

Desarrollo de marcadores moleculares microsatélites derivados del transcriptoma foliar de *Cordia trichotoma* (Vell.) Arráb. ex Steud.

María Florencia Pomponio¹, Luis Fornes²,
María Cristina Soldati³, Susana Torales³

Palabras Clave: Peteribí, Microsatélites, Transcriptoma

Introducción

Cordia trichotoma (Vell.) Arráb. ex Steud (Boraginaceae) comúnmente conocido como “petiribí” o “afata” es una especie forestal nativa de alto valor debido a la calidad de su madera utilizada principalmente en la fabricación de muebles, revestimientos y como ornamental. En Argentina, se encuentra distribuida naturalmente en la selva de Yungas (NOA) y en la Selva Paranaense (NEA) (Tortorelli 2009) en donde está sujeta a un intenso aprovechamiento forestal de carácter extractivo y selectivo, sin técnicas de manejo adecuadas lo que ha generado un estado de alto riesgo para la especie.

El estudio de la variabilidad genética a través de técnicas de biología molecular nos permite determinar, por ejemplo, el estado de fragmentación y degradación de las poblaciones, identificación clonal o acortar tiempos en los programas de mejoramiento basados en la selección fenotípica. Estas aplicaciones, dependen del conocimiento de las secuencias de ADN (genoma) o ARN (transcriptoma) y de la disponibilidad de marcadores moleculares.

En especies no modelo como *C. trichotoma*, la falta de información genética ha sido una limitante para esta clase de estudios dado principalmente por los altos costos de secuenciación. En este aspecto con el surgimiento de las tecnologías de secuenciación masiva y de herramientas bioinformáticas asociadas, han posibilitado la secuenciación de grandes cantidades de ADN y ARN y el desarrollo de marcadores moleculares en poco tiempo y a un costo relativamente bajo. Dentro de los marcadores, los microsatélites o SSR son segmentos de ADN con un motivo de repetición en tándem. Estos marcadores, presentan una alta reproducibilidad, herencias codominantes y un

alto nivel de polimorfismo. En el género *Cordia*, previamente se han desarrollado marcadores microsatélites derivados de secuencias genómicas al azar en las especies *C. bifurcata* (Spoon & Kesseli 2008) y *C. alliodora* (Marulanda et al. 2011), de los cuales 10 han sido transferidos a *C. trichotoma* (Soldati et al. 2019). Hasta el momento no se han reportado marcadores específicos desarrollados para *C. trichotoma*.

Este estudio tiene como objetivo desarrollar marcadores moleculares SSR a partir de las secuencias obtenidas del transcriptoma foliar de *C. trichotoma* a fin de incluir nuevos marcadores moleculares que complementen los métodos de manejo y mejoramiento tradicionales y avanzar en el conocimiento genético para su conservación y producción sustentable.

Materiales y Métodos

Desarrollo del transcriptoma: El ARN total fue extraído de hojas de 3 genotipos de *C. trichotoma* provenientes de la región de las Yungas (PN Calilegua-Jujuy, R. Acambuco-Salta, y Cedro Solo-Salta) utilizando el método descrito por Chomczynski & Mackey (1995). La biblioteca de cDNA se realizó utilizando un Kit de Roche y luego fue secuenciado un cuarto de placa mediante la metodología 454-GS – FLX System (Roche). Las secuencias obtenidas fueron ensambladas en unigenes (isotigs y singletons) utilizando el programa Newbler 2.9 (Roche, IN, USA) y luego fueron comparadas contra la Base Viridiplantae de proteínas No Redundante del NCBI usando BLASTX (E-value 10⁻¹⁰) (Altschul et al. 1990).

Desarrollo y validación de marcadores microsatélites: Con las secuencias obtenidas se efectuó la búsqueda de motivos SSRs y se diseñaron primers (cebadores) con el programa GMATA (Wang & Wang 2016). Los parámetros para el diseño fueron: amplicon comprendido entre 100-300 pb, tamaño de cebador de entre 20-24 pb, y la exclusión de los motivos mononucleó-

¹ Instituto de Recursos Biológicos (IRB)-CIRN INTA Castelar. Contacto: florencia@inta.gob.ar.

² EEA Famailla INTA.

³ Instituto de Recursos Biológicos (IRB)-CIRN INTA Castelar.

tidos. La validación se realizó en 8 genotipos de la especie, provenientes de las regiones del NOA y el NEA. La extracción de ADN se realizó a partir de hojas según la metodología de CTAB propuesta por Hoisington et al. (1994). Los fragmentos de ADN portadores de marcadores microsatélites se amplificaron mediante la técnica de la reacción en cadena de la polimerasa (PCR), ajustando las condiciones para cada marcador. Los productos de amplificación fueron visualizados en geles de poliacrilamida teñidos con Nitrato de Plata.

Resultados

Como resultado de la secuenciación se obtuvieron un total de 196.717 secuencias con una longitud promedio de 465 bp, de las cuales 14.729 resultaron no redundantes. Sobre estas secuencias se identificaron 1.503 motivos microsatélites con una frecuencia de un 18% teniendo en cuenta las múltiples ocurrencias repetidas en un mismo locus (Tabla 1).

Los motivos predominantes fueron los dinucleótidos con un total de 985 motivos (6,54%), seguido por los trinucleótidos con 501 (33,33%) y en menos medida los tetra y los pentanucleótidos con 12 y 5 respectivamente (0,80% y 0,33%) (Figura 1).

La longitud de SSR más abundantes fueron las de 10 pb (484), seguida de 15 pb (309), 18 pb (173), 12 pb (158) y 14 pb (102). De acuerdo a la distribución en base al tipo de motivo repetitivo, para las repeticiones en tándem de dinucleótidos las frecuencias más abundantes se observaron para los pares TC/GA (202), TA/TA (188), AG/CT (187), AT/AT (180). Los motivos de repetición de trinucleótidos más comunes fueron TGA/TCA (57), AGA/TCT (48), TTC/GAA (40) y ATC/GAT (31).

Para 903 de los 1.503 motivos SSR identificados se diseñaron cebadores con el programa GMATA según los parámetros descritos en Materiales y métodos. De este grupo, 34 pares de cebadores fueron seleccionados para su validación en base a motivos microsatélites localizados en secuencias con potenciales funciones biológicas a determinadas a través del algoritmo BLASTx. Se priorizaron en genes de interés biológico, como por ejemplo los relacionados con características la respuesta a estrés abiótico o reloj circadiano (Tabla 2). Como resultado, un total de 31 marcadores (91%) presentaron productos de amplificación, y tres de ellos (Ct15, Ct21 y Ct25) no amplificaron o presenta-

ron una amplificación inespecífica por lo cual fueron descartados del análisis. El tamaño de los fragmentos amplificados abarcó el rango de 121 pb a 350 pb.

Dentro de los marcadores que amplificaron, 14 no presentaron variación (monomórficos) y 17 marcadores resultaron polimórficos con un número de alelos que van de 2 a 8 (Tabla 2). Los marcadores con un mayor número de alelos fueron Ct4, Ct26 y Ct38 (Tabla 2). En algunos marcadores se halló más de dos alelos por individuo, lo que es esperable dado que la especie presenta poliploidía ($2n = 104$, $x = 13$) (Las Peñas 2003). Esto también ha sido reportado por Soldati et al. (2019).

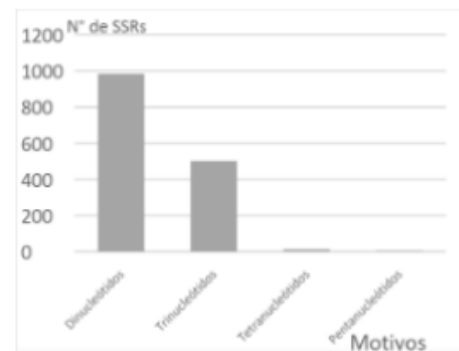


Fig. 1. Número de SSR según los tipos de motivos

Tabla 1. Resultados de la búsqueda de SSRs

Parámetros	Valores
Número total de secuencias analizadas	14.729
Tamaño total de secuencias examinadas (pb)	11.329.61
Número total de SSR identificados	1.503
Número de secuencias que contienen SSR	1.262
Número de marcadores únicos	798
Número de secuencias que contienen más de 1 SSR	188
Número de SSRs con primer diseñado	901
Número de SSRs sin primer	602

Tabla 2. Marcadores SSRs analizados

Nombre	(Motivo) de repet	Tamaño del fragmento (bp)		Polimorfismo	N° de alelos	Blast-Descripción función de la secuencia
		Exp	Obs			
C11	(TCA)5	121	121	M	1	CONSTANS-like protein 16a
C12	(TAC)5	203	203	P	5	Serine-threonine kinase receptor-associated protein
C14	(TGC)5	193	193	P	8	Putative WRKY transcription factor 32
C17	(GGC)5	178	200	P	2	Glycine-rich RNA-binding protein 4, mitochondrial
C18	(AT)5	146	146	P	2	Stress-response A/B barrel domain-containing protein UP3 isoform X2
C111	(GA)5	232	270	M	1	Transcription factor APETALA2
C113	(TGA)5	195	195	P	3	Transcription factor MYB1R1-like
C114	(AT)7	238	238	P	4	Probable protein phosphatase 2C 5
C115	(GAA)6	180	NA	-	-	Auxin early response protein AUX/IAA27
C116	(AG)11	204	204	M	1	WD40 repeat-containing protein HOS15
C117	(CT)5	198	198	M	1	WD40 repeat-containing protein MS14-like
C119	(AAG)5	182	182	P	3	Peroxidase 42-like
C120	(CT)5	153	180	P	3	FRIGIDA-like protein 4a
C121	(GTT)5	201	NA	-	-	Heat stress transcription factor B-2b
C124	(GA)7	239	350	M	1	Heat shock protein 90-2-like
C125	(CA)5	162	NA	-	-	Zinc finger protein GIS2-like
C126	(TTC)6	240	240	P	8	Protein SUPPRESSOR OF FRI 4-like
C127	(CAG)5	217	180	M	1	Transcriptional corepressor LEUNIG
C128	(TGA)5	217	217	M	1	Transcription factor TCP4-like
C129	(TG)5	183	183	M	1	Potassium channel SKOR
C130	(TGG)8	239	239	P	6	SDSL esterase/lipase
C131	(GTG)5	212	212	P	2	BELL-like homeodomain protein 2
C132	(GAT)7	134	134	P	5	Ethylene-responsive transcription factor 4
C133	(TC)5	266	300	M	1	Transcription factor PIF4
C134	(TCT)5	242	310	M	1	Bax inhibitor 1-like protein
C135	(TC)5	190	250	M	1	PREDICTED: allantoinase
C136	(AG)5	122	220	M	1	Protein REVELLE 7-like isoform X2
C137	(TA)6	185	185	P	6	Protein HOTHREAD-like isoform X1
C138	(TCT)5	162	162	P	8	Ethylene-responsive transcription factor RAP2-4-like
C139	(AAT)7	218	218	P	3	Protein SRC2 homolog
C141	(GA)5	198	198	M	1	Phenylalanine ammonia-lyase-like isoform X1
C142	(AT)8	190	190	P	3	Patanin-like protein 2
C145	(GT)5	156	156	P	5	PREDICTED: syntaxin-22-like
C146	(TCC)5	153	153	M	1	Auxin-related protein 2-like

Discusión y Conclusiones

En el presente trabajo se desarrollaron marcadores SSR para *C. trichotoma* a partir de secuencias de RNA desarrolladas con la tecnología 454 FLX Roche. Estos representan los primeros marcadores desarrollados en la especie.

La frecuencia de SSRs observada (18%) es ligeramente superior a la reportada en otras especies forestales nativas como *Prosopis alba* (11%, Torales et al. 2013), y similar a la reportada en *Nothofagus* (15%, Torales et al. 2012). Los motivos más abundantes fueron los dinucleótidos, al igual que en *Handroanthus impetiginosus* (Com. Pers.).

Se halló un alto nivel de polimorfismo en microsatélites localizados en secuencias de interés biológico. Estos marcadores al encontrarse en regiones expresadas podrán ser útiles para evaluar la diversidad genética funcional en la especie.

Este enfoque constituye un punto de partida para la caracterización de la diversidad genética y asistir a los programas de mejoramiento en curso, como así también a los programas de conservación en la especie.

Bibliografía Citada

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215: 403-410.
Chomczynski P MK. 1995. Modification of the TRI-

ZOL reagent procedure for isolation of RNA from polysaccharide-and proteoglycane-rich source. *Biotechniques* 19: 942-945.

Hoisington D, Khairallah M, González de LeÛn D. 1994. *Laboratory Protocols: CIMMYT Applied Molecular Genetics Laboratory*. Second Edition, México, DF. pp 2-3.

Las Peñas ML. 2003. Los cromosomas somáticos de *Cordia ecalyculata* y *C. trichotoma* (Boraginaceae). *Bol. Soc. Argent. Bot.* 38(3-4): 319-323.

Marulanda ML, Lopez AM, Uribe M, Ospina CM. 2011. Genetic variability of *Cordia alliodora* (R. & P.) Oken progenies. *Colombia Forestal* 14(2): 119-135.

Soldati C, Vitoria R, Ledesma T, Schmidt P, Rotundo C, Rodriguez G, Fornes L, Zelener N. 2019. VIII Reunión GeMFO. ISBN 978-987-86-1630-8.

Spoon TR, Kesseli RV. 2008. Development of microsatellite markers in *Cordia bifurcata* (Boraginaceae) and cross-species amplification in *Cordia inermis* and *Cordia pringlei* *Molecular Ecology Resources* 8: 989-992.

Torales SL, Rivarola M, Pomponio MF, et al. 2012. Transcriptome survey of Patagonian southern beech *Nothofagus nervosa* (= *N. alpina*): assembly, annotation and molecular marker discovery. *BMC genomics* 13: 291.

Torales S, Rivarola M, Pomponio MF, et al. 2013. De novo assembly and characterization of leaf transcriptome for the development of functional molecular markers of the extremophile multipurpose tree species *Prosopis alba*. *BMC Genomics* 14: 705.

Tortorelli LA. 2009. *Maderas y bosques argentinos*. Tomo I y II 2ª ed. Buenos Aires: Orientación Gráfica Editora. Argentina.

Wang X WL. 2016. GMATA: An Integrated Software Package for Genome-Scale SSR

Mining, Marker Development and Viewing. *Front. Plant Sci* 7: 1350.