Polymorphic Microsatellite Markers for a Wind-Dispersed Tropical Tree Species, Triplaris cumingiana (Polygonaceae)

Authors: Na Wei, and Christopher W. Dick
Source: Applications in Plant Sciences, 2(9)
Published By: Botanical Society of America
URL: https://doi.org/10.3732/apps.1400051
POLYMORPHIC MICROSATELLITE MARKERS FOR A WIND-DISPERSED TROPICAL TREE SPECIES, *Triplaris cumingiana* (Polygonaceae)\(^1\)

**NA WEI**\(^2,4\) AND CHRISTOPHER W. DICK\(^2,3\)

\(^2\)Department of Ecology and Evolutionary Biology, University of Michigan, 830 North University Avenue, Ann Arbor, Michigan 48109-1048 USA; and \(^3\)Smithsonian Tropical Research Institute, Apartado Postal 0843-03092, Balboa, Ancón, Republic of Panama

- **Premise of the study:** Novel microsatellite markers were characterized in the wind-dispersed and dioecious neotropical tree *Triplaris cumingiana* (Polygonaceae) for use in understanding the ecological processes and genetic impacts of pollen- and seed-mediated gene flow in tropical forests.
- **Methods and Results:** Sixty-two microsatellite primer pairs were screened, from which 12 markers showing five or more alleles per locus (range 5–17) were tested on 47 individuals. Observed and expected heterozygosities averaged 0.692 and 0.731, respectively. Polymorphism information content was between 0.417 and 0.874. Linkage disequilibrium was observed in one of the 66 pairwise comparisons between loci. Two loci showed deviation from Hardy–Weinberg equilibrium. An additional 14 markers exhibiting lower polymorphism were characterized on a smaller number of individuals.
- **Conclusions:** These microsatellite markers have high levels of polymorphism and reproducibility and will be useful in studying gene flow and population structure in *T. cumingiana*.

**Key words:** gene flow; microsatellite marker; PacBio sequencing platform; single-molecule real-time sequencing; *Triplaris cumingiana*; wind dispersal.

*Triplaris cumingiana* Fisch. & C. A. Mey. ex C. A. Mey. (Polygonaceae) is a wind-dispersed, dioecious tree species found in humid forests of lower Central America and western South America (Croat, 1978). It forms an obligate mutualistic (Polygonaceae) relationship with the sting ant *Pseudomyrmex triplarinus* (Croat, 1978) as an ant-herbivore defense. Unlike most dioecious tree species that have inconspicuous unisexual flowers (Bawa and Opler, 1975), flower sexual dimorphism is pronounced in *T. cumingiana*, which produces bright red bracts signaling flowers on female trees during the dry season in Panama. The dioecious mating system, which permits unambiguous identification of the maternal and paternal contribution to seedlings, combined with the ease of sexual determination, make *T. cumingiana* of particular interest for studies of pollen- and seed-mediated gene flow in tropical tree species. To investigate the ecological and genetic impacts of pollen and seed dispersal in *T. cumingiana* as compared to tropical trees of alternative pollination and dispersal syndromes, we developed polymorphic microsatellite markers in *T. cumingiana*. We used single-molecule real-time sequencing (SMRT) implemented in the PacBio RS platform (Pacific Biosciences, Menlo Park, California, USA) because it is capable of generating long reads (Wei et al., 2014).

**METHODS AND RESULTS**

Genome shotgun sequences were obtained using PacBio’s high accuracy mode of circular consensus sequencing (Wei et al., 2014). In brief, genomic DNA from one reproductive-sized tree of *T. cumingiana* that grows in the 50-ha Forest Dynamics Plot (FDP) on Barro Colorado Island, Panama (9°10′N, 79°51′W; tag no. 199017), was used for PacBio 500-bp SMRTbell library preparation. Following sonication, DNA fragments averaging 500 bp were ligated with two 55-nucleotide hairpin adapters, and then sequenced on the PacBio RS platform using C2 chemistry. Four SMRT cells generated a total of 178,122 circular consensus reads. Quality control was performed to remove homopolymer-rich sequences and poor-quality portions of individual reads (for details, see Wei et al., 2014). The resulting high-quality sequences were used for microsatellite searching and primer design in QDD version 2.1 (Meglécz et al., 2010). In total, 795 microsatellite loci were obtained, in which 686 loci contained pure repeat motifs (524 di-, 143 tri-, 15 tetra-, three penta-, and one hexanucleotide repeat motifs). From these pure microsatellites, loci of 29 repeats with dinucleotide motifs and loci of 27 repeats with other repeat motifs were retained for marker validation. This test array was composed of 62 microsatellite loci (32 di-, 25 tri-, three tetra-, one penta-, and one hexanucleotide motifs).

To test these primers, we isolated genomic DNA using a modified cetyltrimethylammonium bromide (CTAB) method (Doyle and Doyle, 1987) from lyophilized leaves of 47 *T. cumingiana* adult trees growing in the 50-ha FDP (voucher no. Pérez 1862, Smithsonian Tropical Research Institute herbarium [STRI], Panama). After the initial check of primer amplification on three individuals, we found 39 primer pairs generated easily interpretable allelic patterns, of which 11 loci were monomorphic. We then tested the remaining 28 polymorphic markers on another nine individuals, for which two loci showing apparent null alleles were excluded from further analyses. Microsatellite loci of 255 alleles based on these 12 samples were screened on an additional 35 individuals. The 8-μL PCR reactions contained 4 ng DNA, 6.25 nM HEX-labeled or 9.40 nM FAM-labeled M13 primer (TGTAAACGACGGCCAGT), 0.075 μM M13-tagged forward primer, 0.3 μM reverse primer, 4 μM MgCl\(_2\), 4 μL GoTaq Colorless Master Mix (Promega Corporation, Madison, Wisconsin, USA) including 200 μM of each dNTP and 1 unit *Taq* DNA polymerase, and H\(_2\)O. PCRs were carried out using two different thermocycling conditions. For most of the tested primers, we used a touchdown protocol (a; Table 1): 94°C for 4 min; 28 cycles of

---

\(^1\) Manuscript received 20 June 2014; revision accepted 22 July 2014.

This work was supported by a CTPS-ForestGEO grant from the Smithsonian Tropical Research Institute and the Center for Tropical Forest Science.

\(^4\) Author for correspondence: weina@umich.edu
doi:10.3732/apps.1400051


This work is licensed under a Creative Commons Attribution License (CC-BY-NC-SA).
Applications in Plant Sciences 2014 2 (9): 1400051
Wei and Dick—Triplaris cumingiana microsatellites

94°C for 30 s, 59°C (decreasing 0.2°C per cycle) for 40 s, and 72°C for 60 s; 10 cycles of 94°C for 30 s, 53°C for 40 s, and 72°C for 60 s; and a final extension at 72°C for 10 min. When the above protocol produced weak PCR amplicons, we followed a nontouchdown protocol (b; Table 1): 94°C for 4 min; 28 cycles of 94°C for 30 s, 54.5°C for 40 s, and 72°C for 60 s; and a final extension at 72°C for 10 min. PCR amplicons were multiplexed by combining one HEX-labeled locus of 1.6 μL, and one FAM-labeled locus of 1.4 μL, with 11.5 μL Hi-Di formamide (Life Technologies) and 0.05 μL GeneScan 500 ROX Size Standard (Life Technologies), before loading to a single lane on an ABI 3730 DNA Analyzer (Life Technologies). Alleles were called using GeneMarker version 2.4.1 (SoftGenetics, State College, Pennsylvania, USA).

We examined marker characteristics including allele richness, observed and expected heterozygosity, and probability of exclusion (PE2, when one parent was known; PE3, of a parent pair) using GenAlEx version 6.5 (Peakall and Smouse, 2012). Polymorphism information content (PIC) was estimated using PowerMarker version 3.25 (Liu and Muse, 2005). Exact tests of Hardy–Weinberg equilibrium (HWE) and locus pairwise linkage disequilibrium (LD) were conducted in GENEPOP version 4.2.2 (Rousset, 2008). The $P$ values of HWE and LD tests were adjusted for multiple comparisons using Holm’s correction (Holm, 1979).

We first focused on 12 markers with an average allelic richness of nine (range 5–17) (Table 2). Observed heterozygosity ranged from 0.426 to 0.936 (mean = 0.692), and expected heterozygosity was between 0.475 and 0.885 (mean = 0.731). Locus PIC averaged 0.705, and overall exclusion probability was 0.998 (with one parent known) and 1.000 (of a parent pair). We observed LD between the loci TRI_27 and TRI_31 (Holm’s adjusted $P < 0.007$). These 12 microsatellite markers should provide resolution for studying gene flow and genetic structure in T. cumingiana. In addition, we provide information on 14 less-polymorphic loci (2–4 alleles per locus tested on 12 individuals) (Appendix 1), which can be potential candidate markers if more genetic information is required.

**CONCLUSIONS**

We characterized novel microsatellite markers in the dioecious, insect-pollinated, wind-dispersed tropical tree *T. cumingiana* to improve understanding of the processes of pollen- and seed-mediated gene flow in tropical forests. This will be done in parallel with studies of tree species with alternative pollination and dispersal syndromes. These markers will also be useful for studying the ecological responses of *T. cumingiana* (e.g., dispersal, recruitment) to rapid changes in temperature and rainfall patterns, as the distribution of this species is associated with

**TABLE 1. Characteristics of 12 polymorphic microsatellite markers developed in Triplaris cumingiana.**

<table>
<thead>
<tr>
<th>Locus</th>
<th>Primer sequences (5’–3’)*</th>
<th>Repeat motif</th>
<th>Allele size range (bp)</th>
<th>$T_e$ (°C)</th>
<th>GenBank accession no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRI_01</td>
<td>F: GCCATATACCCAGACACAGC</td>
<td>(AAT)$_3$</td>
<td>337–418</td>
<td>b</td>
<td>KF680412</td>
</tr>
<tr>
<td>TRI_07</td>
<td>F: GCCGATGATGATCAATCTC</td>
<td>(ACAT)$_3$</td>
<td>220–364</td>
<td>b</td>
<td>KF680415</td>
</tr>
<tr>
<td>TRI_09</td>
<td>F: GAGTTCGGAGCTGAGTTTC</td>
<td>(AAAG)$_3$</td>
<td>194–242</td>
<td>a</td>
<td>KF680417</td>
</tr>
<tr>
<td>TRI_20</td>
<td>F: ATTCATGACAGCCCGACTT</td>
<td>(ACAT)$_7$</td>
<td>196–238</td>
<td>a</td>
<td>KF680421</td>
</tr>
<tr>
<td>TRI_26</td>
<td>F: TCCGCTTGCTGCTGAGTCG</td>
<td>(ACAT)$_7$</td>
<td>196–238</td>
<td>a</td>
<td>KF680426</td>
</tr>
<tr>
<td>TRI_27</td>
<td>F: TCCTGCAGAGTCCTCAAAGC</td>
<td>(AG)$_{17}$</td>
<td>154–238</td>
<td>a</td>
<td>KF680427</td>
</tr>
<tr>
<td>TRI_31</td>
<td>F: GCAAACTGATAATGGCTTACC</td>
<td>(AT)$_5$</td>
<td>200–224</td>
<td>b</td>
<td>KF680430</td>
</tr>
<tr>
<td>TRI_38</td>
<td>F: TGGGTAGCTGTCGATGTTG</td>
<td>(AT)$_{12}$</td>
<td>109–127</td>
<td>b</td>
<td>KF680432</td>
</tr>
<tr>
<td>TRI_40</td>
<td>F: TACACGCGCCCTCCTTCG</td>
<td>(AG)$_{10}$</td>
<td>232–254</td>
<td>a</td>
<td>KF680433</td>
</tr>
<tr>
<td>TRI_45</td>
<td>F: TACAGGGGACTGTCG</td>
<td>(AG)$_{26}$</td>
<td>106–122</td>
<td>a</td>
<td>KF680437</td>
</tr>
<tr>
<td>TRI_49</td>
<td>F: TGCGGCTGCTTCCTTTTC</td>
<td>(AG)$_{19}$</td>
<td>123–149</td>
<td>a</td>
<td>KF680440</td>
</tr>
<tr>
<td>TRI_55</td>
<td>F: AACTGTGACGCTGACAGG</td>
<td>(AG)$_{17}$</td>
<td>288–304</td>
<td>a</td>
<td>KF680444</td>
</tr>
</tbody>
</table>

**Note:** $T_e$ a annealing temperature.

* M13 tail (TGTAAAACGACGGCCAGT) added to the 5’ end of each forward primer.  
* A = number of alleles per locus; $H_a$ = observed heterozygosity; $H_e$ = expected heterozygosity; PE$_2$ = probability of exclusion of a parent; PE$_3$ = probability of exclusion of a parent pair; PIC = polymorphism information content.

**TABLE 2. Summary statistics of microsatellite marker polymorphism tested on 47 reproductively mature trees of Triplaris cumingiana, growing in the 50-ha Forest Dynamics Plot on Barro Colorado Island, Panama.**

<table>
<thead>
<tr>
<th>Locus</th>
<th>$A$</th>
<th>$H_a$</th>
<th>$H_e$</th>
<th>PE$_2$</th>
<th>PE$_3$</th>
<th>PIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRI_01</td>
<td>13</td>
<td>0.915</td>
<td>0.877</td>
<td>0.605</td>
<td>0.909</td>
<td>0.866</td>
</tr>
<tr>
<td>TRI_07</td>
<td>17</td>
<td>0.800</td>
<td>0.847</td>
<td>0.554</td>
<td>0.890</td>
<td>0.835</td>
</tr>
<tr>
<td>TRI_09</td>
<td>8</td>
<td>0.786</td>
<td>0.740</td>
<td>0.366</td>
<td>0.759</td>
<td>0.717</td>
</tr>
<tr>
<td>TRI_20</td>
<td>7</td>
<td>0.830</td>
<td>0.775</td>
<td>0.398</td>
<td>0.769</td>
<td>0.747</td>
</tr>
<tr>
<td>TRI_26</td>
<td>8</td>
<td>0.511*</td>
<td>0.691</td>
<td>0.303</td>
<td>0.700</td>
<td>0.664</td>
</tr>
<tr>
<td>TRI_27</td>
<td>15</td>
<td>0.936</td>
<td>0.885</td>
<td>0.626</td>
<td>0.920</td>
<td>0.874</td>
</tr>
<tr>
<td>TRI_31</td>
<td>7</td>
<td>0.652</td>
<td>0.790</td>
<td>0.419</td>
<td>0.786</td>
<td>0.762</td>
</tr>
<tr>
<td>TRI_38</td>
<td>9</td>
<td>0.511*</td>
<td>0.825</td>
<td>0.484</td>
<td>0.834</td>
<td>0.804</td>
</tr>
<tr>
<td>TRI_40</td>
<td>8</td>
<td>0.660</td>
<td>0.655</td>
<td>0.269</td>
<td>0.670</td>
<td>0.631</td>
</tr>
<tr>
<td>TRI_45</td>
<td>6</td>
<td>0.489</td>
<td>0.481</td>
<td>0.127</td>
<td>0.466</td>
<td>0.454</td>
</tr>
<tr>
<td>TRI_49</td>
<td>5</td>
<td>0.809</td>
<td>0.736</td>
<td>0.317</td>
<td>0.666</td>
<td>0.688</td>
</tr>
<tr>
<td>TRI_55</td>
<td>5</td>
<td>0.426</td>
<td>0.475</td>
<td>0.116</td>
<td>0.384</td>
<td>0.417</td>
</tr>
<tr>
<td>Mean</td>
<td>9.0</td>
<td>0.692</td>
<td>0.731</td>
<td>0.998*</td>
<td>1.000*</td>
<td>0.705</td>
</tr>
</tbody>
</table>

**Note:** $A$ = number of alleles per locus; $H_a$ = expected heterozygosity; $H_e$ = observed heterozygosity; PE$_2$ = probability of exclusion with one parent known; PE$_3$ = probability of exclusion of a parent pair; PIC = polymorphism information content.

* Significant deviation from Hardy–Weinberg equilibrium after Holm’s correction (adjusted $P < 0.007$).

* Cumulative probability of exclusion over multiallelic loci.

http://www.bioone.org/loi/apps
high soil phosphorous content and high dry-season intensity (Condit et al., 2013).

**LITERATURE CITED**


**APPENDIX 1**. Fourteen additional polymorphic microsatellite markers of *Triplaris cumingiana* screened on 12 individuals sampled from the 50-ha Forest Dynamics Plot on Barro Colorado Island, Panama. PCRs follow a touchdown protocol (see text for details).

<table>
<thead>
<tr>
<th>Locus</th>
<th>Primer sequences (5′–3′)</th>
<th>Repeat motif</th>
<th>Allele size range (bp)</th>
<th>A</th>
<th>H&lt;sub&gt;o&lt;/sub&gt;</th>
<th>H&lt;sub&gt;e&lt;/sub&gt;</th>
<th>PIC</th>
<th>GenBank accession no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRI_06</td>
<td>F: CCTTTCCAACAAAGGCTTACC</td>
<td>(ACT)&lt;sub&gt;7&lt;/sub&gt;</td>
<td>272–281</td>
<td>2</td>
<td>0.083</td>
<td>0.080</td>
<td>0.077</td>
<td>KF680414</td>
</tr>
<tr>
<td>TRI_13</td>
<td>F: TGTGATACCACAAAGGCGAGG</td>
<td>(AGG)&lt;sub&gt;15&lt;/sub&gt;</td>
<td>115–118</td>
<td>2</td>
<td>0.667</td>
<td>0.444</td>
<td>0.346</td>
<td>KF680419</td>
</tr>
<tr>
<td>TRI_28</td>
<td>F: TCACACTGTTGTCGCTC</td>
<td>(AAT)&lt;sub&gt;14&lt;/sub&gt;</td>
<td>108–114</td>
<td>2</td>
<td>0.167</td>
<td>0.444</td>
<td>0.346</td>
<td>KF680428</td>
</tr>
<tr>
<td>TRI_30</td>
<td>F: AAAGGGAGGAGAAGATGTTG</td>
<td>(AAT)&lt;sub&gt;11&lt;/sub&gt;</td>
<td>125–212</td>
<td>4</td>
<td>0.250</td>
<td>0.358</td>
<td>0.338</td>
<td>KF680429</td>
</tr>
<tr>
<td>TRI_36</td>
<td>F: CACCTGATGGTTGTCAGAAC</td>
<td>(AG)&lt;sub&gt;9&lt;/sub&gt;</td>
<td>163–177</td>
<td>2</td>
<td>0.222</td>
<td>0.444</td>
<td>0.555</td>
<td>KF680431</td>
</tr>
<tr>
<td>TRI_44</td>
<td>F: CACCTGATGGTTGTCAGAAC</td>
<td>(AT)&lt;sub&gt;28&lt;/sub&gt;</td>
<td>148–166</td>
<td>4</td>
<td>0.250</td>
<td>0.705</td>
<td>0.651</td>
<td>KF680436</td>
</tr>
<tr>
<td>TRI_51</td>
<td>F: CACCTGATGGTTGTCAGAAC</td>
<td>(AC)&lt;sub&gt;30&lt;/sub&gt;</td>
<td>147–153</td>
<td>3</td>
<td>0.455</td>
<td>0.368</td>
<td>0.425</td>
<td>KF680441</td>
</tr>
<tr>
<td>TRI_52</td>
<td>F: CACCTGATGGTTGTCAGAAC</td>
<td>(AG)&lt;sub&gt;23&lt;/sub&gt;</td>
<td>115–121</td>
<td>4</td>
<td>0.182</td>
<td>0.380</td>
<td>0.446</td>
<td>KF680442</td>
</tr>
<tr>
<td>TRI_54</td>
<td>F: CACCTGATGGTTGTCAGAAC</td>
<td>(AG)&lt;sub&gt;26&lt;/sub&gt;</td>
<td>121–123</td>
<td>2</td>
<td>0.083</td>
<td>0.080</td>
<td>0.077</td>
<td>KF680443</td>
</tr>
<tr>
<td>TRI_56</td>
<td>F: CACCTGATGGTTGTCAGAAC</td>
<td>(AG)&lt;sub&gt;15&lt;/sub&gt;</td>
<td>138–142</td>
<td>3</td>
<td>0.818</td>
<td>0.661</td>
<td>0.652</td>
<td>KF680445</td>
</tr>
<tr>
<td>TRI_57</td>
<td>F: CACCTGATGGTTGTCAGAAC</td>
<td>(AG)&lt;sub&gt;14&lt;/sub&gt;</td>
<td>147–155</td>
<td>3</td>
<td>0.833</td>
<td>0.517</td>
<td>0.420</td>
<td>KF680446</td>
</tr>
<tr>
<td>TRI_58</td>
<td>F: CACCTGATGGTTGTCAGAAC</td>
<td>(AG)&lt;sub&gt;14&lt;/sub&gt;</td>
<td>266–268</td>
<td>2</td>
<td>0.636</td>
<td>0.483</td>
<td>0.471</td>
<td>KF680447</td>
</tr>
<tr>
<td>TRI_59</td>
<td>F: CACCTGATGGTTGTCAGAAC</td>
<td>(AG)&lt;sub&gt;13&lt;/sub&gt;</td>
<td>190–192</td>
<td>2</td>
<td>0.667</td>
<td>0.444</td>
<td>0.346</td>
<td>KF680448</td>
</tr>
<tr>
<td>TRI_62</td>
<td>F: CACCTGATGGTTGTCAGAAC</td>
<td>(AG)&lt;sub&gt;11&lt;/sub&gt;</td>
<td>175–187</td>
<td>3</td>
<td>0.091</td>
<td>0.169</td>
<td>0.281</td>
<td>KF680450</td>
</tr>
</tbody>
</table>

Note: A = number of alleles per locus; H<sub>e</sub> = expected heterozygosity; H<sub>o</sub> = observed heterozygosity; PIC = polymorphism information content.