Microsatellite discovery and sequence comparisons in an endangered North American hardwood tree
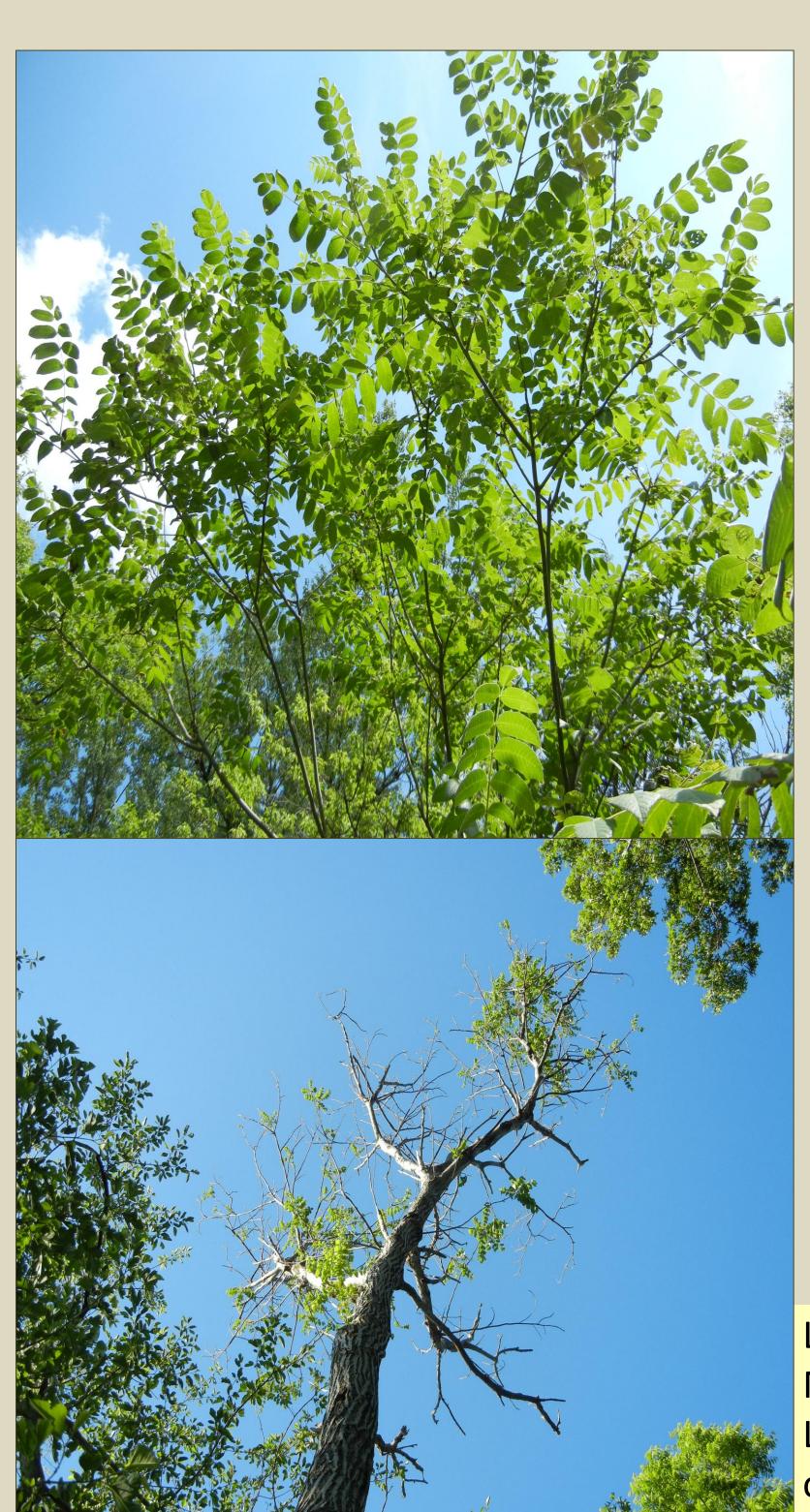
Nick LaBonte and Keith E. Woeste

Purdue University Department of Forestry and Natural Resources, USDA Forest Service Hardwood Tree Improvement and Regeneration Center

Introduction

Butternut (*Juglans cinerea*) is a rare hardwood tree native to most of the eastern United States. It has been decimated in recent decades by an introduced fungal pathogen, *Ophiognomonia clavigignenti-juglandacearum,* which causes a disease known as butternut canker. Research on butternut at Purdue University and the HTIRC is focused on understanding its ecology and the possibility that heritable canker resistance exists in remaining wild populations of butternut. Microsatellite markers have been developed for butternut and some of its close relatives in the past decade. These have proven useful in analyses of rangewide and regional population genetics as well as interspecific hybrid identification. In order to further butternut breeding research, we used next-generation sequencing to identify new microsatellite markers in butternut and Japanese walnut (*Juglans ailantifolia*). We also utilized the expanding database of non-model species sequence information to localize isolated microsatellite sequences to the nuclear genome or the chloroplast and mitochondria of butternut.

BLAST analysis

We used the BLASTer online program developed by Diagrid to compare our microsatellite sequences with the 'NCBI-genomic other' and 'NCBI-EST other' databases. The Diagrid development team is based at Purdue's Rosen Center for Advanced Computing. Japanese walnut and butternut were compared separately. We were interested in localizing the microsatellite-containing sequences to the nucleus or the organelle genomes. Minimum E-value was set to 10 in the search, but during analysis hits under 50 base pairs were discarded because brief segments of repetitive sequence in microsatellites have an unacceptably high rate of false assignment. Significant hits were considered those over 50 aligned base pairs for queries of the other-genomic and other-EST databases. In addition, the Purdue Genomics Center provided us with BLAST results from the UniProtK database, which was queried against both contigs and reads not assembled into contigs. Results are summarized in tables 2, 3, and 4.

Table 2. BLAST summary. Hit contigs indicates the number of contigs with significant hits.

	JC	JA
Total contigs	2653	1298
Hit contigs: EST-other	575	54
Hit contigs: Genomic-other	35	11
Hit contigs: UniProtK	151	64
Mitochondrion	3	2
Chloroplast	0	0

Table 3 (Below). Species breakdown for top BLAST hits, Japanese walnut contigs.
Table 4 (Right). Species breakdown for top BLAST hits, butternut contigs.

Genomic		EST		UniProtK	
Species	hits	Species	hits	Species	hits
Theobroma cacao	3	*Juglans hindsii x regia*	26	*Vitis vinifera*	41
Fragaria vesca	3	*Fagus sylvatica*	5	*Ricinus communis*	6
Vitis vinifera	1	*Juglans regia*	4	*Ailuropoda melanoleuca*	2
Citrullus lanatus	1	*Alnus glutinosa*	3	*Cucumis melo*	2
Ricinus communis	1	*Quercus petraea*	2	*Glycine max*	2
Medicago truncata	1	*Quercus robur*	2	*Betula pendula*	1
Arabidopsis lyrata	1	*Mimulus guttatus*	2	*Malus domestica*	1
Other	0	Other	9	Other	9

Genomic		EST		UniProtK	
Species	hits	Species	hits	Species	hits
Ricinus communis	9	*Juglans hindsii x regia*	164	*Vitis vinifera*	80
Theobroma cacao	7	*Carya illinoensis*	49	*Ricinus communis*	15
Fragaria vesca	5	*Juglans regia*	44	*Populus trichocarpa*	15
Carica papaya	5	*Quercus robur*	36	*Arabidopsis thaliana*	4
Cucumis sativa	2	*Fagus sylvatica*	27	*Callithrix jacchus*	4
Medicago trunculata	2	*Alnus glutinosa*	16	*Arabidopsis lyrata*	4
Citrullus lanatus	1	*Quercus petraea*	16	*Medicago truncata*	4
Other	1	*Malus x domestica*	12	Other	24
		Vitis vinifera	12		
		Citrus spp	10		
		Other	189		

Left (Top): A healthy butternut on the Kettle Moraine State Forest in southern Wisconsin.
Left (Bottom): A butternut (Wisconsin) with severe canker symptoms. Photos by N. LaBonte.

Conclusions and future research

Microsatellite markers will remain a valuable tool for tree breeders, particularly in minor species like butternut. The most polymorphic loci we discovered will be used in our current project, an estimation of the heritability of disease resistance in a large wild population of butternut. We also plan to use sequence comparison software to query this new sequence data against a previously developed microsatellite library for black walnut (*Juglans nigra*). Some microsatellite primers developed for black walnut also work well for butternut, and we hope to gain insight into the extent of sequence similarity and divergence between these sympatric relatives.

Microsatellite sequence enrichment

We isolated DNA from frozen butternut (JC) and Japanese walnut (JA) leaves using standard phenol-chloroform extraction protocols developed at the HTIRC. DNA was viewed on agarose gels and analyzed with a Nanodrop spectrophotometer (Thermo Scientific) for concentration so that 5 micrograms of DNA could be used for the enrichment. Enrichment for microsatellite sequences was performed using the protocol of Glenn and Schable (2005), and DynaBeads (Invitrogen) were used to bind microsatellite-containing sequences. Di-, tri-, and tetra-nucleotide oligo probes were used. Enrichment products were sequenced using a Roche 454 sequencer at the Purdue Genomics Core facility. We obtained 40,783 reads in Japanese walnut, and 67,928 in butternut. Contigs (2653 butternut and 1298 Japanese walnut) were viewed in Tablet (Milne et al. 2010) and primers were designed. Of 40 primer pairs screened in the summer of 2012, nearly all amplified, and nine were polymorphic. Three had greater than 5 alleles, a reasonable number for butternut (Table 1).

Table 1. Microsatellite loci derived from butternut (JC)

Contig no.	Length (bp)	Repeat	Alleles (JA and JC)
13	390	TCC, TAC	5
92	167	TG	8
149	238	TG	2
167	217	CT	4
194	181	CTT	4
224	235	TTC	3
242	246	TC	8
965	164	CT	3
1453	270	Complex	2

References
Glenn, T.C. and N.A. Schable. 2005. Isolating microsatellite DNA loci. Methods in Enzymology 395:202-222.
Milne, I, Bayer, M, Cardle, L, Shaw, P, Stephen, G, Wright, F and Marshall, D. 2010. Tablet - next generation sequence assembly visualization. Bioinformatics 26(3), 401-402.
Christopher Thompson; Brian Raub (2012), "BLASTer," http://diagrid.org/resources/blastgui