. Based on the current size of the remaining colonies and a presumed growth rate of 15 cm per year, it can certainly be stated that these clones are some of the oldest plants in the eastern US.

Acknowledgments

The authors wish to thank Dr. Melvin Shemluck, Quinsigamond College, Worcester, MA for assistance in collecting samples in Pennsylvania.

Literature Cited

Center for Plant Conservation. 2006. CPC National Collection Plant Profile. Available online at http://www.centerforplantconservation.org/ASP/CPC_ViewProfile.asp?CPCNum=1999. Accessed January 5, 2006.

Coville, F.V. 1919. The threatened extinction of the box huckleberry. Science L (1280):30–34.

Crable, A. 1999. Meet the world's oldest—and hardest working—plant. Lancaster New Era. Aug. 20, 1999.

Dirr, M.A. 1998. Manual of Woody Landscape Plants, 5th Edition. Stipes Publishing LLC, Champaign, IL. 1187 pp.

Doyle, J.J., and J.L. Doyle. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochemistry Bulletin 19:11–15.

Felsenstein, J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. Evolution 39:783–791.

Gray, A. 1846. *Chloris boreali-americana*. Memorial American Acadamy of Arts and Science, New Series 3:54.

Invitrogen Life Technologies. 2003. AFLP Analysis System I/AFLP Starter Primer Kit Instruction Manual, Version B. Invitrogen Corporation, Carlsbad, CA.

Kardolus, J.P., H.J. Van Eck, and R.G. Van den Berg. 1998. The potential of AFLPs in biosystematics: A first application in *Solanum* taxonomy (Solanaceae). Plant Systematics and Evolution 210:87–103.

Koopman, W.J.M., M.J. Zevenbergen, and R.G. Van den Berg. 2001. Species relationships in *Lactuca* S.L. (Lactuceae, Asteraceae) inferred from AFLP fingerprints. American Journal of Botany 88:1881–1887.

Krussmann, G. 1977. Manual of Cultivated Broad-leaved Trees and Shrubs. Vol II, E-Pro. Timber Press, Portland, OR. 445 pp.

Mace, E.S., R.N. Lester, and C.G. Gebhardt. 1999a. AFLP analysis of genetic relationships among the cultivated eggplant, *Solanum melongena* L., and wild relatives (Solanaceae). Theoretical and Applied Genetics 99:626–633.

Mace, E.S., C.G. Gebhardt, and R.N. Lester. 1999b. AFLP analysis of genetic relationships in the tribe Datureae (Solanaceae). Theoretical and Applied Genetics 99:634–641.

Michaux, A. 1803. Flora Boreali-Americana: Sistens Caracteres Plantarum Quas in America Septenrionali Collegit et Detexit Andreas Michaux. 2 vols. Parisiis; Argentorati [Strasbourg]: Apud Fratres Levrault.

Moldenke, H.N. 1950. The oldest flowering plant. Horticulture 18:456–457.

Pooler, M.R., L.G.H. Riedel, S.E. Bentz, and A.M. Townsend. 2002. Molecular markers used to verify interspecific hybridization between hemlock (*Tsuga*) species. Journal of the American Society for Horticultural Science 127:623–627.

Pooler, M.R., L.R. Dix, and R.J. Griesbach. 2006. Genetic diversity of the endangered box huckleberry (*Gaylussacia brachycera*) based on AFLP markers. Journal of the Torrey Botanical Society 133(3):439–448.

Rohlf, F.J. 1992. NTSYS-pc. Numerical taxonomy and multivariate analysis system, version 1.70 Exeter Publishing, Ltd., Setauket, NY.

Smith, H., and D. Smith. 1971. The box huckleberry, *Gaylussacia brachycera*. Castanea 36:81–89.

US Department of Agriculture, Natural Resource Conservation Service (USDA, NRCS). 2006. The PLANTS Database, Version 3.5. Available online at http://plants.usda.gov. Data compiled from various sources by Mark W. Skinner. National Plant Data Center, Baton Rouge, LA.

Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. van de Lee, M. Hornes, A. Frijters, J. Pot, Peleman, M. Kuiper, and M. Zabeau. 1995. AFLP: A new technique for DNA fingerprinting. Nucleic Acids Research 23:4407–4414.

Ward, H.A. 1920. A new station for *Gaylussacia brachycera*. Rhodora 22:167–168.

Wherry, E.T. 1934. The box huckleberry as an illustration of the need for field work. Bulletin of the Torrey Club 61:81–84.

Wherry, E.T. 1972. Box huckleberry as the oldest living protoplasm. Castanea 37: 94–95.

Wilber, R.L., and S. Bloodworth. 2004. Notes on the box huckleberry, *Gaylussacia brachycera* (Ericaceae) and its unexpected presence in North Carolina. Rhodora 106:371–377.

Wilson, H.D., J. Doebley, and M. Duvall. 1992. Chloroplast DNA diversity among wild and cultivated members of *Cucurbita* (Cucurbitaceae). Theoretical and Applied Genetics 84:859–865.

Willaman, J.J. 1965. The oldest living plants. Morris Arboretum Bulletin 16:65–66.

Yap, I.V., and R.J. Nelson. 1996. WinBoot: A program for performing bootstrap analysis of binary data to determine the confidence limits of UPGMA-based dendrograms. International Rice Research Institute, Manila, Philippines.