United States
Department of
Agriculture

Forest Service

**Southern
Research Station**

e-Research Note
SRS–24

April 2018

DNA Fingerprinting Sets for Four Southern Pines

Craig Echt and Sedley Josserand

Abstract

DNA markers can provide valuable genetic information for forest tree research, breeding, conservation, and restoration programs. When properly evaluated, selected sets of DNA markers can be used to efficiently get information about genetic diversity in regions, forests, or stands, or in seed lots and orchards. Selected markers also can be used to determine parentage or verify clonal identity of individual trees in tree improvement programs and seed production orchards. With these purposes in mind, we developed sets of informative markers for four species of southern pines: shortleaf pine (*Pinus echinata* Mill.), slash pine (*P. elliottii* Englem.), longleaf pine (*P. palustris* Mill.), and loblolly pine (*P. taeda* L.). We selected 38 markers for all 4 species and then selected markers within each species: 9 for shortleaf pine, 16 for slash pine, 10 for longleaf pine, and 11 for loblolly pine. All markers were originally developed for use in loblolly pine. Once we optimized marker sets for each species, however, there were few markers left in common between the different marker sets. In this report, we provide all the methods and information needed for any molecular biology lab to use the markers in any of the four pine species.

Keywords: Genotyping, microsatellite markers, pine genetics.

Introduction

The most common forest and plantation tree species in the Southern United States are pines, specifically shortleaf pine (*Pinus echinata* Mill.), slash pine (*P. elliottii* Englem.), longleaf pine (*P. palustris* Mill.), and loblolly pine (*P. taeda* L.), which in 2012 accounted for 84 percent of all forest planting stock (Oswalt and others 2014). Longleaf pine and shortleaf pine are also the focus of major restoration initiatives (America's Longleaf Restoration Initiative, Shortleaf Pine Initiative). For restoration efforts in particular, it is important to have the means to assess the genetic integrity of planting stocks (Echt and others 2013, Mijangos and others 2015).

Simple sequence repeats (SSRs) of DNA, also known as microsatellite DNA, are very useful as molecular markers, particularly for establishing genetic identity of individuals and populations (Grover and Sharma 2016, Nybom and others 2014). In pines, SSR markers have been used for integrated genome mapping of loblolly and slash pines (Echt and others 2011, Westbrook and others 2015). The genetically informative nature of SSR markers makes them especially useful for determining parentage, verifying clonal identity of trees in breeding and seed production orchards, and measuring genetic diversity in populations for conservation and restoration management. SSR markers developed for loblolly pine have been applied to genetic studies in shortleaf pine and longleaf pine, two other pines that are also in the Australes taxonomic subsection of the *Pinus* genus (Josserand and others 2011, Koppleman and others 2007, Nelson and others 2007, Stewart and others 2010).

We have developed sets of SSR markers to use in conservation genetic studies and individual tree genotype verification for four pine species: shortleaf pine, slash pine, longleaf pine, and loblolly pine. After evaluating hundreds of SSR markers, we selected sets of about a dozen markers for each pine species that are suitable for general genotyping of populations and individuals. The resulting DNA fingerprinting sets were used to measure genetic diversity in sample populations.

Craig Echt, Research Plant Molecular Geneticist, and **Sedley Josserand**, Biological Technician, U.S. Department of Agriculture, Forest Service, Southern Research Station, Forest Genetics and Biological Foundations Work Unit, 23332 Success Road, Saucier, MS 39574.

Methods

Needle samples were taken either from clones of individual trees in breeding and seed orchards or from natural populations. The set of samples for each species was constructed as a synthetic population intended to be representative of the genetic diversity typically seen throughout the natural range, in breeding populations, or in restoration seed sources. We analyzed 39 samples from shortleaf pine, 30 from slash pine, 34 from longleaf pine, and 27 from loblolly pine. These sample sets were established at different times and were from various unrelated genotyping projects for different clients and collaborators; appendix table A.1 lists sample names, sources, and States of provenance for each species.

DNA was purified from needle tissue. DNA samples for shortleaf pine were provided by the U.S. Department of Agriculture Forest Service National Forest Genetics Laboratory (NFGEL) with material sourced from various Region 8 (Southern Region) National Forest seed orchards (R8 NFSO). Slash pine DNA samples were provided by the University of Florida Cooperative Forest Genetics Research Program, which included materials sourced from the Western Gulf Tree Improvement Program. Some longleaf pine DNA were provided by NFGEL from R8 NFSO material, and some longleaf pine DNA samples were extracted in our lab using the DNAeasy Plant Mini Kit (Qiagen). All loblolly DNA samples were extracted in our lab with material sourced from ArborGen, Inc. and the Forest Service Harrison Experimental Forest clonal archive.

All DNA concentrations were quantified by ultraviolet (UV) absorbance using a NanoDrop™ spectrophotometer (Thermo Scientific™) and standardized to 10 ng/µl for use in a polymerase chain reaction (PCR). All PCR forward primer oligonucleotides included on their 5′ end the M13-forward (-29) sequence CACGACGTTGTAAAACGAC. All reverse primer oligonucleotides included on their 5′ end the sequence GTTTCTT, which forces a non-templated dA addition to the amplified fragments (Brownstein and others 1996). Primer pairs containing both of these tail sequences added 26 base pairs (bp) to the size of amplified fragments. Fluorescent dye (6-FAM, VIC, NED, or PET) was incorporated into amplicons by including a 5′ dye-labelled M13-forward (-29) primer in the PCR (Schuelke 2000). We used the following PCR reagent composition in a 10 µl reaction volume: 20 ng pine DNA (dried), 40 nM forward primer, 160 nM reverse primer, 160 nM dye-labelled primer, 66 µM dNTPs, pH 9.0 buffer (2.0 mM $MgCl_2$, 10 mM Tris-Cl, 50 mM KCl), and approximately one unit of Platinum® *Taq* polymerase (Life Technologies Corporation, Carlsbad, CA, USA). For all markers we used the following PCR thermocycling protocol: 2 minutes at 94 °C; followed by 20 cycles of 30 seconds at 94 °C, 30 seconds at x °C, and 1 minute at 72 °C, where x = 65 °C -0.5 °C per cycle; followed by 24 cycles of 30 seconds at 92 °C, 30 seconds at 55 °C, 1.5 minutes at 72 °C; followed by 15 minutes at 72 °C. Completed reactions were refrigerated until analyzed.

PCR amplification products and ABI PRISM® GS 600 LIZ® internal size standards (Life Technologies Corporation) were separated by capillary electrophoresis using POP-4® polymer in an ABI PRISM® 3130xl Genetic Analyzer (Life Technologies Corporation). Electrophoresis was run with a 36-cm capillary array using the following instrument Run Module settings as they appear in the machine parameter menu: oven temp 60, polymer fill vol 6500, current stability 5, PreRun voltage 15, PreRun time 180, injection voltage 1.2, injection time 18, voltage number of steps 40, voltage step interval 15, data delay time 1, run voltage 15, run time 1830.

For electrophoresis, markers were pooled four to a capillary channel (each marker labeled with a different fluorescent dye) and grouped in sets such that their expected allele size ranges in a population did not overlap. Allele sizing using a third order least-squares algorithm and allele size binning were both determined with ABI PRISM® GENEMAPPER® 4.0 software (Life Technologies Corporation), followed by manual inspection and editing of the allele assignments, as needed. We standardized SSR marker allele identification between capillary electrophoresis runs or with different dye labels with the aid of a control genotype sample (loblolly pine clone 20-1010 or 7-56).

PCR primer information for all markers, except PtRIP_0031 and PtTX4114, was previously reported (Echt and others 2011). PtRIP_0031 was developed for use in loblolly pine, but was first reported for use in shortleaf pine (Nelson and others 2007). PtTX4114 and all other PtTX markers were first reported as loblolly pine SSR markers by Auckland and others (2002). Primer and SSR motif sequences for each marker are listed in table 1 for convenient reference.

Table 1—SSR marker primer pair DNA sequences and repeat sequence motifs

Marker ID	Forward primer sequence	Reverse primer sequence	SSR motif
NZPR0143	GAAAGCATTAGCCATCTACATTCA	TCATTGTGCATGCATTTATAATCTC	AG
NZPR0563	GCATTTCTTGTTGCTATTTTCAA	GCACAAGTCCCATTTCCATT	AC
PpSIFG_3147	CACGTGGTTTCCTCCAGTTT	CAATGCGTTCTGCATATTGG	AT
PtRIP_0031	CCAACCAATGTGGTTCATCA	AGGAAAATAGAAGGGAATAAGACC	AAT
PtRIP_0079	TGATTTGATCCCTCTAGGCG	AATCTTGAAAAGAAATTCAATATGAGA	AAT
PtRIP_0211	GACGAGGGGGTCTCATACACCAA	TGCATAGAGGATGTATTTCTTGGA	AAT
PtRIP_0968	TCTACGACAAAACCACGTAGTG	CATGTGGCTTTGTGGCATAT	AC
PtRIP_0984	TGTGACCTGAAAATTCCCCT	GGCTTGCAACCAGTTCCATA	AC
PtRIP_1077	AACATTCTAGCATGCCCCAC	TTGTGGTGGATGTCTCTCCTC	AC, AT
PtSIFG_0193	CCCATGCATCAATTCAAGTT	TGTGCGTGGATATGGAAAAA	AT
PtSIFG_0408	ACATCCCTCAATCATGCAAA	TGAGGCCAAGCTCGATAACT	AAT
PtSIFG_0437	TCTATGATGGAAGGCCCAAC	GTTCTGCTTGCCCTCTCAAC	CTG
PtSIFG_0493	GAGAACATCTGCCTTGAGCC	CTGGCATGATGGGTTTCTCT	CTG
PtSIFG_0561	GCCAACTGCAATAACAGCAA	CCGGCAAGAGCATCATTATT	GCAGAA
PtSIFG_0566	ACTTAGTGGGAAAGGGGGAA	TTCCTCAGCCAAAAGCTCTC	GGGAAG
PtSIFG_0629	CATGGGCGAGATCAAGAGAT	GAAAGGAAAGGAAACCTCCG	AACGGA
PtSIFG_0737	GCAAGGGGAATTGCTTATGA	GGGATCGCATCAGCTGTAAT	CAG, CAGCAT
PtSIFG_0745	AAGAAGGGCGGACTAGGAGC	GTGAACCCACAATTCCCAAC	AGGTTG, GGCTGA
PtSIFG_1008	GAGAACATCTGCCTTGAGCC	CTGGCATGATGGGTTTCTCT	CTG
PtSIFG_1190	CAGGTGGCTTGGATTTCATT	TCATTCAAGCGTCCTGCTTA	TCC
PtSIFG_4102	CTTTGTTGACCCCTGCATTT	TTGGCTTAGCTAAAAGGGTGA	AT
PtSIFG_4218	AAAGGCAGCAGTCGGTAGAA	AAACCAAGTTTGCCTGATCG	AT
PtSIFG_4222	CACCCTTTTCACGCAAGAAT	GCCTACGCATTATCCTTCCA	AT
PtSIFG_4232	GAAAAAGAAGAGAAGAATCAACGC	CTTCAAATGCCCTTCGACAT	AT
PtSIFG_4233	AGGGAAACCGCGGATTATAG	CCGGAATGAAGATTGCAGTT	AT
PtSIFG_4249	TCCTTGGTTTGTGCTTTTCC	ATATGCCCGTTGGCAGTAAC	AT
PtSIFG_4304	CATGCATGTGTGGAGGAGTT	CTCATGTGCTTTGATCCCCT	AG
PtSIFG_4380	CCTATCCCACAAAGACGGAA	AACTCAAAAACCTGGGGCTT	AT
PtSIFG_4502	AGCGTAATCAACTGGGAACG	GTTCATTCATGCTCAGCGAA	AT
PtTX3013	GCTTCTCCATTAACTAATTCTA	TCAAAATTGTTCGTAAAACCTC	AAC
PtTX3034	TCAAAATGCAAAAGACG	ATTAGGACTGGGGATGAT	AG
PtTX3052	CCTCACTAGGAGGCTACGGAAGAG	AAAGACTCCTTGATGTTGTGAACA	ATC
PtTX3081	GCCGAGGAAGCAAGCAACCAA	CCTCGGCAGCCAAATCCTTCA	AAC
PtTX4003	GCGATAAGCATACCTACACT	TTACTAAAATGGGGATGAAA	AC
PtTX4058	AAGTGTTGGGAGAAAAATGTAAT	CTCCTTCTGTCCCTATCCTCT	AG
PtTX4092	GGATGATACTTTCCATGAGTTAGG	TCTAGTCCAGATCTTGGTCCAC	AAG
PtTX4114	ACACATGTCTTGAGGAGT	TCAATTTGATCTATAACTTTCACC	AG
SsrPt_ctg9249	CTGCTCCCTCAGCTCTTCC	AGACGTCACTGCCATTACCC	AAG

Allele sizes for loblolly pine reference genotypes, clones 20-1010 and 7-56, are provided in table 2 to aid calibration and standardization of allele sizes by laboratories using different allele detection methods, genotyping platforms, or PCR protocols that might affect measured values of allele lengths. Samples of the two reference genotypes are readily available on request from the clonal archives at the Harrison Experimental Forest, Saucier, MS.

GenBank® record identifiers of the DNA sequences used to develop PCR primers for all markers selected for the current report are listed in table 2, along with the linkage group positions for 32 markers that were previously mapped in loblolly pine (Echt and others 2011, Westbrook and others 2015).

Population allele frequency metrics were used to characterize genetic diversity in each sample population. For individual markers and their population means, we calculated number of alleles (N_a), effective allele number (N_e), observed frequency of heterozygosity (H_o), expected frequency of heterozygosity (H_e), fixation index *(F)*, and probability of identity *(PI)* using the program GenAlEx 6.501 software (Peakall and Smouse 2012). For a marker set, the multilocus *PI* is the product of the individual marker *PI* values. Estimated individual marker null allele frequencies (p_{null}) and the null-corrected population inbreeding coefficient *(null-corrected F)* for each marker set were calculated simultaneously with a population inbreeding model using maximum likelihood, as implemented in the program INEst 1.0 (Chybicki and Burczyk 2009).

Results

To develop DNA fingerprinting sets for the different southern pine species, we evaluated hundreds of pine SSR markers over the course of several independent projects, all of which had been previously screened for use in loblolly pine (Echt and others 2011). Markers for this study were selected based on their consistency of PCR amplifications to provide easily scored allelic profiles and which had informative levels of allelic diversity in a diverse population of trees. For all four species, we selected 38 markers (tables 1 and 2). The number of markers selected in each species is 9 for shortleaf pine, 16 for slash pine, 10 for longleaf pine, and 11 for loblolly pine (table 3). Because the markers we selected were those best suited for use in a particular species, we found few markers that amplified equally well in two or more species or were equally genetically informative in more than one species. For example, although all markers amplified in loblolly pine, only one of the eleven selected loblolly pine markers was used in another species (PtSIFG_0493 with shortleaf pine); across all four species, only seven markers were used in two or more species (table 3).

We report various genetic diversity metrics for each marker and species in table 3. To easily compare genetic diversity among populations, effective allele number (N_e) is often more useful than the standard of expected heterozygosity (H_e) because N_e scales linearly with increasing allele diversity, while H_e, being bound between zero and one, scales asymptotically; this difference is especially evident with the more polymorphic markers. Observed and expected heterozygosity, H_o and H_e, respectively, are traditional measures of genetic diversity and used to calculate the fixation index, *F*. The average *F* value of a population is a useful metric for conservation genetic studies to compare levels of inbreeding among natural populations or in seed stores intended for forest restoration; values near zero indicate random mating (no inbreeding) within a population. Inbreeding values, however, are subject to genotype miscoding due to the presence in a population of null SSR alleles, which are alleles that do not amplify in PCR (DeWoody and others 2006). When paired in an individual with a standard allele, the heterozygous genotype will be scored as if it were homozygous for the standard allele, thus erroneously increasing H_o and *F*. When at low frequency in a population, a null allele will have negligible effect on N_e or H_e, but can have a noticeable impact on *F* or parentage analysis. To aid in selecting particular markers for particular purposes, we provide in table 3 the estimated null allele frequency for each marker. Table 3 also shows that the simultaneously estimated *null-corrected F* population inbreeding coefficient for each species takes on a value of zero, which indicates a zero probability that two alleles at a random locus in a random individual in a population are identical by descent (Chybicki and Burczyk 2009).

The probability of identity *(PI)* of a marker estimates the average probability that two unrelated individuals will have the same genotype when drawn from the same randomly mating population (Peakall and Smouse 2012). When multiple loci are used, individual *PI* values are multiplied together to provide the multilocus *PI* for the population. The multilocus *PI* listed for each species in table 3 is the product of the individual marker *PI* values, so the more markers considered, the lower the multilocus *PI* value. To adjust for different number of markers selected for each species, we standardized *PI* for each species by using just the eight markers with the lowest *PI* (the most genetically informative markers): 8.4×10^{-9} for shortleaf, 4.8×10^{-10} for slash, 4.8×10^{-8} for longleaf, and 9.5×10^{-7} for loblolly. What this tells us is that for any marker set of a species, the best eight

Table 2—SSR marker sequence record identifier, linkage group, and reference genotypes

Marker ID	GenBank® ID[a]	LG[b]	20-1010 alleles[c]	7-56 alleles[c]
NZPR0143	Pr032754273	2	126/130	124/130
NZPR0563	BV728916	8	289/289	271/293
PpSIFG_3147	BV728652	6	184/190	186/190
PtRIP_0031	BV683043	NA	251/276	260/285
PtRIP_0079	BV683053	12	159/159	156/175
PtRIP_0211	BV683076	1	162/168	168/177
PtRIP_0968	BV683124	2	223/227	217/231
PtRIP_0984	BV683125	1	235/248	244/252
PtRIP_1077	BV683137	4	247/247	241/247
PtSIFG_0193	BV728742	11	256/258	256/258
PtSIFG_0408	BV728749	9	342/342	347/347
PtSIFG_0437	BV728751	3	202/202	202/202
PtSIFG_0493	BV728661	2	314/317	314/317
PtSIFG_0561	BV728753	12	456/456	466/466
PtSIFG_0566	BV728755	1	129/129	135/135
PtSIFG_0629	BV728666	9	163/163	157/157
PtSIFG_0737	BV728669	10	450/462	450/456
PtSIFG_0745	BV728671	10	509/509	503/509
PtSIFG_1008	BV728723	NA	221/221	221/221
PtSIFG_1190	BV728679	9	312/312	312/312
PtSIFG_4102	BV728704	NA	221/221	221/221
PtSIFG_4218	BV728761	1	225/225	223/225
PtSIFG_4222	BV728762	8	346/346	342/346
PtSIFG_4232	BV728785	NA	291/291	291/291
PtSIFG_4233	BV728685	7	122/124	122/136
PtSIFG_4249	BV728687	5	382/382	382/390
PtSIFG_4304	BV728793	12	418/420	420/420
PtSIFG_4380	BV728709	NA	444/444	444/444
PtSIFG_4502	BV728695	5	349/352	349/349
PtTX3013	BV728853	12	154/154	154/154
PtTX3034	BV728857	8	220/224	227/230
PtTX3052	BV728827	11	266/286	266/286
PtTX3081	BV728831	9	261/265	260/260
PtTX4003	BV728866	5	185/185	181/181
PtTX4058	BV728868	5	164/168	164/174
PtTX4092	BV728870	10	159/162	162/180
PtTX4114	BV728876	NA	130/134	136/138
SsrPt_ctg9249	BV728813	5	178/178	175/175

[a] GenBank® ID prefixes are specific to these National Center for Biotechnology Information databases: Pr – Probe, BV – Nucleotide.

[b] LG = linkage group number genetically mapped in *Pinus taeda*.

[c] Clones 20-1010 and 7-56 are reference *P. taeda* genotypes; observed allele sizes in base pairs (bp) include the 26 bp length of the primer tails used for PCR.

NA = not available.

Table 3—Population genetic statistics for SSR markers in four southern pine species

Species, sample count	Marker ID	N_a	N_e	H_o	H_e	F	PI	Allele size range	p_{null}
P. echinata, *N* = 39	PtRIP_0031	20	10.6	0.92	0.92	0.00	0.02	248–311	0.00
	PtRIP_0079	13	5.2	0.77	0.82	0.06	0.06	153–196	0.01
	PtRIP_0211	10	4.9	0.65	0.81	0.20	0.07	159–186	0.07
	PtSIFG_0437	5	2.0	0.44	0.51	0.14	0.30	199–207	0.05
	PtSIFG_0493	3	1.9	0.51	0.48	-0.06	0.37	300–317	0.00
	PtSIFG_1008	4	1.3	0.23	0.22	-0.07	0.63	218–227	0.00
	PtTX3013	6	2.9	0.64	0.67	0.04	0.18	147–167	0.02
	PtTX3034	8	5.7	0.86	0.84	-0.02	0.05	212–232	0.00
	PtTX3052	8	3.9	0.72	0.76	0.05	0.10	237–283	0.02
				mean[a]			*multilocus Pi*[b]		*null-corrected F*[c]
		8.6	4.3	0.64	0.67	0.04	5.3×10^{-9}		0.00
P. elliottii, *N* = 30	NZPR0143	10	5.8	0.73	0.84	0.13	0.05	120–146	0.04
	NZPR0563	5	3.1	0.63	0.69	0.08	0.16	163–285	0.02
	PtRIP_0079	5	3.3	0.76	0.71	-0.07	0.13	163–175	0.00
	PtRIP_0968	12	8.3	0.90	0.90	-0.01	0.03	208–242	0.00
	PtSIFG_0408	5	2.4	0.67	0.59	-0.12	0.23	331–345	0.00
	PtSIFG_0561	3	2.1	0.50	0.53	0.05	0.35	461–473	0.01
	PtSIFG_0629	4	1.3	0.23	0.22	-0.08	0.63	152–170	0.00
	PtSIFG_0745	5	2.5	0.67	0.61	-0.09	0.23	496–522	0.00
	PtSIFG_4218	8	6.5	0.87	0.86	-0.01	0.04	222–238	0.00
	PtSIFG_4232	3	2.0	0.54	0.50	-0.07	0.36	262–290	0.00
	PtSIFG_4249	4	1.5	0.30	0.32	0.07	0.48	378–388	0.03
	PtSIFG_4380	6	3.0	0.66	0.68	0.04	0.16	441–453	0.04
	PtSIFG_4502	4	1.9	0.43	0.49	0.12	0.33	346–352	0.00
	PtTX3052	6	1.9	0.57	0.49	-0.16	0.30	266–289	0.00
	PtTX3081	10	6.9	0.90	0.87	-0.03	0.04	230–270	0.00
	PtTX4092	12	4.6	0.80	0.80	0.00	0.06	139–172	0.00
				mean			*multilocus PI*		*null-corrected F*
		6.4	3.6	0.64	0.63	-0.01	9.6×10^{-14}		0.00
P. palustris, *N* = 34	NZPR0143	16	9.1	0.85	0.90	0.06	0.02	118–152	0.01
	PpSIFG_3147	5	2.5	0.53	0.61	0.13	0.21	183–191	0.05
	PtRIP_0984	6	2.5	0.65	0.62	-0.05	0.20	237–247	0.00
	PtSIFG_0561	4	1.9	0.50	0.48	-0.03	0.33	442–466	0.00
	PtSIFG_0745	6	4.1	0.79	0.77	-0.04	0.10	492–522	0.01
	PtSIFG_4102	8	5.1	0.71	0.82	0.14	0.06	217–237	0.03
	PtSIFG_4218	5	3.7	0.74	0.74	0.00	0.12	229–237	0.00
	PtTX3052	4	1.9	0.47	0.48	0.03	0.36	264–285	0.00
	PtTX4003	4	1.5	0.41	0.36	-0.15	0.44	175–181	0.00
	PtTX4058	8	2.1	0.50	0.53	0.06	0.24	149–171	0.00
				mean			*multilocus PI*		*null-corrected F*
		6.6	3.4	0.62	0.63	0.02	7.6×10^{-9}		0.00

continued

Table 3 (continued)—Population genetic statistics for SSR markers in four southern pine species

Species, sample count	Marker ID	N_a	N_e	H_o	H_e	*F*	*PI*	Allele size range	p_{null}
P. taeda, *N* = 27	PtRIP_1077	10	4.0	0.81	0.77	-0.07	0.09	233–263	0.00
	PtSIFG_0193	4	1.9	0.52	0.47	-0.10	0.35	253–260	0.00
	PtSIFG_0493	3	2.2	0.59	0.56	-0.06	0.30	308–317	0.00
	PtSIFG_0566	4	2.9	0.63	0.66	0.05	0.19	123–141	0.01
	PtSIFG_0737	6	3.4	0.81	0.72	-0.14	0.13	438–462	0.00
	PtSIFG_1190	4	1.6	0.37	0.38	0.03	0.41	303–312	0.02
	PtSIFG_4222	3	1.9	0.44	0.49	0.09	0.32	342–346	0.03
	PtSIFG_4233	5	2.1	0.63	0.54	-0.18	0.27	122–138	0.00
	PtSIFG_4304	3	1.9	0.56	0.48	-0.15	0.33	416–420	0.00
	PtTX4114	9	6.1	0.89	0.85	-0.04	0.05	130–148	0.00
	SsrPt_ctg9249	4	2.0	0.56	0.51	-0.09	0.34	175–184	0.00
				mean			*multilocus PI*		*null-corrected F*
		5.0	2.7	0.62	0.58	-0.06	4.6×10^{-8}		0.00

N_a = number of alleles; N_e = effective allele number; H_o = observed frequency of heterozygosity; H_e = expected frequency of heterozygosity; *F* = fixation index; p_{null} = estimated null allele frequency; *PI* = probability of identity

Allele size range = observed allele sizes in base pairs for the sample population; includes 26 bp from the forward and reverse primer tails.

[a] Means were calculated from values that we rounded to two significant digits and thus may differ slightly from the raw results of the respective analysis software packages.

[b] Multilocus PI is the product of individual marker PI values and was calculated from values that we rounded to two significant digits and thus may differ slightly from the raw results of the respective analysis software packages.

[c] Null-corrected F is the population inbreeding coefficient estimated with null allele frequencies.

markers alone can provide a multilocus genotype that has less than a million-to-one chance of being duplicated within a rather modestly sized random mating population. While multilocus *PI* values reported here are specific to the sample populations, they can be considered generally indicative of the level of genotype discrimination that these markers can provide to studies of any population of random individuals.

The allele size range for each marker in each sample population is provided in table 3 to assist other labs for pooling sets of markers in ways that can increase genotyping efficiency (so that alleles of a marker do not overlap and confound genotyping efforts). For example, we constructed multiplexed sets of longleaf pine markers so that its ten markers can be run in just four PCR reactions, each containing two or three primer pairs. Following PCR, the four reactions were mixed and all ten markers run in a single electrophoresis lane, each being unambiguously distinguished by both dye color and fragment size. The longleaf pine markers within each of the following four sets were labeled with a different dye primer for multiplexed PCR: NZPR0143, PtSIFG_0745, and PtSIFG_4102; PtSIFG_0561, PtTX3052, and PtTX4058; PpSIFG_3147 and PtSIFG_4218; and PtRIP_0984 and PtTX3052. To find sets of primers that work well in a single PCR, markers were sorted by the size ranges of their alleles, and those with non-overlapping ranges were grouped. Then, those within a size-selected group that had similar amplification efficiency (typical peak height of chromatographic profile after separation by capillary electrophoresis) were grouped into candidate sets of primer pairs for testing in multiplexed PCR. Those sets that had all markers amplify well were selected, typically sets of two or three primer pairs; attempts to find higher levels of multiplexing were not successful. This general approach could be used to create multiplex sets for the other species.

Acknowledgments

We thank ArborGen, Inc., the National Forest Genetics Laboratory, University of Florida Cooperative Forest Genetics Research Program, the Virginia Department of Forestry, and the Western Gulf Tree Improvement Program for use of samples and provenance data. We also thank Valerie Hipkins, C. Dana Nelson, and Kevin Potter for their thoughtful and very helpful reviews.

Literature Cited

America's Longleaf Restoration Initiative [website]. 2017. http://www.americaslongleaf.org/. [Date accessed: December 15, 2017].

Auckland, L.D.; Bui, T.; Zhou, Y.; Williams, C.G. 2002. Conifer microsatellite handbook. Raleigh, NC: Corporate Press. 57 p.

Brownstein, M.J.; Carpten, J.D.; Smith, J.R. 1996. Modulation of non-templated nucleotide addition by *Taq* DNA polymerase: primer modifications that facilitate genotyping. BioTechniques. 20: 1004–1010.

Chybicki, I.J.; Burczyk, J. 2009. Simultaneous estimation of null alleles and inbreeding coefficients. Journal of Heredity. 100(1): 106–113.

Dewoody, J.; Nason, J.D.; Hipkins, V.D. 2006. Mitigating scoring errors in microsatellite data from wild populations. Molecular Ecology Resources. 6(4): 951–957.

Echt, C.; Josserand, S.; Hipkins, V.; Crane, B. 2013. DNA marker approaches to species conservation and restoration. In: Proceedings of the 32nd Southern Forest Tree Improvement Conference; June 10–13, 2013; Clemson, SC. [Place of publication unknown]: [Publisher unknown]: 47–51. http://www.sftic.org. [Date accessed: December 13, 2017].

Echt, C.S.; Saha, S.; Krutovsky, K.; Wimalanathan, K. [and others]. 2011. An annotated genetic map of loblolly pine based on microsatellite and cDNA markers. BMC Genetics. 12: 17. http://treesearch.fs.fed.us/pubs/38008. [Date accessed: December 13, 2017].

Grover, A.; Sharma, P.C. 2016. Development and use of molecular markers: past and present. Critical Reviews in Biotechnology. 36(2): 290–302. doi:10.3109/07388551.2014.959891. [Date accessed: December 15, 2017].

Josserand, S.; Echt, C.; Deemer, D.; Nelson, C. 2011. A DNA marker resource for longleaf pine genetics [poster; published abstract]. In: Proceedings of the 31st Southern Forest Tree Improvement Conference; June 2011; Biloxi; MS. [Place of publication unknown]: [Publisher unknown]: 127. http://www.sftic.org. http://www.sftic.org/f/SFTIC%2031%20Proceedings-2011.pdf. [Date accessed: December 19, 2017].

Koppelman, J.; Parsons, E; Scott; B; Collantes, J. [and others]. 2007. Phylogeographic analyses and evaluation of shortleaf pine population structure in Missouri. In: Kabrick, J.M.; Dey, D.C.; Gwaze, D., eds. Shortleaf pine restoration and ecology in the Ozarks: proceedings of a symposium; November 7–9, 2006; Springfield, MO. Gen. Tech. Rep. NRS-P-15. Newtown Square; PA: U.S. Department of Agriculture, Forest Service, Northern Research Station: 104.

Mijangos, J.L.; Pacioni, C.; Spencer, P.B.S.; Craig, M.D. 2015. Contribution of genetics to ecological restoration. Molecular Ecology. 24(1): 22–37. doi:10.1111/mec.12995. [Date accessed: December 15, 2017].

Nelson, C.D; Josserand, S; Echt, C.S.; Koppelman, J. 2007. Loblolly pine SSR markers for shortleaf pine genetics. In: Kabrick, J.M.; Dey, D.C.; Gwaze, D., eds. Shortleaf pine restoration and ecology in the Ozarks: proceedings of a symposium; 2006 November 7–9; Springfield; MO. Gen. Tech. Rep. NRS-P-15. Newtown Square; PA: U.S. Department of Agriculture, Forest Service, Northern Research Station: 95–98. http://www.treesearch.fs.fed.us/pubs/12800. [Date accessed: December 13, 2017].

Nybom, H.; Weising, K.; Rotter, B. 2014. DNA fingerprinting in botany: past, present, future. Investigative Genetics. 5: 1. doi:10.1186/2041-2223-5-1. [Date accessed: December 15, 2017].

Oswalt, S.N.; Smith, W.B.; Miles, P.D.; Pugh, S.A. 2014. Forest resources of the United States, 2012: a technical document supporting the Forest Service 2010 update of the RPA Assessment. Gen. Tech. Rep. WO-91. Washington, DC: U.S. Department of Agriculture, Forest Service, Washington Office. 218 p. https://www.srs.fs.usda.gov/pubs/gtr/gtr_wo091.pdf. [Date accessed: December 19, 2017].

Peakall, R.; Smouse, P.E. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. Bioinformatics. 28: 2537–2539.

Schuelke, M. 2000.An economic method for the fluorescent labeling of PCR fragments. Nature Biotechnology. 18: 233–234.

Shortleaf Pine Initiative [website]. 2017. http://shortleafpine.net/. [Date accessed: December 15, 2017].

Stewart, J.F.; Liu, Y.; Tauer, C.G.; Nelson, C.D. 2010. Microsatellite versus AFLP analyses of pre-management introgression levels in loblolly pine (*Pinus taeda* L.) and shortleaf pine (*P. echinata* Mill.). Tree Genetics & Genomes. 6: 853–862.

Westbrook, J.W.; Chhatre, V.E.; Wu, L.-S. [and others]. 2015. A consensus genetic map for *Pinus taeda* and *Pinus elliottii* and extent of linkage disequilibrium in two genotype-phenotype discovery populations of *P. taeda*. G3: Genes | Genomes | Genetics. 5: 1685–1694. http://treesearch.fs.fed.us/pubs/48896. [Date accessed: December 13, 2017].

Appendix

Table A.1—Pine samples, sources, and origins

Sample ID	Source institution[a]	Tree type	State of provenance[b]	Species
MM50	R8 NFSO	clone	Virginia	*P. echinata*
MM51	R8 NFSO	clone	Virginia	*P. echinata*
MM52	R8 NFSO	clone	Virginia	*P. echinata*
MM53	R8 NFSO	clone	Virginia	*P. echinata*
MM54	R8 NFSO	clone	Virginia	*P. echinata*
MM55	R8 NFSO	clone	Virginia	*P. echinata*
MM56	R8 NFSO	clone	Virginia	*P. echinata*
MM57	R8 NFSO	clone	Virginia	*P. echinata*
MQ09	R8 NFSO	clone	Tennessee	*P. echinata*
MQ10	R8 NFSO	clone	Tennessee	*P. echinata*
MQ11	R8 NFSO	clone	Tennessee	*P. echinata*
MQ12	R8 NFSO	clone	Tennessee	*P. echinata*
MQ13	R8 NFSO	clone	Tennessee	*P. echinata*
MQ14	R8 NFSO	clone	Tennessee	*P. echinata*
MQ15	R8 NFSO	clone	Tennessee	*P. echinata*
MQ16	R8 NFSO	clone	Tennessee	*P. echinata*
MT77	R8 NFSO	clone	Arkansas	*P. echinata*
MT78	R8 NFSO	clone	Arkansas	*P. echinata*
MT80	R8 NFSO	clone	Arkansas	*P. echinata*
MT81	R8 NFSO	clone	Arkansas	*P. echinata*
MT82	R8 NFSO	clone	Arkansas	*P. echinata*
MT83	R8 NFSO	clone	Arkansas	*P. echinata*
MT84	R8 NFSO	clone	Arkansas	*P. echinata*
MV13	R8 NFSO	clone	Arkansas	*P. echinata*
MV14	R8 NFSO	clone	Tennessee	*P. echinata*
MV15	R8 NFSO	clone	Tennessee	*P. echinata*
MV16	R8 NFSO	clone	Tennessee	*P. echinata*
MV17	R8 NFSO	clone	Tennessee	*P. echinata*
MV18	R8 NFSO	clone	Tennessee	*P. echinata*
MV19	R8 NFSO	clone	Tennessee	*P. echinata*
MV20	R8 NFSO	clone	Tennessee	*P. echinata*
MV27	R8 NFSO	clone	Mississippi	*P. echinata*
MV29	R8 NFSO	clone	Mississippi	*P. echinata*
MV30	R8 NFSO	clone	Mississippi	*P. echinata*
MV31	R8 NFSO	clone	Mississippi	*P. echinata*
MV32	R8 NFSO	clone	Mississippi	*P. echinata*
MV33	R8 NFSO	clone	Mississippi	*P. echinata*
MV34	R8 NFSO	clone	Mississippi	*P. echinata*
MZ11	R8 NFSO	clone	Arkansas	*P. echinata*
FBRC_3031	UF CFGRP	clone	Georgia	*P. elliottii*
FBRC_3032	UF CFGRP	clone	Florida	*P. elliottii*
FBRC_3033	WGTIP	clone	Texas	*P. elliottii*
FBRC_3034	WGTIP	clone	Louisiana	*P. elliottii*
FBRC_3035	WGTIP	clone	Louisiana	*P. elliottii*

continued

Table A.1 (continued)—Pine samples, sources, and origins

Sample ID	Source institution[a]	Tree type	State of provenance[b]	Species
FBRC_3036	WGTIP	clone	Louisiana	*P. elliottii*
FBRC_3037	WGTIP	clone	Texas	*P. elliottii*
FBRC_3038	UF CFGRP	clone	Florida	*P. elliottii*
FBRC_3039	WGTIP	clone	Louisiana	*P. elliottii*
FBRC_3040	WGTIP	clone	Louisiana	*P. elliottii*
FBRC_3041	WGTIP	clone	Louisiana	*P. elliottii*
FBRC_3042	UF CFGRP	clone	Georgia	*P. elliottii*
FBRC_3043	UF CFGRP	clone	Florida	*P. elliottii*
FBRC_3044	WGTIP	clone	Louisiana	*P. elliottii*
FBRC_3045	UF CFGRP	clone	Florida	*P. elliottii*
FBRC_3046	WGTIP	clone	Louisiana	*P. elliottii*
FBRC_3047	WGTIP	clone	Louisiana	*P. elliottii*
FBRC_3048	UF CFGRP	clone	Georgia	*P. elliottii*
FBRC_3049	UF CFGRP	clone	Georgia	*P. elliottii*
FBRC_3050	WGTIP	clone	Texas	*P. elliottii*
FBRC_3051	UF CFGRP	clone	Florida	*P. elliottii*
FBRC_3052	UF CFGRP	clone	Georgia	*P. elliottii*
FBRC_3053	WGTIP	clone	Texas	*P. elliottii*
FBRC_3054	WGTIP	clone	Texas	*P. elliottii*
FBRC_3055	UF CFGRP	clone	Florida	*P. elliottii*
FBRC_3056	WGTIP	clone	Louisiana	*P. elliottii*
FBRC_3057	WGTIP	clone	Louisiana	*P. elliottii*
FBRC_3058	UF CFGRP	clone	Georgia	*P. elliottii*
FBRC_3059	WGTIP	clone	Louisiana	*P. elliottii*
FBRC_3060	UF CFGRP	clone	Florida	*P. elliottii*
HU93	R8 NFSO	clone	North Carolina	*P. palustris*
HU94	R8 NFSO	clone	North Carolina	*P. palustris*
HV29	R8 NFSO	clone	North Carolina	*P. palustris*
HV30	R8 NFSO	clone	North Carolina	*P. palustris*
IQ70	R8 NFSO	clone	Louisiana	*P. palustris*
IQ71	R8 NFSO	clone	Louisiana	*P. palustris*
IQ99	R8 NFSO	clone	Florida	*P. palustris*
IR01	R8 NFSO	clone	Florida	*P. palustris*
IR25	R8 NFSO	clone	Mississippi	*P. palustris*
IR26	R8 NFSO	clone	Mississippi	*P. palustris*
IR52	R8 NFSO	clone	Alabama	*P. palustris*
IR53	R8 NFSO	clone	Alabama	*P. palustris*
IR69	R8 NFSO	clone	Alabama	*P. palustris*
IR70	R8 NFSO	clone	Alabama	*P. palustris*
IS53	R8 NFSO	clone	Texas	*P. palustris*
IS54	R8 NFSO	clone	Texas	*P. palustris*
IU91	R8 NFSO	clone	Georgia	*P. palustris*
IU92	R8 NFSO	clone	Georgia	*P. palustris*
LL910	HEF	native	Mississippi	*P. palustris*
LL911	HEF	native	Mississippi	*P. palustris*
VDF _010	IFCO	seedling	Alabama	*P. palustris*
VDF_001	IFCO	seedling	Georgia	*P. palustris*
VDF_002	IFCO	seedling	Georgia	*P. palustris*

continued

Table A.1 (continued)—Pine samples, sources, and origins

Sample ID	Source institution[a]	Tree type	State of provenance[b]	Species
VDF_011	IFCO	seedling	Alabama	*P. palustris*
VDF_019	IFCO	seedling	North Carolina	*P. palustris*
VDF_020	IFCO	seedling	North Carolina	*P. palustris*
VDF_028	IFCO	seedling	Florida	*P. palustris*
VDF_030	IFCO	seedling	Florida	*P. palustris*
VDF_037	IFCO	seedling	Mississippi	*P. palustris*
VDF_038	IFCO	seedling	Mississippi	*P. palustris*
VDF_046	IFCO	seedling	South Carolina	*P. palustris*
VDF_047	IFCO	seedling	South Carolina	*P. palustris*
VDF_20-803	VDF	native	Virginia	*P. palustris*
VDF_20-806	VDF	native	Virginia	*P. palustris*
029NCS	HEF	clone	Georgia	*P. taeda*
11-1060	HEF	clone	North Carolina	*P. taeda*
146NCS	HEF	clone	Virginia	*P. taeda*
158NCS	HEF	clone	Mississippi	*P. taeda*
20-1010	HEF	clone	South Carolina	*P. taeda*
26WGF	HEF	clone	Texas	*P. taeda*
2UFL	HEF	clone	Florida	*P. taeda*
313NCS	HEF	clone	Alabama	*P. taeda*
433NCS	HEF	clone	North Carolina	*P. taeda*
4sARL	HEF	clone	Alabama	*P. taeda*
6-1031	HEF	clone	North Carolina	*P. taeda*
7-56	HEF	clone	North Carolina	*P. taeda*
8-1070	HEF	clone	North Carolina	*P. taeda*
Arb-1	ArborGen	clone	South Carolina	*P. taeda*
Arb-10	ArborGen	clone	SCx(NCxTX)	*P. taeda*
Arb-11	ArborGen	clone	SCx(NCxTX)	*P. taeda*
Arb-12	ArborGen	clone	SCx(NCxTX)	*P. taeda*
Arb-2	ArborGen	clone	South Carolina	*P. taeda*
Arb-3	ArborGen	clone	South Carolina	*P. taeda*
Arb-4	ArborGen	clone	South Carolina	*P. taeda*
Arb-5	ArborGen	clone	South Carolina	*P. taeda*
Arb-6	ArborGen	clone	South Carolina	*P. taeda*
Arb-7	ArborGen	clone	South Carolina	*P. taeda*
Arb-8	ArborGen	clone	South Carolina	*P. taeda*
Arb-9	ArborGen	clone	SCxGA	*P. taeda*
B-145-L	HEF	clone	Texas	*P. taeda*
LP5	HEF	clone	Texas	*P. taeda*

[a] Institutions: R8 NFSO – Region 8 (Southern Region) National Forest Seed Orchards, with DNA samples provided by the National Forest Genetics Laboratory, Placerville, CA; UF CFGRP – University of Florida Cooperative Forest Genetics Research Program; WGTIP – Western Gulf Tree Improvement Program; IFCO – International Forest Company, Moultrie, GA, seeds planted by VDF and samples collected by Chris Maier (Southern Research Station); VDF – Virginia Department of Forestry; HEF – Harrison Experimental Forest, Saucier, MS.

[b] Provenances: SCxGA – a clone from South Carolina crossed with clone from Georgia; SCx(NCxTX) – a wide cross with clones from South Carolina, North Carolina, and Texas.