Development of EST-SSR Markers for Taxillus nigrans (Loranthaceae) in Southwestern China Using Next-Generation Sequencing

Authors: Miao, Ning, Zhang, Lei, Li, Maoping, Fan, Liqiang, and Mao, Kangshan

Source: Applications in Plant Sciences, 5(8)

Published By: Botanical Society of America

URL: https://doi.org/10.3732/apps.1700010

Applications in Plant Sciences 2017 5(8): 1700010

PRIMER NOTE

DEVELOPMENT OF EST-SSR MARKERS FOR *TAXILLUS NIGRANS* (LORANTHACEAE) IN SOUTHWESTERN CHINA USING NEXT-GENERATION SEQUENCING[1]

NING MIAO[2], LEI ZHANG[2], MAOPING LI[2], LIQIANG FAN[2], AND KANGSHAN MAO[2,3]

[2]MOE Key Laboratory of Bio-Resources and Eco-Environment, College of Life Science, Sichuan University, Chengdu 610064,
People's Republic of China

- *Premise of the study:* We developed transcriptome microsatellite markers (simple sequence repeats) for *Taxillus nigrans* (Loranthaceae) to survey the genetic diversity and population structure of this species.
- *Methods and Results:* We used Illumina HiSeq data to reconstruct the transcriptome of *T. nigrans* by de novo assembly and used the transcriptome to develop a set of simple sequence repeat markers. Overall, 40 primer pairs were designed and tested; 19 of them amplified successfully and demonstrated polymorphisms. Two loci that detected null alleles were eliminated, and the remaining 17, which were subjected to further analyses, yielded two to 21 alleles per locus.
- *Conclusions:* The markers will serve as a basis for studies to assess the extent and pattern of distribution of genetic variation in *T. nigrans*, and they may also be useful in conservation genetic, ecological, and evolutionary studies of the genus *Taxillus*, a group of plant species of importance in Chinese traditional medicine.

Key words: Chinese traditional medicine; conservation; Loranthaceae; microsatellite marker; next-generation sequencing; *Taxillus nigrans*; transcriptome.

Taxillus nigrans (Hance) Danser (Loranthaceae) is a mistletoe species that is found attached to many canopy tree species in low mountains, hills, and river basins in subtropical areas of southwestern China at elevations of 300–1300 m. Flowering can occur throughout the year, and the fruiting period is mainly in November. The entire plant of this species can be used as raw material for Chinese traditional medicine (Jiang, 1998). However, because the range of the species has undergone rapid expansion mediated by birds in the urban area of Chengdu (Sichuan Province, China), it forms large groves on garden tree species and is sometimes harmful to its host trees, so that individuals of this species are often removed by gardeners. To date, apart from some basic taxonomic data on the species (Gong et al., 2004) and genome studies on other species of *Taxillus* Tiegh. (Rist et al., 2011; Wei et al., 2017), nearly all published research has focused on aspects relating to its medicinal value, for example, the extraction and identification of medicinal components and the optimization of extraction methods (Li et al., 2006, 2009; Zhang et al., 2016; Zhao et al., 2016). There is little information on the genetic diversity and population structure of the species. We are also interested in developing genetic approaches for identification of individuals and assignment testing, which will help in understanding how this species expands its distribution and jumps from host to host in urban areas as well as in the field.

Simple sequence repeat (SSR) markers, also known as microsatellites or short tandem repeats, are highly polymorphic and are therefore useful as molecular markers in population genetic studies (Zhang et al., 2012; Jiang et al., 2015). Transcriptome sequencing has proven to be a powerful and cost-effective tool that has greatly accelerated the process of discovering molecular markers, including single nucleotide polymorphisms (SNPs) and SSRs (Ashrafi et al., 2012; Qi et al., 2016). In this study, we sequenced and assembled the transcriptome of *T. nigrans* and developed a set of expressed sequence tag (EST)–SSR markers for population genetic studies of *T. nigrans*. We also tested the transferability of these markers in herbarium samples of *T. delavayi* (Tiegh.) Danser and five individuals of *Scurrula parasitica* L. (collected from the field), another Loranthaceae parasite that co-occurs with *T. nigrans*.

METHODS AND RESULTS

Approximately 10 μg (400 ng/μL) of total RNA was extracted from fresh leaf material of one individual of *T. nigrans* using TRIzol Reagent (Invitrogen, Carlsbad, California, USA). Subsequently, mRNA was isolated using magnetic oligo (dT) beads (Illumina, San Diego, California, USA); it was then fragmented into short fragments using the Ambion RNA Fragmentation Kit (Ambion, Austin, Texas, USA) according to the manufacturer's protocols. First-strand cDNA synthesis was performed using reverse transcriptase (Invitrogen) with random primers, and second-strand cDNA was synthesized by RNase H and DNA Polymerase I (Invitrogen). Finally, the transcriptome was sequenced on an Illumina HiSeq 2000 system at Novogene (Beijing, China). Prior to the assembly, a stringent filtering process of raw sequencing reads was conducted. The number of

---

[1]Manuscript received 11 February 2017; revision accepted 19 June 2017.

The authors thank the National Key Research and Development Program of China (2016YFD0600203), the National Natural Science Foundation of China (31622015, 31200477), and the Department of Science and Technology of Sichuan Province (2013JY0026, 2015JQ0018).

[3]Author for correspondence: maokangshan@163.com

doi:10.3732/apps.1700010

TABLE 1. Characteristics of 19 polymorphic microsatellite loci developed for *Taxillus nigrans*.

Locus	Primer sequences (5′–3′)	Repeat motif	Allele size (bp)	T_a (°C)[a]	Fluorescent dye	GenBank accession no.	r	Protein[b]	Organism[c]	E-value[d]
#TR7149	F: GGCAAAATCAACCGAGAAGA R: CTTATCGCATTCACACACGC	$(CT)_{12}$	164	60 60	6-FAM	KY412965	0.0156	NEN1-like	*Populus euphratica*	3×10^{-7}
#TR11564	F: CCTCAGGAACCTGCAAGAAG R: CGACACAGGACAGCTGTGAA	$(AAG)_{16}$	215	60 60	HEX	KY412966	0.0626	WD repeat-containing protein RUP2	*Elaeis guineensis*	8×10^{-8}
TR24412	F: TTTCTTCACCAGCCGAAGTC R: AAGCGAACTCGAATCACTGC	$(CT)_{21}$	122	60 60	6-FAM	KY412967	0.0747	Predicted gene, 39330	*Oryza sativa*	4×10^{-4}
TR47466	F: AGTCCTTCGTTCCCGATACC R: TCTGATGGGCTTACTCCGTC	$(AT)_{24}$	231	60 60	TAMRA	KY421968	0.3990	Unknown	*Vigna angularis*	2×10^{-21}
#TR51334	F: GTAAGATCCCAAACCGAGGC R: CTGCACTCTTCCATACGGCT	$(AG)_{26}$	206	60 60	TAMRA	KY421969	0.0008	Transmembrane protein, putative	*Medicago truncatula*	1×10^{-21}
TR56117	F: TCTTTCCATTCCAGCGACTC R: CTCGATTCTACTCCGCGGTT	$(TC)_{15}$	166	60 61	TAMRA	KY421970	0.1183	LOC107411880	*Ziziphus jujuba*	2×10^{-8}
TR59209	F: TGTGCGTTTGTTTGTTCGTT R: AGGAATCGAACAGGAGGGTC	$(TC)_{15}$	157	60 60	6-FAM	KY421971	0.1826	LOC107268204	*Cephus cinctus*	2×10^{-8}
#TR83979	F: CCTCCGTCTCTCCCTCTCTC R: TCGTCCTCTTCCACTATGCC	$(CT)_{22}$	245	60 60	HEX	KY421972	0.0748	At3g02290	*Oryza sativa*	2×10^{-15}
TR85804	F: TCCTCTCTCTCCGGCCTTAT R: CCTTGCTAATTCCACCACCA	$(AG)_{33}$	219	60 60	HEX	KY421973	0.0705	CARUB_v10002273mg	*Capsella rubella*	1×10^{-26}
TR87965	F: TGGAGATCTTGGCTTCGTTC R: TAACTCGCTTTGCCACCTTC	$(AG)_{14}$	216	60 60	6-FAM	KY421974	0.1080	DDB1- and CUL4-associated factor 13	*Theobroma cacao*	2×10^{-12}
#TR88317	F: GAGGGGAGGGTGCTTGTAAT R: TTGCAGGAACAGGTATGGCT	$(TAT)_{15}$	129	60 60	6-FAM	KY421975	0.2512	Restricted Tev Movement 1-like	*Nicotiana tomentosiformis*	1×10^{-15}
TR90181	F: AATGACCGTATCCTGAACGC R: AAGCGACCTCATCCAACATC	$(TA)_{25}$	217	59 59	HEX	KY421976	−0.0708	LOC107270001	*Cephus cinctus*	1×10^{-22}
#TR91417	F: AGAGGAATTGGCATCGTCAG R: TCCAACTCACACTTGCCTCA	$(GA)_{26}$	213	60 60	6-FAM	KY421977	0.1064	LOC105638199	*Jatropha curcas*	2×10^{-21}
#TR97121	F: CTGGAGTCGTAGTAGCCCGA R: TCCTCCTCACTCATTGCCTC	$(AG)_{15}$	204	60 60	HEX	KY421978	0.0685	Transcription factor bHLH35	*Vigna angularis*	3×10^{-13}
TR98683	F: TGGCTACCCTCCTATCTCCC R: AGACTCGAAGGCCTCTGGTT	$(CT)_{15}$	255	60 60	HEX	KY421979	0.0839	LOC104597466	*Nelumbo nucifera*	2×10^{-9}
TR105177	F: CAGCATGCATTGCTAGGAGA R: TGGGAAATGGACGTTGTTCT	$(GA)_{29}$	217	60 60	TAMRA	KY421980	0.0798	LOC103319601	*Prunus mume*	1×10^{-25}
TR120023	F: CTTGATCTTCTGGTGCGGTT R: CCGTCACTGCTCTCCTTCAT	$(GA)_{14}$	161	60 60	TAMRA	KY421981	0.1191	LOC104727032	*Camelina sativa*	4×10^{-9}
*#TR85478	F: GTCGTCATGGACTCTTCGCT R: ACTGGGACACATTCCTGCAT	$(TC)_5$	228	60 60	TAMRA	KY421982	ND	EUTSA_v10007584mg	*Eutrema salsugineum*	5×10^{-4}
*TR87192	F: CCTTTGGAGGGGTTCAACTT R: TGGCCGAAGTTTTAGTCAGC	$(GCG)_4$	271	60 60	HEX	KY421983	0.4202	LOC103986576	*Musa acuminata*	0.11

Note: ND = not done; r = null allele frequency; T_a = annealing temperature,

[a] The annealing temperature for each primer is listed, and the final annealing temperature for each PCR reaction is given as the average annealing temperature of the adopted primer pair.

[b] Information from BLAST analysis on the protein most closely matching the EST.

[c] Organism from which the BLAST match was obtained.

[d] E-value associated with the BLAST match.

* Null alleles ($r > 0.4$).

\# Primers successfully amplified for *Taxillus delavayi*.

Table 2. Genetic properties of 17 newly developed polymorphic microsatellite loci in three populations of *Taxillus nigrans*. Loci exhibiting null alleles are not included.[a]

	Sichuan University ($n = 100$)			Tazishan ($n = 30$)			Huanhuaxi ($n = 30$)		
Locus	A	H_o	H_e	A	H_o	H_e	A	H_o	H_e
TR7149	7	0.717	0.815	5	0.900	0.728	10	0.967	0.844
TR11564	5	0.667	0.781	4	0.767	0.672	5	0.667	0.727
TR24412	6	0.551	0.628	4	0.633	0.691	7	0.621	0.722
TR47466	6	0.333	0.453	2	0.034	0.034	4	0.367	0.476
TR51334	11	0.525	0.776	2	0.966	0.499	5	0.967	0.577
TR56117	9	0.583	0.745	7	0.724	0.737	8	0.567	0.787
TR59209	10	0.626	0.789	6	0.310	0.596	6	0.643	0.786
TR83979	17	0.737	0.859	8	0.931	0.829	10	0.828	0.757
TR85804	18	0.808	0.876	9	0.633	0.799	17	0.833	0.898
TR87965	7	0.646	0.786	5	0.586	0.703	6	0.667	0.764
TR88317	11	0.347	0.714	5	0.607	0.702	4	0.517	0.644
TR90181	14	1.000	0.786	7	0.963	0.747	5	1.000	0.621
TR91417	10	0.717	0.809	6	0.400	0.665	7	0.700	0.749
TR97121	2	0.380	0.476	2	0.500	0.408	2	0.400	0.464
TR98683	14	0.690	0.860	10	0.833	0.815	15	0.933	0.813
TR105177	20	0.764	0.893	8	0.733	0.807	10	0.931	0.835
TR120023	21	0.802	0.885	6	0.633	0.776	6	0.552	0.797

Note: A = number of alleles sampled; H_e = expected heterozygosity; H_o = observed heterozygosity; n = number of individuals sampled.

[a] Voucher and locality information are provided in Appendix 1.

low-quality ($Q \le 3$) bases in a single read was restricted to less than 50%, and paired reads were discarded if the number of unknown nucleotide bases in either of the paired reads exceeded 3% following the sequencing company's protocol (Novogene). After removing the adapter sequences and ambiguous reads, the clean reads obtained were de novo assembled using Trinity (release 2013-02-25; Grabherr et al., 2011) with default settings. The final assembly was composed of 299,147 unigenes and had an N50 size of 1056 bp. Raw transcriptome read data were deposited in the National Center for Biotechnology Information (NCBI) Short Read Archive (accession no. SRP105083).

SSRs were detected using the Perl script MISA (Thiel et al., 2003) with a motif size of one to six nucleotides and thresholds of eight, four, four, three, three, and three repeat units for mono-, di-, tri-, tetra-, penta-, and hexanucleotide SSRs, respectively. We selected 83,954 microsatellite loci and used the primer design software package Primer3 version 2.3.6 (Untergasser et al., 2012) to design primer sets. Following random browsing across the output files of these primer sets, 40 markers were selected based on length (19–20 bp), GC content (40–65%) of the primers, and annealing temperatures (59–61°C) of the primer sets. Nineteen of the 40 tested markers were selected based on PCR success rate and degree of polymorphism (difference in band length), and these were used to genotype individual *Taxillus* plants (Table 1).

Table 3. Fragment sizes detected in cross-amplification tests of the 19 newly developed microsatellite markers in *Taxillus delavayi* and *Scurrula parasitica*.[a]

Locus	*Taxillus delavayi* ($n = 2$)	*Scurrula parasitica* ($n = 5$)
TR7149	167	152–163
TR11564	192	193
TR24412	—	124
TR47466	—	272
TR51334	182	174–182
TR56117	—	155–185
TR59209	—	125–143
TR83979	244	177–211
TR85804	—	179–255
TR87965	—	197
TR88317	100	100–130
TR90181	—	205–207
TR91417	196	196–204
TR97121	332	353
TR98683	—	244–260
TR105177	—	189
TR120023	—	152
TR85478	229	229
TR87192	—	269

Note: — = amplification failed; n = number of individuals sampled.

[a] Voucher and locality information are provided in Appendix 1.

Genomic DNA was extracted from the silica-dried leaves of 160 individuals from three populations of *T. nigrans*, two individuals of *T. delavayi*, and five individuals of *S. parasitica* (Appendix 1) using the cetyltrimethylammonium bromide (CTAB) method of Doyle and Doyle (1987). PCR reactions were performed in 25-µL volumes containing 12.5 µL 2× PCR buffer, 300.0 µM each dNTP, 0.3 µM each primer, 1.25 unit *Taq* DNA polymerase (Vazyme Biotech, Nanjing, China), and ca. 50 ng of genomic DNA. The cycling conditions were as follows: 95°C for 5 min, followed by 35 cycles of 95°C for 45 s, 55°C for 30 s, and 72°C for 45 s; the reactions were completed by a final elongation step at 72°C for 10 min. The PCR products were checked on 1% agarose gels to confirm PCR success and then sent to TsingKe (Chengdu, China) for microsatellite genotyping. Primer pairs were synthesized with the forward primer of each pair 5′ end-labeled with either 6-FAM, TAMRA, or HEX (Applied Biosystems, Foster City, California, USA), and amplicons were analyzed on an ABI PRISM 3100 genetic analyzer. The microsatellite genotype at each locus for each individual was determined using GeneMarker (SoftGenetics, State College, Pennsylvania, USA). Allele sizes at each locus were then scored and checked for possible genotyping errors, such as stuttering, large allele dropouts, or null alleles, using CERVUS (Dakin and Avise, 2004). In total, null alleles (null allele frequency $[r] > 0.4$) were detected at two loci (Table 1). These loci were eliminated, and the remaining 17 microsatellite loci were subjected to further analyses (Table 2).

These 17 microsatellite loci were highly polymorphic, with two to 21 alleles per locus. We used GenAlEx version 6 (Peakall and Smouse, 2006) to calculate the number of alleles and the observed and expected heterozygosity at each locus (Table 2). When using GIMLET version 1.3.3 (Valière, 2002), a minimum of two loci and six loci are needed to estimate, respectively, the unbiased probability that a genotype is shared by two individuals (P_{ID}) in a population, and the probability that a genotype is shared by two siblings ($P_{ID(sib)}$).

In the cross-species transferability test, eight of the 19 loci were successfully genotyped in two individuals of *T. delavayi* taken from herbarium specimens (Table 3). In contrast, all polymorphic loci were successfully amplified in *S. parasitica* (Table 3). The difference in success between *T. delavayi* and *S. parasitica* may have been due to a higher proportion of degraded DNA from *T. delavayi* herbarium specimens.

CONCLUSIONS

We developed and amplified a set of polymorphic EST-SSR markers for *T. nigrans.* These new SSR markers will serve as a basis for studies assessing the genetic diversity and population structure of *T. nigrans*. Our research will be useful for conservation genetic, ecological, and evolutionary studies of the genus *Taxillus*, a group of plant species of importance in Chinese traditional medicine. We plan to use these markers to explain the rapid demographic expansion and host specificity of *T. nigrans* in urban areas in southwestern China.

LITERATURE CITED

Ashrafi, H., T. Hill, K. Stoffel, A. Kozik, J. Q. Yao, S. R. Chin-Wo, and A. Van Deynze. 2012. *De novo* assembly of the pepper transcriptome (*Capsicum annuum*): A benchmark for in silico discovery of SNPs, SSRs and candidate genes. *BMC Genomics* 13: 571.

Dakin, E. E., and J. C. Avise. 2004. Microsatellite null alleles in parentage analysis. *Heredity* 93: 504–509.

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

Gong, Z. N., Y. F. Wang, Q. L. Liang, Z. T. Wang, L. S. Xu, and G. J. Xu. 2004. A chemotaxonomic study of 27 species of the Loranthaceae plant from China. *Guangxi Zhi Wu* 24: 493–496, 514.

Grabherr, M. G., B. J. Haas, M. Yassour, J. Z. Levin, D. A. Thompson, I. Amit, X. Adiconis, et al. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* 29: 644–652.

Jiang, D. C., G. L. Wu, K. S. Mao, and J. J. Feng. 2015. Structure of genetic diversity in marginal populations of black poplar (*Populus nigra* L.). *Biochemical Systematics and Ecology* 61: 297–302.

Jiang, M. L. 1998. Flora Reipublicae Popularis Sinicae, Vol. 24. Science Press, Beijing, China (in Chinese).

Li, Y. H., D. Lu, M. H. Zhao, and K. X. Zhu. 2006. Research on the developments and applications for medicinal plants of Loranthaceae in Guangxi. *Guangxi Medical Journal* 28: 1695–1698.

Li, Y. H., J. L. Ruan, S. L. Chen, D. Lu, K. X. Zhu, M. H. Zhao, and H. H. Pei. 2009. Study on medicinal plants of Loranthaceae resources in China. *World Science and Technology (Modernization of Traditional Chinese Medicine and Materia Medica)* 11: 665–669.

Peakall, R., and P. E. Smouse. 2006. GenAlEx 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* 6: 288–295.

Qi, W. C., F. Lin, Y. H. Liu, B. Q. Huang, J. H. Cheng, W. Zhang, and H. Zhao. 2016. High-throughput development of simple sequence repeat markers for genetic diversity research in *Crambe abyssinica. BMC Plant Biology* 16: 139.

Rist, L., R. Uma Shaanker, and J. Ghazoul. 2011. The spatial distribution of mistletoe in a southern Indian tropical forest at multiple scales. *Biotropica* 43: 50–57.

Thiel, T., W. Michalek, R. K. Varshney, and A. Graner. 2003. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theoretical and Applied Genetics* 106: 411–422.

Untergasser, A., I. Cutcutache, T. Koressaar, J. Ye, B. C. Faircloth, M. Remm, and S. G. Rozen. 2012. Primer3–New capabilities and interfaces. *Nucleic Acids Research* 40: e115.

Valière, N. 2002. GIMLET: A computer program for analysing genetic individual identification data. *Molecular Ecology Notes* 2: 377–379.

Wei, S., X. Ma, L. Pan, J. Miao, J. Fu, L. Bai, Z. Zhong, et al. 2017. Transcriptome analysis of *Taxillus chinensis* (DC.) Danser seeds in response to water loss. *PLoS ONE* 12: e0169177.

Zhang, H., F. Y. Huang, B. W. Su, K. X. Zhu, H. L. Lu, S. G. Yin, M. Guo, and Y. H. Li. 2016. Impacts of different host trees on the quality of *Taxillus chinensis. World Science and Technology (Modernization of Traditional Chinese Medicine and Materia Medica)* 18: 1182–1187.

Zhang, Q., J. Li, Y. B. Zhao, S. S. Korban, and Y. P. Han. 2012. Evaluation of genetic diversity in Chinese wild apple species along with apple cultivars using SSR markers. *Plant Molecular Biology Reporter* 30: 539–546.

Zhao, H. W., J. Wang, Y. Cui, H. J. Wang, J. M. Wang, and K. Ma. 2016. A research of effective chemicals of *Taxillus chinensis* Danser on osteoporosis and channel tropism based on disease-effect-bioanalysis. *World Science and Technology (Modernization of Traditional Chinese Medicine and Materia Medica)* 18: 626–631.

Appendix 1. Voucher specimen information for Loranthaceae used in this study.

Species	*N*	Population code	Locality	Geographic coordinates	Voucher specimen accession no.[a]
Taxillus nigrans (Hance) Danser	100	SCU	Sichuan University, Sichuan	30°37′48″N, 104°4′48″E	SZ-00545040, SZ-00545041, SZ-00545042, SZ-00545043, SZ-00545044
T. nigrans	30	TZT	Tazishan, Sichuan	30°38′7″N, 104°7′15″E	SZ-00545045, SZ-00545046, SZ-00545047
T. nigrans	30	HH	Huanhuaxi, Sichuan	30°39′28″N, 104°1′55″E	SZ-00545048, SZ-00545049, SZ-00545050
T. delavayi (Tiegh.) Danser	1	Individual	Maerkang, Sichuan	31°54′46″N, 102°11′24″E	SZ-00280020
T. delavayi	1	Individual	Muli, Sichuan	27°55′55″N, 101°16′43″E	SZ-00280006
Scurrula parasitica L.	5	TZS	Tazishan, Sichuan	30°38′7″N, 104°7′15″E	SZ-00545051

Note: *N* = number of individuals sampled.

[a] All voucher specimens are deposited at the herbarium of Sichuan University (SZ), Sichuan, China.