Thirty Polymorphic Nuclear Microsatellite Loci From Black Walnut

K. Woeste, R. Burns, O. Rhodes, and C. Michler

Black walnut (*Juglans nigra* L) is a large tree, native to the eastern United States, that is prized for its high-quality timber and edible nut. Thirty (GA/CT)n nuclear microsatellite markers were identified from black walnut for use in population genetic studies, genome mapping, DNA genotyping of important clones, studies of gene flow, and tree breeding. The markers were polymorphic based on a diversity panel of 10 black walnut individuals from eight Midwestern U.S. states.

Black walnut (*Juglans nigra* L.) is a large tree that is native throughout the eastern United States from New England to Texas (Fowells 1965). Black walnut is prized as a multipurpose species: it provides valuable timber, produces a high-quality edible nut, and is attractive to wildlife (USDA Forest Service Fire Effects Information System website). More than 15 million acres of timberland in 30 states contain black walnut (Schmidt and Kingsley 1997). The vast majority of this species exists in natural stands, with a small percentage grown in plantations. An estimated 15 million cubic feet of black walnut are harvested annually in the United States (USDA Forest Service Forest Inventory and Analysis website). In 1997, 29 million pounds of in-shell black walnut nuts were purchased for processing and about 2 million pounds of nutmeat were sold (Hammonds 1998). Nearly all processed nuts came from uncultivated trees growing in wild populations (Reid 1990).

There have been successful genetic improvement efforts in black walnut for both timber and nut traits (Beineke 1989; Funk 1979; Tourjee 1998). There is now a need for DNA-based genetic markers to investigate critical problems in black walnut breeding and conservation. For example, breeders need a method for genotyping important cultivars to verify their identity (Bish C, personal communication). Efforts to understand the wild black walnut germplasm have been largely limited to provenance tests (Bresnan et al. 1992; Rink 1997). Provenance tests, and the associated morphological and phenological data, have provided important tools for breeders and foresters where they are available, well designed, and well maintained (Guries et al. 1981), but provenance tests of black walnut are expensive and time consuming because the species has a long juvenility and mature trees are large. Genetic markers such as restriction fragment length polymorphisms (RFLPs; Fjellstrom and Parfitt 1994; Fjellstrom et al. 1994), random amplified polymorphic DNA (RAPDs; Woeste et al. 1996), internal transcribed spacer (ITS) sequence polymorphisms (Potter D, personal communication), and allozymes (Arulsekar et al. 1985) have been identified for several members of the Juglandaceae, and they permit a rapid parsing of the genetic differentiation within *J. nigra*. While these markers are more informative than phenotypes in terms of their ability to identify species substructure and diversity, microsatellite DNA markers [simple sequence length polymorphisms (SSLPs)] can provide greater levels of resolution in a cost-effective manner. Microsatellites overcome some of the limitations of other marker systems (Goldstein and Pollock 1997), and a large number of methods for statistical analysis of microsatellite data are available (Luikart and England 1999).

We intend to use the microsatellites published here as part of a larger effort to understand the genetics of black walnut. While the commercial value of this species both for nuts and timber is based almost entirely on exploitation of the wild resource, there are large gaps in our understanding of the genetic structure of wild populations of black walnut, and there is no published information on the effects of timber and nut harvests on the long-term health of the species. In addition, the markers will be used for genetic mapping, DNA genotyping (fingerprinting) of important clones, and studies of gene flow in seed orchards as part of a breeding effort.

Materials and Methods

DNA was isolated from the leaves of three black walnut selections using Nucleon Phytopure DNA extraction columns (Amersham, Buckinghamshire, UK). The trees were part of a walnut breeding and genetics program in the Department of Forestry and Natural Resources at Purdue University. A pooled DNA sample from these trees was used by Genetic Identification Services (San Diego, CA) to create an enriched (GA/CT)n microsatellite library. The library was plated on a selective medium, and 1500 colonies were robotically picked (Genetix, Hampshire, UK) into 96-well plates, miniprepped (Qiagen REAL 96 Prep, Valencia, CA), and sequenced using an ABI 3700 (Perkin-Elmer, Foster City, CA). We analyzed the resulting sequences using Sequencher software (version 3.1.1; Gene Codes, Ann Arbor, MI) and discarded candidate sequences if they contained no discernible microsatellite repeat, or if there was insufficient flanking sequence to construct suitable polymerase chain reaction (PCR) primers. The sequences that remained were assigned to contigs whenever possible. When sequence contigs were available, we derived a consensus sequence for the regions flanking the microsatellites and used it for primer design. Primers (18–20 bp) for amplification of microsatellite-containing sequences (100–400 bp) were designed using Primer 0.5 (Whitehead Institute for Biomedical Research, Cambridge, MA).

To develop a preliminary screening panel, DNA from 10 *J. nigra* individuals representing populations in eight Midwestern U.S. states was isolated from mature leaves using an automated nucleic acid extractor (Autogen, Framingham, MA) and a CTAB extraction buffer modified with 2× PVP and 2× CTAB. PCR amplification of primer pairs was performed with an MJ Research thermal cycler (Waltham, MA) using 20 μl reactions. The PCR reaction mixture contained 20 ng of DNA template, 1.5 mM $MgCl_2$, 0.4 U AmpliTaq Gold (Perkin-Elmer), and 0.8 μM (each) primer. All other components of the PCR mixture were as recommended by the manufacturer (Perkin-Elmer). PCR amplification was for 50 cycles of 92°C for 30 s, 45°C for 1 min, and 72°C for 1 min. All primers were annealed at 45°C. The reaction products were then held at 0°C until aliquots could be loaded into 1.5% Trevigels (Trevigen, Gaithersburg, MD) containing ethidium bromide. Electrophoresis was in 1× TAE buffer, and gels were photographed using

a Stratagene Eagle Eye II digital imaging system (La Jolla, CA). Primers that generated a clear PCR product band of the predicted size were characterized as either clearly polymorphic or probably monomorphic based on the genotypes in the screening population. To confirm that microsatellites were polymorphic, PCR was performed with ABI fluorescent dCTP according to manufacturer's instructions (Perkin Elmer), and 1 μl of the PCR product and 2 μl of CXR 350 bp Ladder Standard (Promega, Fitchburg Center, WI) were combined in a separate tube, denatured for 2 min at 95°C, and loaded onto a Quick-Comb 96-well comb (Sigma, St. Louis, MO). Electrophoresis was in 6% Long Ranger (polyacrylamide) denaturing gels (BMA, Rockland, ME) at 3000 V, 60 mA, 200 W, 51°C for 2.5 h using an ABI 377 (Perkin Elmer) with 36 cm plates and 0.2 mm spacers.

Table 1. Characterization of microsatellites in black walnut

Locus name	Repeat array	Primer sequences (5'–3')	Length (bp)	No. of alleles
WGA2	$(GA)_{30}$	F: GACGACGAAGGTGTACGGAT R: GTACGGCTCTCCTTGCAGTC	169	12 (8)[a]
WGA4	$(GT)_5,(GA)_{15},(GA)_{11}$	F: TGTTGCATTGACCCACTTGT R: TAAGCCAACATGGTATGCCA	241	5 (8)
WGA6	$(GA)_{25}$	F: CCATGAAACTTCATGCGTTG R: CATCCCAAGCGAAGGTTG	157	6 (8)
WGA7	$(GA)_{19}$	F: CAAACAAAATCCGACCGC R: AAACCTCGATGAGCGAAGAA	222	5 (7)
WGA11	$(CT)_{14}$	F: CTCGACAGAAACAGCCACAA R: GGAGTTGTGTGCAGTGCG	220	9 (9)
WGA17	$(GA)_7,(GA)_8$	F: CTGTCACTTGATCACCGTGG R: GAGCTTCTACCAACGCCAAG	202	2 (7)
WGA24	$(T)_8,(CT)_{18},(CT)_4$	F: TCCCCCTGAAATCTTCTCCT R: TTCTCGTGGTGCTTGTTGAG	242	6 (9)
WGA25	$(CT)_{16}$	F: CCACTTTCGCTCTCGATTTC R: TTGCGCAGCCTCTATAAGGT	179	3 (5)
WGA27	$(GA)_{30}$	F: AACCCTACAACGCCTTGATG R: TGCTCAGGCTCCACTTCC	242	8 (9)
WGA32	$(CT)_{19}$	F: CTCGGTAAGCCACACCAATT R: ACGGGCAGTGTATGCATGTA	176	9 (9)
WGA33	$(GA)_{22},(GAGT)_5,(GA)_5$	F: TGGTCTGCGAAGACACTGTC R: GGTTCGTCGTTTGTTGACCT	230	11 (9)
WGA42	$(GA)_{14}$	F: GTGGGTTCGACCGTGAAC R: AACTTTGCACCACATCCACA	241	6 (9)
WGA45	$(T)_4(CT)_{19}$	F: TCGTTACCACCAGCACAGAG R: GACATAGCGAGGGGCTAGG	233	8 (9)
WGA47	$(CT)_{19}$	F: CCCCTTCAAATTAGGGCTTC R: AGCTCCAAACATTGGAAGGA	150	6 (9)
WGA53	$(CT)_8$	F: CTCCCTTCCGGTAAGACCTC R: GAGGCGCGACTAGGTTAGG	176	10 (9)
WGA54	$(GA)_{17}$	F: CTAGGCTTCCCTAGCCGTG R: GGCTCCTCTCGATCTCGAC	218	12 (9)
WGA56	$(CT)_{17}$	F: GTGAAGCTGCAATGAATTGG R: AAAGATGCAACAGACACCCC	152	11 (9)
WGA58	$(CT)_8CCT(CT)_{13}$	F: CCCTAGCCTGGCTTTTCTTT R: ATAGCCTCCACTGGTTGTGG	201	8 (9)
WGA60	$(GA)_{22}$	F: CAAGCTTCCGTATTGGTGGT R: TATACCGCAAGCTCGCAAC	234	11 (9)
WGA65	$(CT)_{20}$	F: CACCGTCTTATGCCATCCTT R: GTGCACTGTGGACGAAGAGA	161	4 (9)
WGA69	$(GA)_4ATATAA(GA)_{16}$	F: TTAGTTAGCAAACCCACCCG R: AGATGCACAGACCAACCCTC	182	5 (9)
WGA70	$(GA)_{14}$	F: TGTAATTGGGGAATGTTGCA R: TGGGAGACACAATGATCGAA	165	8 (9)
WGA71	$(GA)_6,(G)_{12}$	F: ACCCGAGAGATTTCTGGGAT R: GGACCCAGCTCCTCTTCTCT	212	3 (9)
WGA72	$(CT)_{14}$	F: AAACCACCTAAAACCCTGCA R: ACCCATCCATGATCTTCCAA	151	9 (9)
WGA73	$(CT)_9AT(CT)_9$	F: AGATCAAGCCTCCACCACC R: TGCGGCTGAGACATATTGAG	230	6 (9)
WGA74	$(GA)_{18},(GA)_{11}$	F: AAGCTCATGGTGTGGAGGAG R: TTTTCCCTCACTCTGTGCG	184	6 (9)
WGA78	$(GA)_{11}$	F: CATGCAAGCTTGTCGTGG R: AACCCCCACGAATTCTCTCT	212	5 (8)
WGA79	$(GA)_{12}$	F: CACTGTGGCACTGCTCATCT R: TTCGAGCTCTGGACCACC	206	9 (9)
WGA80	$(CT)_{20}$	F: GCAACGAATGTGAGAAAGCA R: TGTGGAATTGTGAGCGGATA	216	7 (9)
WGA82	$(GA)_{23}$	F: TGCCGACACTCCTCACTTC R: CGTGATGTACGACGGCTG	175	11 (9)

[a] Number in parentheses is the number of individuals analyzed.

Results and Discussion

Our initial screening of 1500 colonies from an enriched (GA/CT)n library yielded 450 unique microsatellite-containing sequences from *J. nigra*. The remaining colonies produced sequences containing no discernible microsatellite repeats or insufficient flanking DNA to construct PCR primers. Of the 450 positive sequences, 141 (30%) were grouped into contigs of two or more sequences (presumably derived from the same locus). Because our microsatellite-containing sequences were derived from the pooled DNA of three *J. nigra* individuals, 67% (95/141) of these contigs contained sequences that differed in the number of (GA/CT)n repeats. This indicates that the three arbitrarily chosen *J. nigra* individuals that we used to construct our microsatellite library were polymorphic at about two-thirds of the microsatellite loci we identified.

Thirty of the microsatellites showed clear polymorphism in the screening population (Table 1). Alleles ranged from 150 to 242 bp, a range that should facilitate multiplexing of samples. About 66% of the microsatellite sequences contained perfect (GA/CT)n repeats. The average number of (GA/CT)n repeats was 18.2, and the range was between 8 and 30 repeats. The remaining 34% of the microsatellites contained repeats that were interrupted, and five microsatellites contained repeats other than (GA/CT)n, including WGA33, which contained the tetranucleotide repeat $(GAGT)_5$.

An average of 8.6 individuals was analyzed at each locus, and we observed an average of more than seven alleles per locus. This high level of polymorphism may have been the result of how the microsatellite library was developed and screened. Sequence contigs polymorphic among the three individuals that were used to make the library were selected for analysis first. We will not know if the allelic richness found in these 30 microsatellites is characteristic of all walnut microsatellite loci in the species until more of the library has been screened with larger populations. By comparison, Fjellstrom and Parfitt (1994) found an average of 1.47 alleles per RFLP locus in five *J. nigra* populations containing 11 individuals each.

The black walnut microsatellite loci we describe here may be useful for population-level studies, genetic mapping, plant breeding, and cultivar identification. The ordering of these microsatellites into a ge-

netic map of black walnut will enhance their value as markers for breeding and diversity studies. Sampling from hierarchically structured wild populations will permit an estimation of population genetic parameters and provide insight into the reproductive biology of black walnut. Confirming sequence homology for the SSLPs that amplify in several members of the genus will increase their suitability for use in comparative genome or mapping studies, studies of hybrid zones, and population genetic studies.

From the USDA Forest Service, North Central Research Station, Hardwood Tree Improvement and Regeneration Center, Department of Forestry and Natural Resources, 1159 Forestry Building, Purdue University, West Lafayette, IN 47907. We gratefully acknowledge the technical assistance of Dr. Phillip San Miguel, director of the Purdue University Genomics Core Facility. The use of trade names is for the information and convenience of the reader and does not imply official endorsement or approval by the U.S. Department of Agriculture or the Forest Service of any product to the exclusion of others that may be suitable. Address correspondence to Keith Woeste at the address above or e-mail: kwoeste@fs.fed.us.

© 2002 The American Genetic Association

References

Arulsekar S, Parfitt DE, and McGranahan GH, 1985. Isozyme gene markers in *Juglans* species. J Hered 76:103–106.

Beineke WF, 1989. Twenty years of black walnut genetic improvement at Purdue University. North J Appl For 6: 68–71.

Bresnan F, Geyer W, Lynch KD, and Rink G, 1992. Black walnut provenance performance in Kansas. North J Appl For 9:41–43.

Fjellstrom RG and Parfitt DE, 1994. Walnut (*Juglans* spp.) genetic diversity determined by restriction fragment length polymorphisms. Genome 37:690–700.

Fjellstrom RG, Parfitt DE, and McGranahan, GH, 1994. Genetic relationships and characterization of Persian walnut (*Juglans regia* L.) cultivars using restriction fragment length polymorphisms (RFLPs). J A Soc Hort Sci 119:833–839.

Fowells HA, 1965. Silvics of forest trees of the United States. USDA Agriculture Handbook 71. Washington, DC: USDA.

Funk DT, 1979. Black walnuts for nuts and timber. In: Nut tree culture in North America (Jaynes RA, ed). Hamden, CT: Northern Nut Growers Association; 51–73.

Goldstein DB and Pollock DD, 1997. Launching microsatellites: a review of mutation processes and methods of phylogenetic inference. J Hered 88:335–342.

Guries RP, Brown S, and Kress J, 1981. A guide to forest tree collections of known source or parentage in northeastern and north central U.S. and adjacent Canadian provinces. Research Bulletin R-3142. Madison, WI: University of Wisconsin-Madison, Agricultural Experiment Station.

Hammonds B, 1998. Status report on the eastern black walnut nut industry, nut markets, by-products and future challenges. In: Nut production handbook for eastern black walnut (Jones JE, Mueller R, and Van Sambeek JW, eds). Republic, MO: Southwest Missouri Resource Conservation & Development; 25–28.

Luikart G and England PR, 1999. Statistical analysis of microsatellite DNA data. Tree 14:253–255.

Reid W, 1990. Eastern black walnut: potential for commercial nut producing cultivars. In: Advances in new crops (Janick J and Simon JE, eds). Portland: Timber Press; 327–331.

Rink G, 1997. Genetic variation and selection potential for black walnut timber and nut production. In: Knowledge for the future of black walnut: proceedings of the Fifth Black Walnut Symposium, Springfield, MO, July 28–31, 1996. General Technical Report NC-191 (Van Sambeek J, ed). St. Paul, MN: USDA Forest Service, North Central Forest Experiment Station; 58–62.

Schmidt TL and Kingsley, NP, 1997. Status of black walnut in the United States. In: Knowledge for the future of black walnut: proceedings of the Fifth Black Walnut Symposium, Springfield, MO, July 28–31, 1996. General Technical Report NC-191 (Van Sambeek J, ed). St. Paul, MN: USDA Forest Service, North Central Forest Experiment Station; 9–22.

Tourjee KR, 1998. Missouri eastern black walnut breeding program. In: Nut production handbook for eastern black walnut (Jones JE, Mueller R, and Van Sambeek JW, eds). Republic, MO: Southwest Missouri Resource Conservation & Development; 90–96.

USDA Forest Service. Forest Inventory and Analysis Timber Product Output (TPO) Database Retrieval System (visited/last modified Feb 26, 2001). ‹http://srsfia.usfs.msstate.edu/rpa/tpo/›

USDA Forest Service, Rocky Mountain Research Station, Fire Sciences Laboratory. Fire Effects Information System (visited Feb 26, 2001). ‹http://www.fs.fed.us/database/feis/›

Woeste K, McGranahan GH, and Bernatzky R, 1996. Randomly amplified polymorphic DNA loci from a walnut backcross [(*Juglans hindsii* × *J. regia*) × *J. regia*]. J Am Soc Hort Sci 121:358–361.

Received April 23, 2001
Accepted September 20, 2001

Corresponding Editor: James L. Hamrick