



Genetic characterization of *Liriodendron* seed orchards with EST-SSR markers

Xinfu Zhang¹, Alanna Carlson¹, Zhenkun Tian², Margaret Staton³, Scott E. Schlarbaum⁴, John E. Carlson⁵ and Haiying Liang^{1*}

*Correspondence: hliang@clemsun.edu



CrossMark

← Click for updates

¹Department of Genetics and Biochemistry, Clemson University, Clemson SC 29634, USA.

²Beijing Forestry University, Beijing, China.

³Department of Entomology and Plant Pathology, The University of Tennessee, Knoxville, TN 37996-4563, USA.

⁴Department of Forestry, Wildlife & Fisheries, The University of Tennessee, Knoxville, TN 37996-4563, USA.

⁵Department of Ecosystem Science and Management and The Department of Plant Science, Pennsylvania State University, University Park, PA 16802, USA.

Abstract

Liriodendron tulipifera L., is a wide-spread, fast-growing pioneering tree species native to eastern North America. Commonly known as yellow-poplar, tulip tree, or tulip-poplar, the species is valued, both ecologically and economically. It is perhaps the most commonly used utility hardwood in the USA, and is planted widely for reforestation and, in varietal forms, as an ornamental. Although most seedlings used for reforestation today derive from collections in natural populations, two known seed orchards, established from plus-tree selections, i.e. superior phenotypes, in the 1960's and 1970's have been used for local and regional planting needs in Tennessee and South Carolina. However, very little is known about the population genetics of yellow-poplar nor the genetic composition of the existing seed orchards. In this study, 194 grafted yellow-poplar trees from a Clemson, SC orchard and a Knoxville, TN orchard were genetically characterized with 15 simple sequence repeat (SSR) markers developed from expressed sequence tags (ESTs). Of the 15 EST-SSR markers, 14 had a polymorphic information content (PIC) of at least 0.5. There was no significant difference between the Clemson and Knoxville orchards in average effective number of alleles (5.93 vs 3.95), observed and expected heterozygosity (H_o : 0.64 vs 0.58; H_e : 0.74 vs 0.70), Nei's expected heterozygosity (0.74 vs 0.58), or Shannon's Information index (1.84 vs 1.51). The larger Clemson orchard exhibited a significantly greater number of observed alleles than the Knoxville orchard (15.3 vs 7.4). Overall, substantial genetic diversity is captured in the Clemson and Knoxville orchards.

Keywords: Genetic diversity, seed orchard, SSR markers, species

Introduction

Liriodendron tulipifera L., commonly known as yellow-poplar or tulip-poplar, is a wide-spread, fast-growing pioneering hardwood species of considerable economic value in the forests of eastern North America. Yellow-poplar is distributed predominantly east of the Mississippi River from the gulf coast to southern Canada (28° to 43° north latitude) [35]. According to the forest inventory analysis [11], as surveyed from 2006-2012, the total saw log volume of *L. tulipifera* on timberland in the United States was 25.9 billion cubic feet, with the majority (65%) located in the southeastern United States. The species is shade intolerant and highly competitive, growing faster than

Acer rubrum L. (red maple) and *Quercus rubra* L. (northern red oak) seedlings under a variety of silvicultural understory treatments (Beckage and Clark 2003). Yellow-poplar is often seen as a pioneering species in old fields. As a component of 16 forest cover types, this species' degree of dominance has created differentiation between the ecological communities [46]. In addition, yellow-poplar is valued as a nectar source for honey production, as a source of wildlife food (mast), and as a large shade tree in urban plantings [3]. The wood of yellow-poplar is used in a diverse range of products, such as in furniture, pallets and framing construction as well as pulp [12,41]. Chemical extracts from yellow-poplar wood or leaves have proven useful,

such as sesquiterpenes which have an anti-tumor effect and antifeeding for herbivores [27], and antimicrobial alkaloids [2].

L. tulipifera has been cultivated since 1663 [5] and is currently widely planted in eastern forests. Although seed orchards have been established to meet local or regional planting needs in the U.S.A. [6,36], genetic diversity of *Liriodendron* seed orchards in relation to natural stands has not been studied. Because seed orchards is the bridge between breeding and silvicultural activities, genetic diversity of tree seeds orchards determines the genetic quality of future forest stands and forms the basis for further improving the management of genetic resources and for the genetic modification of cultivars to meet new environmental challenges. Thus, the lacking information limits utilization of these *Liriodendron* orchards in a tree improvement program.

The primary goal of our study was to determine the genetic composition and diversity in two *Liriodendron* seed orchards in the southeastern USA. The orchard residing in Knoxville, Tennessee, was established in 1966 and contains 100 grafted ramets, representing 31 genotypes or clones. The Clemson orchard in South Carolina was established in 1976 by grafting multiple ramets of 150 plus trees selected from throughout the 17,500-acre Clemson Experimental Forest by Dr. Roland E. Schoenike (<http://www.clemson.edu/trails/history/schoenike.html#top>). Seeds from this orchard have been used for reforestation efforts for a number of years. Currently there are 165 surviving trees in the Clemson orchard. Besides *L. tulipifera*, the only other *Liriodendron* species is *Liriodendron chinense*, which is native to China and Vietnam.

Although the two species separated 10~16 million years ago [32], they are quite similar morphologically and are cross fertile [26,34], and the hybrids exhibit heterosis [31,39]. Because the incomplete records suggest that the Clemson orchard may contain *L. chinense* or hybrids, we first used the sequence of a chloroplast gene, maturase *K* (*matK*), to discriminate the two *Liriodendron* species and their hybrids. Then we investigated the genetic diversity and allele richness among selections of this unique native species in each orchard as a first step toward contrasting orchard-produced seedling diversity with natural diversity. We chose simple sequence repeat (SSR) markers (also called microsatellites) in the study, because SSR markers are co-dominant, easily reproduced and scored, highly polymorphic, abundant through the genome, and have higher information content than isoenzyme and dominant markers [45].

Materials and methods

Plant materials and DNA isolation

Fresh leaves of all *Liriodendron* trees (165) from the Clemson seed orchard and 31 trees from the Knoxville seed orchard were collected in the spring of 2013 and stored in plastic bags at -80°C prior to DNA isolation. All these trees represented different clones as validated by the SSR markers used in this study. Leaves from a *Liriodendron tulipifera* tree (accession

number 70921 H) from the US National Arboretum (collected by Kevin Conrad) were also included in the study. Total genomic DNA was isolated from leaves using a CTAB protocol as described in [16] and suspended in TE buffer (Tris base 6.1g/L, EDTA 0.37 g/L, pH 8). The quality and concentrations of genomic DNA from individual plants were determined with a NanoDrop 3300 (Thermo Scientific, Wilmington, Delaware, USA) and by electrophoresis on 0.8% agarose gels.

Distinguishing between two *Liriodendron* species based on maturase *K* sequence

The record of the 165 surviving *Liriodendron* trees in the Clemson orchard is not complete. Therefore, the sequence of a chloroplast gene, maturase *K* (*matK*) was used to discriminate between the species/hybrids. The *matK* sequence was amplified with forward (5'-CGATCTATTCATTCAATATTTTC-3') and reverse primers (5'-TCTAGCACACGAAAGTCGAAGT-3') in a 12.5- μ l reaction containing 6.875 μ l ddH₂O, 1 μ l MgCl₂ (25 mM), 0.5 μ l forward primer (10 μ M), 0.5 μ l reverse primer (10 μ M), 0.25 μ l dNTPs (10 mM each), 0.25 μ l BSA (0.8 μ g/ μ l), 0.125 μ l Taq Polylmerase (5u/ μ l), 0.5 μ l DNA (~20ng/ μ l), 2.50 μ l 5X PCR buffer (-Mg).

The conditions for polymerase chain reactions (PCR) were as follows: 5 minutes of initial denaturation at 94°C, 35 cycles of touch-down PCR with 30 seconds of denaturation at 94°C, 30 seconds of annealing at 60-50°C (first cycle 60°, then each subsequent cycle 1°C lower than the previous until 51°C annealing temperature, followed by 25 cycles each with a 50°C annealing temperature), and 3 minutes of extension at 72°C, and a final extension at 72°C for 10 minutes. Before being sequenced with 1 μ l of 10 μ M forward or reverse primer, PCR products were cleaned with ExoAP mix (89 μ l H₂O+ 10 μ l 5000U/mL Antarctic Phosphatase +1 μ l 20000U/mL Exonuclease I) for 30 minutes in a reaction containing 1 μ l of PCR product and 1 μ l of ExoAP mix, followed by a heat inactivation step at 80°C for 15min. An 834 bp-segment of maturase *K* gene from each tree was used for alignment with MUSCLE and curated with Gblocks, and a phylogenetic tree was built with maximum likelihood (PhyML) (<http://www.phylogeny.fr/>) [7].

The maturase *K* gene sequence of *L. tulipifera* (GI: 5731451), *L. chinense* (GI: 7239759), and a hybrid (GI: 389955358) available in GenBank were included in the analysis.

L. tulipifera EST-SSR markers, PCR amplification, and allele sizing

Twenty simple sequence repeat (SSR) markers (also called microsatellites) were used to investigate the genetic composition of the *Liriodendron* seed orchards. These markers included seven Expressed Sequenced Tags (EST)-SSR markers (LT002, LT015, LT021, LT086, LT096, LT131, LT157) previously characterized by electrophoresis on 8% polyacrylamide gels [42] and thirteen new markers (LTCU19, LTCU40, LTCU51, LTCU53, LTCU125, LTCU139, LTCU142, LTCU143,

LTCU145, LTCU150, LTCU151, LTCU152, LTCU154) mined from a comprehensive EST dataset [22]. PCR amplification for each marker was performed with genomic DNA of *Liriodendron* trees from the Clemson and Knoxville seed orchards and the US National Arboretum. For a more cost-effective 153 primer screening, a M13 tail (5'-CACGACGTTGTAACGAC-3') was added to the 5'-end of the forward primer of each marker pair in order to amplify the fragments using a complementary adapter with a fluorescent dye (6-FAM, VIC, NED, or PET) at its 5'-end (Applied Biosystems, Foster City, California, USA). Polymerase chain reactions were carried out in a 12.5- μ L solution comprising: approximate 75 ng DNA template, 0.052 U/ μ L Promega Taq DNA polymerase, 0.16 nM forward primer, 0.4 nM reverse primer, 0.4 nM fluorescent M13 primer, 0.24 mM each dNTPs, and 1.2 \times Promega PCR buffer. The PCR profile consisted of an initial denaturation at 94°C for 3 minutes followed by 10 cycles of 1 minute at 94°C, 1 minute at annealing temperature (Table 1), and 1 minute 15 seconds at 72°C, and then 35 cycles of 1 minute at 94°C, 1 minute at 58°C, and 1 minute at 72°C, with a final extension of amplified DNA at 72°C for 5 minutes.

An aliquot of 1.5 μ L PCR products were treated with 1.5 μ L of 10-fold diluted ExoSAP-IT (Affymetrix Inc. Cleveland, OH, USA) to remove single stranded primers which might influence fragment analysis at 37°C for 30 minutes and then at 80°C for 15 minutes. After being diluted to 100 ng/ μ L, 1 μ L of each sample was mixed with 0.1 μ L of LIZ600 and 8.9 μ L of Hi-Di Formamide, denatured at 95°C for 5 minutes, and then put on ice for 10 minutes before being separated on an ABI 3730 Genetic Analyzer. The Dye set was DS-33 (6-FAM, VIC, NED, PET and LIZ). Allele sizes were scored with GeneMapper (4.0) (Applied Biosystems, Foster City, California, USA). Functional annotation of EST-SSRs was performed by applying a homology search of reassembled ESTs against the non-redundant (nr) NCBI database using the BLASTx algorithm [1].

Data analysis

MICRO-CHECKER [38] was employed to check for potential genotyping errors arising from large allele drop-out and stuttering. Observed and expected heterozygosities and polymorphic information content (PIC) were calculated using Cervus 2.0 [25]. Deviations from Hardy–Weinberg equilibrium and the Shannon's Information index were calculated with GENEPOP (<http://genepop.curtin.edu.au/>, Raymond and Rousset 1995).

Results and discussion

Distinguishing between two *Liriodendron* species with maturase *K* (*matK*) sequence

It is not clear when *L. chinense* was first introduced to U.S.A., but the two *Liriodendron* species hybridize readily [26] and efforts in crossing have been well-documented [e.g., 31,34]. The two species are similar morphologically, except that *L. chinense* is smaller in stature and has larger, more deeply

lobed leaves and smaller flowers. However, our attempt to tell these two species apart by morphology failed: the leaf shape varied depending on age (Supplementary figure S1) and the flowers were located at a too high for sampling. Molecular techniques including biochemical analysis [34], isozymes [14], and fingerprinting with random amplified polymorphic DNA (RAPD) [21] have been explored in discrimination of *Liriodendron* species and their hybrids. In 2012, Zhang et al., reported an SSR marker that amplified a 190-bp fragment from *L. chinense*, a 180-bp fragment from *L. tulipifera*, and both 190- and 180-bp fragments from hybrid. In this study, *matK* sequence was employed. The *matK* gene locates within the intron of the *trnK* and codes for maturase like protein involved in Group II intron splicing [37]. The *trnKUUU-matK* region, ranging from approximately 2.2 kb (liverworts) to 2.6 kb (seed plants) in size, is universally present in land plants and only few exceptions of a secondary loss or reorganizations are known to date [40]. Because the *matK* gene evolves more rapidly, compared to other plastid genes, it has become a valuable marker for lower-level phylogenetic reconstruction of systematic and evolutionary studies. The Clemson seed orchard contains 165 surviving trees. The *matK* sequence was amplified from each of the 165 trees in the Clemson orchard (Supplementary figure S2). When the amplicons were pair-end sequenced, an 834-bp segment of high quality was obtained for each tree, representing 55% of the full-length gene. There were only eight nucleotides different between the two *Liriodendron* species within the 834-bp segment. As shown in Figure 1 and Supplementary figure S3, only Tree#CU24 and #134 were not *L. tulipifera*. Their hybrid status was established by having seven nucleotides different from *L. tulipifera* and 2 nucleotides different from *L. Chinense*. These results confirm the record of hybrids being planted in the Clemson Orchard. Thus, these two hybrids were excluded in the genetic composition analysis. Further, our study indicates that *L. tulipifera*, *L. Chinense*, and their hybrids contain unique nucleotide compositions in *matK* sequence that can be utilized in distinguishing the species and hybrids.

Amplification of EST-SSR loci in *Liriodendron*

No evidence for large allele dropout was found for any of the 20 markers. Stuttering occurred in five markers: LT131, LT157, LTCU40, LTCU142, and LTCU143, and these five markers were excluded in further analyses. All of the remaining 15 markers were polymorphic in both Clemson and Knoxville orchards. The 20 markers were also tested on one *L. tulipifera* and one *L. chinense* tree from the US National Arboretum. Eleven loci were heterozygous and one locus failed in the *L. tulipifera* tree. PCR amplification for all 20 markers was successful in the *L. Chinense* tree, although sizing in an ABI 3730 Genetic Analyzer failed for LT157 and LTCU142, due to stuttering. Fourteen loci were heterozygous in the *L. chinense* tree. This indicates a high frequency of transferability of *L. tulipifera* EST-SSR markers in *L. Chinense*, supporting the previous

Table 1. Characteristics of 20 EST-SSR loci.

Marker name	Repeat motif	Forward primer sequence (5'-3')	Reverse primer sequence (5'-3')	Expected size (bp)	Stuttering ^a	Annealing temperature °C
LT002	(GCA)8	CCTACCACCAGCA ATACCTA	TCTCGTCGCTGAAGAT ATG	189	N	59
LT015	(CCGAAC)5	TCCGTTATCTCTCT CCAAAA	CTAGACAGGTGCTCGG ATAC	110	N	59
LT021	(TTC)8	CAAATACCATTGC ACCTTGT	ACGCATCCTCTTCCAC TAC	180	N	57
LT086	(CTT)10	AAGACAGGACTTT CCACTGA	GAACGAACCTAACCA AATGA	274	N	55
LT096	(CT)20	TGCAACCTAACAA GATGTGT	TGAAAAGCAACCAAG TTACC	272	N	55
LT131	(AC)22	GCAGCATCTCCTC ATATTCT	TTGCAGTTGAGCTATT GTTG	240	Y	55
LT157	(TTC)6	AGTTGCCCTTTAGC TTCTTT	GCCACAGAGTTTGGGA AGTA	222	Y	55
LTCU19	(AG)10	GTGGATTGCAAAG GCAGAGT	AAAACAAAAGCAAGC AAGCC	183	N	57
LTCU40	(ATG)8	TTGCGTAAATGCA TCCAAAA	GAAGCCtaTGCAAGAT GCAA	181	Y	55
LTCU51	(CT)18	ATCACCATCTTCCT CATCGC	AAACCATTCCAACCAT CCAA	198	N	55
LTCU53	(TG)14	CGGATCTTTCTCTT TCCATCC	AAGAAGATTGCAGAG GCAGAA	223	N	55
LTCU125	(TC)8	CGAAAGACATTCC CATCACA	CCATTACAATCCACAG CCAA	205	N	55
LTCU139	(TCT)10	GAATAACCGCTCT TTTGGGA	AAGCCAAGTGCCAAA GAAGA	164	Y	55
LTCU142	(AAT)8	TGGTGCATATGGG CTTAGAA	TATTCCTCCAGCTTCT CCTT	171	Y	55
LTCU143	(TG)13	AAAAATGCTAATC CAATAACTTTTCG	TATCCAACCGATCACC CATT	160	N	55
LTCU145	(GA)18	TTGAAGTCCAGAT TGATTGATTG	GCCTAGGGaGATGtTTT TGG	157	N	55
LTCU150	(TC)10	TCTTCAAACCAAG GCTGTTG	GCACTACATCCCTTTTc CCA	167	N	55
LTCU151	(TC)11	TGAGGTGACTTTG GCTTTTG	GACCCgaGCTGTAAAA TGGA	189	N	55
LTCU152	(CA)17	CATCCAAATGCAG CAGAAAT	ATTCCCACTCGGTTGA ACAC	177	N	55
LTCU154	(CT)10	GATGAAGGAGAAT TCTATATTTCTGA	CCAGCCAAGAAAGAA AATGG	156	N	55

^aY:Yes; N:No.

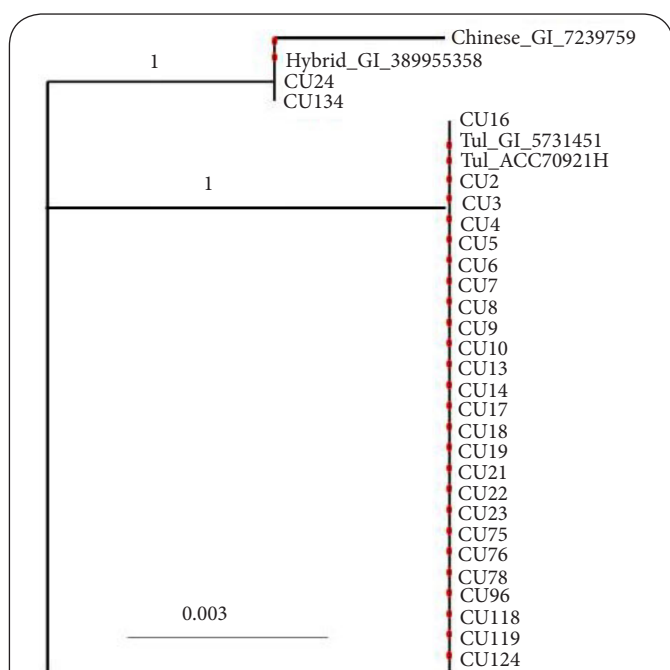


Figure 1. Sequence comparison of maturase K gene. The sequences were aligned with MUSCLE the phylogenetic tree was built with maximum likelihood (PhyML). The maturase K gene sequence of *L. tulipifera* (GI: 5731451), *L. chinense* (GI: 7239759), and a hybrid (GI: 389955358) available in GenBank were included in the analysis. The complete illustration of the tree is included in the [Supplementary figure S1](#).

findings of 72.4% success rate by [42] and 82.1% by [43]. This is expected because EST-SSRs have generally demonstrated a high frequency of cross-species transferability despite less polymorphism compared to genomic SSRs [9,12,44]. Among the 194 *L. tulipifera* trees included in the study, the number of alleles per locus ranged from 3 to 26 (mean=13.0) (Table 2). The observed and expected heterozygosities (*Ho* and *He*) ranged from 0.17 to 0.89 and from 0.19 to 0.93, with averages of 0.62 and 0.74, respectively. The polymorphic information content (PIC) ranged from 0.17 to 0.92, with an average of 0.71. Overall, 14 of the 15 markers had a PIC \geq 0.5.

Many genomic resources, such as expressed sequence tag (EST) databases [15,22,23] and genomic DNA libraries [24], have been developed for *L. tulipifera*. Through these resources, several thousand putative SSR markers have been identified by *in silico* mining. However, only 345 *L. tulipifera* SSR markers having been tested for polymorphism by polyacrylamide denaturing gels [42,43]. Compared to other species, *Liriodendron* has lacked development of polymorphic and informative SSR markers.

As a result, no genetic linkage maps of *Liriodendron* have been reported. This is in contrast with the species' ecological and economic value and phylogenetic position as a basal angiosperm.

Table 2. Statistics of the 15 markers analyzed by cervus.

Locus	K	N	HO	HE	PIC
LT002	6	174	0.72	0.69	0.65
LT015	6	172	0.55	0.60	0.54
LT021	3	173	0.17	0.19	0.18
LT086	7	170	0.36	0.53	0.50
LT096	15	165	0.65	0.76	0.73
LTCU19	12	174	0.60	0.71	0.69
LTCU51	18	172	0.74	0.87	0.85
LTCU53	13	167	0.63	0.84	0.82
LTCU125	18	164	0.88	0.89	0.88
LTCU143	14	164	0.74	0.80	0.76
LTCU145	11	172	0.84	0.86	0.84
LTCU150	15	172	0.51	0.74	0.72
LTCU151	11	158	0.51	0.76	0.73
LTCU152	19	143	0.65	0.93	0.92
LTCU154	26	160	0.74	0.93	0.92
Average	13	167	0.62	0.74	0.71

K: number of alleles; N: number of individuals; HO: observed heterozygosity; He: expected heterozygosity; PIC: polymorphic information content.

Genetic composition of the *L. tulipifera* orchards

While there were only two loci (LT002 and LT015) significantly deviating from Hardy-Weinberg proportions in the Clemson population, there were 10 deviating loci in the Knoxville population ($p>0.05$) (Supplementary Table S1 and S2). This may be due to insufficient sample size from the Knoxville population. As shown in Tables 3 and 4, the Clemson orchard had higher values for observed number of alleles (15.3 vs 7.4), effective number of alleles (5.9 vs 4.0), observed (*Ho*, 0.64 vs 0.58) and expected heterozygosity (*He*, 0.74 vs 0.70), Nei's expected heterozygosity (0.74 vs 0.58), and Shannon's Information index (1.85 vs 1.51) than the Knoxville orchard. However, the differences were not statistically significant ($p=0.05$, t-Test) except for observed number of alleles. The different number of trees from the orchards included in the study, 163 from the Clemson vs 31 from the Knoxville, may be a contributing factor.

This is the first report of genetic composition of *Liriodendron* cultivated populations in North America and has provided the basic data of genetic diversity and allele richness among selections of this unique native species. [43] examined 27 trees from a cultivated population of *L. tulipifera* in the Jurong, Jiangsu Province of China with 39 polymorphic EST-SSR loci through electrophoreses in 6% polyacrylamide denaturing gels and visualization with silver nitrate staining. It was found that the number of alleles per locus ranged from three to 18 and the average *Ho* and *He* were 0.68 and 0.78, respectively. Compared to this cultivated population in China, the two

Table 3. Genetic variation at 15 EST-SSR loci characterized in the clemson orchard.

Clemson orchard (163 trees)									
Locus	Sample size	Na	Ne	Obs_Hom	Obs_Het	Exp_Hom	Exp_Het	Nei's	I
LT002	326	6.00	3.16	0.28	0.72	0.31	0.69	0.68	1.35
LT015	324	8.00	2.30	0.46	0.54	0.43	0.57	0.57	1.11
LT021	324	6.00	1.99	0.54	0.46	0.50	0.50	0.50	1.01
LT086	318	9.00	1.84	0.69	0.31	0.54	0.46	0.46	1.00
LT096	308	18.00	3.75	0.38	0.62	0.26	0.74	0.73	1.80
LTCU19	326	15.00	3.20	0.40	0.60	0.31	0.69	0.69	1.71
LTCU51	324	18.00	7.74	0.26	0.74	0.13	0.87	0.87	2.26
LTCU53	310	13.00	4.84	0.30	0.70	0.20	0.80	0.79	1.89
LTCU125	308	26.00	11.46	0.05	0.95	0.08	0.92	0.91	2.75
LTCU143	306	14.00	4.84	0.23	0.77	0.20	0.80	0.79	1.81
LTUCU145	324	12.00	7.49	0.17	0.83	0.13	0.87	0.87	2.18
LTUCU150	322	14.00	2.90	0.42	0.58	0.34	0.66	0.66	1.61
LTUCU151	294	9.00	2.99	0.54	0.46	0.33	0.67	0.67	1.38
LTUCU152	258	25.00	15.89	0.33	0.67	0.06	0.94	0.94	2.91
LTUCU154	302	37.00	14.49	0.28	0.72	0.07	0.93	0.93	3.03
Mean	312	15.33	5.93	0.36	0.64	0.26	0.74	0.74	1.85
St. Dev.	--	8.53	4.59	0.16	0.16	0.15	0.15	0.15	0.66

Na: Observed number of alleles. Ne: Effective number of alleles (Kimura and Crow 1964). Obs_Hom/Obs_Het: Observed homozygosity/heterozygosity. Ext_Het/Exp_Het: expected homozygosity/heterozygosity (Levene 1949). Nei's (1973) expected heterozygosity. I=Shannon's Information index (Lewontin 1972). St. Dev.: Standard deviation.

Table 4. Genetic variation at 15 EST-SSR loci characterized in the knoxville orchard.

Knoxville orchard (31 trees)									
Locus	Sample size	Na	Ne	Obs_Hom	Obs_Het	Exp_Hom	Exp_Het	Nei's	I
LT002	62	5	3.65	0.32	0.68	0.26	0.74	0.73	1.43
LT015	60	5	3.38	0.43	0.57	0.28	0.72	0.7	1.36
LT021	62	2	1.17	0.84	0.16	0.85	0.15	0.15	0.28
LT086	62	6	2.81	0.35	0.65	0.35	0.65	0.64	1.22
LT096	62	10	4.75	0.32	0.68	0.2	0.8	0.79	1.85
LTCU19	62	9	3.59	0.52	0.48	0.27	0.73	0.72	1.68
LTCU51	60	12	5.84	0.27	0.73	0.16	0.84	0.83	2
LTCU53	62	6	2.77	0.74	0.26	0.35	0.65	0.64	1.32
LTCU125	60	15	6.14	0.3	0.7	0.15	0.85	0.84	2.17
LTCU143	60	8	5.26	0.33	0.67	0.18	0.82	0.81	1.8
LTUCU145	60	8	4.64	0.1	0.9	0.2	0.8	0.78	1.71
LTUCU150	62	5	3.29	0.68	0.32	0.29	0.71	0.7	1.33
LTUCU151	56	3	2.26	0.39	0.61	0.43	0.57	0.56	0.89
LTUCU152	62	9	5.88	0.42	0.58	0.16	0.84	0.83	1.93
LTUCU154	58	8	3.83	0.24	0.76	0.25	0.75	0.74	1.62
Mean	61	7.4	3.95	0.42	0.58	0.29	0.70	0.58	1.51
St. Dev.	--	3.4	1.44	0.20	0.20	0.17	0.17	0.16	0.46

Na: Observed number of alleles. Ne: Effective number of alleles (Kimura and Crow 1964). Obs_Hom/Obs_Het: Observed homozygosity/heterozygosity. Ext_Het/Exp_Het: expected homozygosity/heterozygosity (Levene 1949). Nei's (1973) expected heterozygosity. I=Shannon's Information index (Lewontin 1972). St. Dev.: Standard deviation.

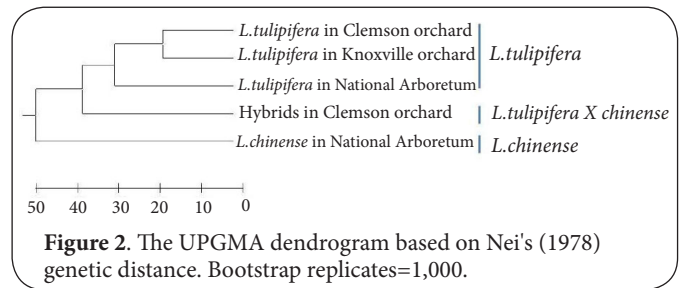
US orchards had slightly lower values of average *Ho* and *He*, with 0.64 (*Ho*) and 0.74 (*He*) in the Clemson orchard and 0.58 (*Ho*) and 0.70 (*He*) in the Knoxville orchard. These values are comparable to those reported in a *L. chinense* cultivated population in China, which had a *Ho* and *He* of 0.48 and 0.74 [43]. Other forest tree species have similar heterozygosities as well, for example, a *Pinus merkusii* parental and seedling populations had a *He* of 0.55 and 0.49, respectively [10,30]. Reported 0.48 (*Ho*) and 0.63 (*He*) in a white spruce plantation and 0.49 (*Ho*) and 0.63 (*He*) in a white spruce improvement selection population. It is noteworthy that genetic diversity of natural *L. tulipifera* populations has been reported [e.g., 18,32]. However these studies utilized either allozymes or amplified fragment length polymorphism (AFLP) markers, which usually have lower information content than SSR markers. None of the reported expected heterozygosities from these studies exceeded 0.29. Overall, substantial genetic diversity is captured in the Clemson and Knoxville seed orchards.

Conclusion

The data obtained in this study will be useful in future applications such as prediction of genetic gain and gene diversity in the seed orchards. Nei's genetic distance between the two orchards was 0.39, which was the lowest among all comparisons (Table 5). The *L. chinense* and *L. tulipifera* trees from the National Arboretum exhibited the largest genetic distance (1.17). The two orchards and the *L. tulipifera* sample from the US National Arboretum grouped together in the UPGMA dendrogram. The genetic distance of the hybrids in the Clemson orchard was closest to the Clemson orchard (0.50), followed by the Knoxville orchard (0.80) and *L. chinense* from the National Arboretum (0.88), and then by the *L. tulipifera* from the National Arboretum (1.17) (Figure 2). With a widespread range of distribution, *L. tulipifera* has adapted to many different ecological conditions and is one of the species becoming increasingly dominant in forests due to its quick respond to increases in light to the forest floor and rapid initial growth rate [8]. Its increasingly important roles in forestry and wood products is making studying *Liriodendron* of great interest. Our study provides a first look at the genetic diversity and allele richness among selections of this unique native species,

Table 5. Nei's (1978) unbiased identity (above diagonal) and distance (below diagonal).

Pop ID	Clemson	Knoxville	NA <i>L. chinense</i>	Hybrids in Clemson	NA <i>L. chinense</i>
Clemson	--	0.6792	0.6234	0.6051	0.4051
Knoxville	0.3869	--	0.4648	0.4484	0.3495
NA <i>L. tulipifera</i>	0.4725	0.7662	--	0.3608	0.3097
Hybrids in clemson	0.5025	0.8021	1.0195	--	0.4138
NA <i>L. chinense</i>	0.9037	1.0513	1.1721	0.8823	--



and provides a foundation for further genetic and breeding exploration. The polymorphic markers developed in this study will serve as a resource enabling the future study of population dynamics and adaptive variation in *Liriodendron*.

Additional files

- Supplementary Table S1
- Supplementary Table S2
- Supplementary figure S1
- Supplementary figure S2
- Supplementary figure S3

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

Authors' contributions	XZ	AC	ZT	MS	SES	JEC	HL
Research concept and design	--	--	--	--	--	✓	✓
Collection and/or assembly of data	✓	✓	✓	✓	✓	--	--
Data analysis and interpretation	✓	--	--	--	--	--	✓
Writing the article	✓	--	--	--	--	--	✓
Critical revision of the article	--	--	--	✓	✓	--	✓
Final approval of article	✓	✓	✓	✓	✓	✓	✓
Statistical analysis	✓	--	--	--	--	--	✓

Acknowledgement

The authors thank Nick Wheeler for his guidance and thorough review of the manuscript. The funding for the study came from the NSF Plant Genome Research program (NSF 1025974) and National Institute of Food and Agriculture, USDA SC-1700449 with a Clemson University Experiment Station technical contribution number of 6215).

Publication history

Editor: Shouan Zhang, University of Florida, USA.

Received: 21-Jan-2015 Revised: 24-Feb-2015

Accepted: 17-Mar-2015 Published: 27-Mar-2015

References

- Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W and Lipman DJ. **Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.** *Nucleic Acids Res.* 1997; 25:3389-402. | [Article](#) | [PubMed Abstract](#) | [PubMed Full Text](#)
- Bae K and Byun J. **Screening of leaves of higher plants for antibacterial action.** *Kor J Pharmacogn.* 1987; 8:1.
- Beck DE. *Liriodendron tulipifera* L. **Yellow-poplar.** In *Silvics of North*

- America. Harwoods. Edited by R.M. Burns and B.H. Honkala. US Dep Agric Agric Handb 654. 1990; 2:406-416. | [Article](#)
4. Beckage B and Clark JS. **Seedling survival and growth of three forest tree 328 species: The role of spatial heterogeneity.** *Ecology*. 2003; 84:1849-1861. | [Pdf](#)
5. Bonner FT and Russell TE. ***Liriodendron tulipifera* L. Yellow-poplar. In Schopmeyer, CS (Tech. Coord.). Seeds of woody plants of the United States.** USDA For Serv Agric Handb. 1974; 450:508-511.
6. Cech F, Brown J and Weingartner D. **Wind damage to a yellow-poplar seed orchard.** *Tree Planters' Notes*. 1976; 27:3-4.
7. Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, Dufayard JF, Guindon S, Lefort V, Lescot M, Claverie JM and Gascuel O. **Phylogeny.fr: robust phylogenetic analysis for the non-specialist.** *Nucleic Acids Res*. 2008; 36:W465-9. | [Article](#) | [PubMed Abstract](#) | [PubMed Full Text](#)
8. Dyer JM. **Using witness trees to assess forest change in southeastern Ohio.** *Can J Bot*. 2001; 31:1708-1718. | [Article](#)
9. Ellis JR and Burke JM. **EST-SSRs as a resource for population genetic analyses.** *Heredity (Edinb)*. 2007; 99:125-32. | [Article](#) | [PubMed](#)
10. Fageria MS and Rajora OP. **Effects of silvicultural practices on genetic diversity and population structure of white spruce in Saskatchewan.** *Tree Gene Genome*. 2014; 10:287-296. | [Article](#)
11. Forest Service. **Forest Inventory and Analysis National Program, The United States Department of Agriculture.** FIADB 5.1.6 ed. 2014. | [Website](#)
12. Han Y, Chagne D, Gasic K, Rikkerink EH, Beever JE, Gardiner SE and Korban SS. **BAC-end sequence-based SNPs and Bin mapping for rapid integration of physical and genetic maps in apple.** *Genomics*. 2009; 93:282-8. | [Article](#) | [PubMed](#)
13. Hernandez R, Davalos JF, Sonti SS, Kim Y and Moody RC. **Strength and stiffness of reinforced yellow-poplar glued laminated beams.** *Res. Pap. FPL-RP-U.S. Department of Agriculture, Forest Service, Forest Products Laboratory, Madison, WI*. 1997. | [Pdf](#)
14. Huang MR and Chen DM. **An isozyme analysis of tulip-tree hybrids (*Liriodendron Chinese X L. tulipifera*).** *J Nanjing For Univ*. 1979; 1:156-158. | [Article](#)
15. Jin H, Do J, Moon D, Noh EW, Kim W and Kwon M. **EST analysis of functional genes associated with cell wall biosynthesis and modification in the secondary xylem of the yellow poplar (*Liriodendron tulipifera*) stem during early stage of tension wood formation.** *Planta*. 2011; 234:959-77. | [Article](#) | [PubMed](#)
16. Kimura M and Crow JF. **The measurement of effective population number.** *Evolution*. 1964; 17:279-288. | [Article](#)
17. Kobayashi N, Horikishi T, Katsuyama H, Handa T and Takayanagi K. **A simple and efficient DNA extraction method for plants, especially woody plants.** *Plant Tissue Culture Biotechnol*. 1998; 4:76-80. | [Article](#)
18. Kovach KE. **Assessment of genetic variation of *Acer rubrum* L. and *Liriodendron tulipifera* L. populations in unmanaged forests of the Southeast United States.** 2009. | [Article](#)
19. Levene H. **On a matching problem arising in genetics.** *Ann Math Stat*. 1949; 20:91-94. | [Article](#)
20. Lewontin RC. **The apportionment of human diversity.** *Evol Biol*. 1972; 6:381-398. | [Pdf](#)
21. Li Z and Wang Z. **RAPD markers used for the hybrid identification and parents choice in *Liriodendron*.** *Sci Silv Sinic*. 2002; 38:169-174. | [Article](#)
22. Liang H, Ayyampalayam S, Wickett N, Barakat A, Xu Y, Landherr L, Ralph P, Xu T, Schlarbaum SE, Leebens-Mack JH and dePamphilis CW. **Generation of a large-scale genomic resource for functional and comparative genomics in *Liriodendron*.** *Tree Gen Genom*. 2011; 7:941-954. | [Article](#)
23. Liang H, Carlson JE, Leebens-Mack JH, Wall PK, Mueller LA, Buzgo M, Landherr LL, Hu Y, DiLoreto DS, Ilut DC, Field D, Tanksley SD, Ma H and dePamphilis CW. **A 378 n EST Database for *Liriodendron tulipifera* L. floral buds: the first EST resource for functional and comparative genomics in *Liriodendron*.** *Tree Genet Genomes*. 2008; 4:419-433. | [Article](#)
24. Liang H, Feng E, Tomkins J, Arumuganathan K, Zhao S, Luo M, Kudrna D, Wing R, Banks J, dePamphilis C, Mandoli D, Schlarbaum S and Carlson JE. **Development of a BAC library resource for yellow poplar (*Liriodendron tulipifera*) and the identification of genomic regions associated with flower development and lignin biosynthesis.** *Tree Genet Genomes*. 2007; 3:215-225. | [Article](#)
25. Marshall TC, Slate J, Kruuk LE and Pemberton JM. **Statistical confidence for likelihood-based paternity inference in natural populations.** *Mol Ecol*. 1998; 7:639-55. | [Article](#) | [PubMed](#)
26. Merkle SA, Hoey MT, Watson-Pauley BA and Schlarbaum SE. **Propagation of *Liriodendron* hybrids via somatic embryogenesis.** *Plant Cell Tissue Organ Cult*. 1993; 34:191-198. | [Article](#)
27. Moon MK, Oh HM, Kwon BM, Baek NI, Kim SH, Kim JS and Kim DK. **Farne-syl protein transferase and tumor cell growth inhibitory activities of lipiferolide isolated from *Liriodendron tulipifera*.** *Arch Pharm Res*. 2007; 30:299-302. | [Article](#) | [PubMed](#)
28. Nei M. **Analysis of gene diversity in subdivided populations.** *Proc Natl Acad Sci U S A*. 1973; 70:3321-3. | [PubMed Abstract](#) | [PubMed Full Text](#)
29. Nei M. **Estimation of average heterozygosity and genetic distance from a small number of individuals.** *Genetics*. 1978; 89:583-90. | [Article](#) | [PubMed Abstract](#) | [PubMed Full Text](#)
30. Nurtjahjaningsih ILG, Saito Y, Tsuda Y and Ide Y. **Genetic diversity of parental and offspring populations in a *Pinus merkusii* seedling seed orchard detected by microsatellite markers.** *Bulletin of the Tokyo University Forest, the Tokyo University Forests*. 2007; 118:1-14. | [Pdf](#)
31. Parks CR, Miller NG, Wendel JF and Mc-Dougal KM. **Genetic divergence within the genus *Liriodendron* (Magnoliaceae).** *Ann Miss Bot Gard*. 1983; 70:658-666. | [Article](#)
32. Parks CR and Wendel JF. **Molecular divergence between Asian and North American species of *Liriodendron* (Magnoliaceae) with implications for interpretation of fossil floras.** *Am J Bot*. 1990; 77:1243-1256. | [Article](#)
33. Raymond M and Rousset F. **GENEPOP (version 1.2): population genetics software for exact tests and ecumenisms.** *J Heredity*. 1995; 86:248-249. | [Article](#)
34. Santamour FS. **Interspecific hybrids in *Liriodendron* and their chemical verification.** *Fort Sci*. 1972; 18:233-236. | [Article](#)
35. Sewell MM, Parks CR and Chase MW. **Intraspecific chloroplast DNA variation and biogeography of North American *Liriodendron* L. (Magnoliaceae).** *Evolution*. 1996; 50:1147-1154. | [Article](#)
36. Thor E. **Tree Breeding at the University of Tennessee, 1959-1976.** *Univ Tenn Ag Exp Stn Bull*, 1976; 48. | [Article](#)
37. Turmel M, Otis C and Lemieux C. **The chloroplast genome sequence of *Chara vulgaris* sheds new light into the closest green algal relatives of land plants.** *Mol Biol Evol*. 2006; 23:1324-38. | [Article](#) | [PubMed](#)
38. van Oosterhout C, Hutchinson WF, Wills DPM and Shipley P. **MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data.** *Mol Ecol Notes*. 2004; 4:535-538. | [Article](#)
39. Wang Z. **Utilization and species hybridization in *Liriodendron*.** Chinese Forestry Press, Beijing. 2005.
40. Wicke S and Quandt D. **Universal primers for the amplification of the plastid trnK/matK region in land plants.** *Anales Jard Bot Madrid*. 2009; 66:285-288. | [Article](#)
41. Williams RS and Feist WC. **Durability of yellow-poplar and sweetgum and service life of finishes after long-term exposure.** *Forest Products J*. 2004; 54:96-101. | [Article](#)
42. Xu M, Sun Y and Li H. **EST-SSRs development and paternity analysis for *Liriodendron* spp.** *New Forests*. 2010; 40:361-382. | [Article](#)
43. Yang AH, Zhang JJ, Tian H and Yao XH. **Characterization of 39 novel EST-SSR markers for *Liriodendron tulipifera* and cross-species amplification in *L. chinense* (Magnoliaceae).** *Am J Bot*. 2012; 99:e460-4. | [Article](#) | [PubMed](#)
44. Yu JK, La Rota M, Kantety RV and Sorrells ME. **EST derived SSR markers for comparative mapping in wheat and rice.** *Mol Genet Genomics*. 2004; 271:742-51. | [Article](#) | [PubMed](#)
45. Zane L, Bargelloni L and Patarnello T. **Strategies for microsatellite isolation: a review.** *Mol Ecol*. 2002; 11:1-16. | [Article](#) | [PubMed](#)
46. Zhang LJ, Oswald BP and Green TH. **Relationships between overstorey species and community classification of the Sipsey Wilderness, Ala-**

bama. *For Ecol Manage.* 1999; **114**:377-383. | [Article](#)

47. Zhang H, Li H, Xu M and Feng Y. **Identification of *Liriodendron tulipifera*, *Liriodendron chinense* and hybrid *Liriodendron* using species-specific SSR markers.** *Sci Silv Sinic.* 2012; **46**:36-39.

Citation:

Zhang X, Carlson A, Tian Z, Staton M, Schlarbaum SE, Carlson JE and Liang H. **Genetic characterization of *Liriodendron* seed orchards with EST-SSR markers.** *J Plant Sci Mol Breed.* 2015; **4**:1.
<http://dx.doi.org/10.7243/2050-2389-4-1>