



UNIVERSIDAD DE CONCEPCION

DIRECCION DE POSTGRADO

FACULTAD DE CIENCIAS FORESTALES-PROGRAMA MAGISTER EN CIENCIAS
FORESTALES

Asociación entre marcadores ISSR y el florecimiento de *Eucalyptus cladocalyx*

F.MUELL

RODRIGO IVAN CONTRERAS SOTO

Profesor Guía: Freddy Mora Poblete
Dpto. de Silvicultura, Facultad de Ciencias Forestales
Universidad de Concepción

CONCEPCION-CHILE

2013

Asociación entre marcadores ISSR y el florecimiento de *Eucalyptus cladocalyx*
F.MUELL

Comisión Evaluadora

Freddy Mora (Profesor guía)

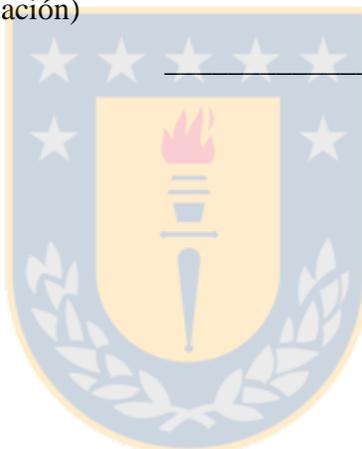
Ingeniero Forestal, Dr. _____

Eduardo Ruiz (Profesor co-guía)

Licenciado en biología, Dr. _____

Iván Maureira (Comisión evaluación)

Ingeniero agrónomo, Dr. _____



Director de Postgrado:

Darcy Ríos Leal

Prof. Biología y Química, Dra. _____

Decano Facultad de Ciencias Forestales:

Manuel Sánchez Olate.

Ingeniero Forestal, Dr. _____

AGRADECIMIENTOS

Gracias a Dios por acompañarme siempre y poner en mi camino a las mejores personas!

Con cariño especial dedico este trabajo a quienes me han enseñado a ver la vida de otra manera, mis padres, Máximo y Margarita, y a mis hermanas Maggi y Katty, quienes me entregaron todo su apoyo constantemente en esta etapa importante de mi vida. De manera especial, agradecer a Maggi por ser mi compañera, por ayudarme y asesorarme cada vez que la necesité. A mis Padrinos, Juan Alberto y Gladys, por todo su apoyo.

Agradezco a mis amigos quienes fueron un apoyo constante durante mi paso por la universidad, gracias por los momentos felices que pasamos. De manera especial agradezco a Felipe Vargas, Paulina Ballesta, Alexis Velásquez, Fran Beltrán, Javier Saavedra, y Claudia Gómez, por las buenas vibras, largas conversaciones y palabras de aliento que más de una vez necesité. Gracias a Juan, Feña, Christopher y Anita, por su importante compañía durante estos años.

Agradezco a mis Profesores guías Eduardo e Iván, y especialmente a Freddy Mora, por la oportunidad que me dio de integrarme al grupo. Agradezco las conversaciones, consejos, ideas, comentarios y enseñanzas entregadas durante cuatro años. Gracias por la paciencia, entrega, y dedicación durante la realización de la tesis. Infinitas gracias!

Gracias equipo CEAF. En especial a mis colegas, Ariel Salvatierra, Paula Pimentel, Pamela Rojas, Rubén Almada, Marcia Bravo y mi jefe Boris Sagredo, por todo el apoyo que me han dado.

Agradezco a los profesores y administrativos de la Facultad de Ciencias Forestales, especialmente a los profesores: Darcy Ríos, Carlos Figueroa y Rodrigo Hasbún, quienes han sido un gran apoyo durante mi estadía en la facultad. De manera muy especial dedico mi trabajo a quienes fueron un pilar fundamental: Marcelo Cuevas, Patricia Espinoza, Margarita Figueroa, Blanca Retamal, Alejandro Alarcón y Juanito, quienes han sido las mejores personas que he conocido, las cuales me regalaron una sonrisa, un saludo y un abrazo cada día que estuve en la facultad. A todos ustedes infinitas gracias!

INDICE

RESUMEN	5
ABSTRACT	6
INTRODUCCIÓN	7
REVISION BIBLIOGRAFICA	11
Información general, <i>Eucalyptus cladocalyx</i>	11
Importancia de la floración y su control genético.....	12
<i>Loci de característica cuantitativa</i> y desafíos para la selección asistida por marcadores.....	15
Floración y QTL	16
Inferencia Bayesiana en la detección de QTL	18
Estructura poblacional y detección de <i>QTLs</i>	20
HIPOTESIS	22
OBJETIVOS	22
CAPITULO 1	23
Abstract	23
Introducción	23
Materiales y Métodos.....	28
Ensayo de progenie y evaluación fenotípica	28
Datos genotípicos	29
Análisis de datos moleculares y estructura poblacional	30
Análisis de modelo bayesiano	31
Resultados y discusión.....	33
Variabilidad de la floración.....	33
Diferenciación genética y estructura poblacional.....	35
Asociación bayesiana entre marcadores ISSR y floración	37
CONCLUSIONES	39
REFERENCIAS	40

RESUMEN

Eucalyptus cladocalyx F. Muell, presenta una distribución natural fraccionada y restringida en el sur de Australia, donde es considerada una especie forestal idónea para el establecimiento de plantaciones productivas en zonas áridas y semiáridas. Por ello, en la región de Coquimbo, norte de Chile, se han establecido ensayos para evaluar productos tales como: componentes de floración para la producción de miel de *Apis mellifera*, madera para la producción de postes, polines, y biomasa para la obtención de bioenergía.

En *E. cladocalyx*, el uso de estrategias convencionales de mejoramiento ha permitido estimar valores genéticos utilizados con fines de selección. Al respecto, el objetivo de la presente investigación permitiría complementar dicha selección al identificar marcadores moleculares ISSR asociados a *loci* de características cuantitativas que controlan en parte la floración en la especie. Para ello, un ensayo de procedencia-progenie de la especie, establecido en septiembre de 2001, se evaluó genéticamente en función de la floración precoz (EF) e intensidad de floración (FI). El experimento de campo consistió en un diseño de bloques completos al azar, con 30 bloques, conteniendo las siguientes poblaciones australianas: Cowell; Marble Range; Mt. Remarkable; Flinders Chase NP y Wirrabara. Se utilizaron ocho marcadores moleculares ISSR para la determinación de la estructura genética poblacional, realizada mediante un análisis de agrupamiento bayesiano, con el programa STRUCTURE. Paralelamente, se utilizaron dos modelos bayesianos para estimar los valores genéticos, y evaluar la asociación ISSR- floración, acá se utilizó un modelo adicional sin efecto de estructura genética.

Se encontraron altos niveles de diferenciación genética entre las procedencias ($F_{st}=38\%$). El análisis de estructura poblacional indica que las procedencias se agrupan en dos diferentes clusters. De acuerdo al análisis bayesiano se encontró un locus asociado a EF, que explica un 17% de la variación fenotípica del carácter, mientras que en FI se encontraron 3 loci, que explicaron entre un 8-16% de la variación fenotípica. Interesantemente, el modelo sin estructura indica que un locus fue falso positivo de asociación. La asociación con motivos GA sugiere la existencia de locus relacionados al locus de temporada de floración (gen *SFL*) que ha sido escasamente estudiado en *Eucalyptus*.

ABSTRACT

Eucalyptus cladocalyx F. Muell has a fragmented and restricted natural distribution in South Australia, where it is considered a suitable forest species for the establishment of productive plantations in arid and semiarid regions. Therefore, in the Region of Coquimbo, northern of Chile, trials have been established to evaluate several traits such as, flowering for the production of honey of *Apis mellifera*, wood for poles and wooden stools, and biomass for bioenergy production.

In *E. cladocalyx*, the use of conventional breeding strategies has allowed to estimate breeding values used for selection purposes. The objective of this research would complement the selection through the use of ISSR molecular markers associated with quantitative trait loci for flowering in the specie. For this, a provenance-progeny test established in September of 2001 was evaluated for early flowering (EF) and flowering intensity (FI). The trial consisted of a random complete block design with 30 blocks, containing the following Australian populations: Cowell, Marble Range, Mt Remarkable, Flinders Chase NP and Wirrabara. We used eight ISSR molecular markers to determine population structure, by using a bayesian cluster analysis in STRUCTURE software. Two bayesian models were implemented for estimate breeding values and evaluate the association between ISSR markers and flowering traits; additionally we evaluate the genetic association in an additional model without population structure effect.

We found high levels of genetic differentiation between provenances ($F_{st} = 38\%$). Population structure analysis indicates that populations were grouped in two different clusters. According to bayesian analysis, we found one locus associated to EF, which explains 17% of the phenotypic variation, and three loci associated with FI, which explained between 8-16% of the phenotypic variation. Interestingly, the model within population structure effect indicates that one locus was false positive of association. The association with GA-motifs suggests the presence of loci related to seasonal flowering locus (SFL gene), which has been poorly studied in *Eucalyptus* sp.

INTRODUCCIÓN

El género *Eucalyptus* pertenece a la familia de las Myrtaceae, representado por 108 especies distribuidas principalmente en el hemisferio sur. La mayoría de éstas son nativas de Australia, Papua, Nueva Guinea, Indonesia y Filipinas. Su amplia diversidad y plasticidad fenotípica entre individuos de la misma especie, le han permitido crecer desde el nivel del mar hasta las altas montañas (Grattapaglia et al. 2012). Plantaciones comerciales de *Eucalyptus* han sido establecidas en diversos ambientes, mayoritariamente en Brasil (3.751.857 ha), India (3.942.600 ha), y China (2.609.700) (Iglesias 2009). Adicionalmente, desde hace algunas décadas se han desarrollado diversos programas de mejoramiento en al menos siete países: Argentina, Australia, Chile, Etiopía, Portugal, España y Uruguay (Potts et al. 2004), donde actualmente existen plantaciones con fines comerciales.

Eucalyptus cladocalyx F. Muell, es un árbol endémico del sur de Australia, distribuido en tres regiones geográficamente fraccionadas en un rango total de 500 km. Dos regiones localizadas en el continente Australiano (Flinders Range y Península Eyre), y la otra en la Isla Kangaroo. La especie ha sido considerada en programas de mejoramiento debido a su capacidad para crecer en zonas de bajas precipitaciones, donde podría ser usada, por ejemplo, para la producción de madera y postes, así como para la remediación de suelos contaminados por la industria minera (Hardwood et al. 2007). En el sur del desierto de Atacama, Región de Coquimbo, Chile, la especie está siendo evaluada genéticamente para la producción de postes y miel (Mora et al. 2009), considerando que sus flores constituyen una fuente de néctar que complementa la errática floración de las especies endémicas y nativas de esta zona (Montenegro et al. 2003). En Chile, diversos estudios acerca de los niveles de floración de *E. cladocalyx*, han confirmado una amplia variabilidad genética, convirtiendo a la característica en un parámetro de expresión diferencial considerando la evaluación de procedencias que representan su distribución natural de origen (Mora et al. 2009). Adicionalmente la especie es preferida por sobre otras, debido a su moderada tolerancia a salinidad (Marcar et al. 1995), resistencia a patógenos (Hanks et al. 1994) y su habilidad para crecer en ambientes áridos y semiáridos (Mora et al. 2009; Bush et al. 2011; Contreras-Soto et al. 2011).

Desde el punto de vista de la evaluación genética de caracteres complejos, la predicción de valores genéticos han sido calculados, tradicionalmente, utilizando la metodología de modelos

lineales mixtos, que calculan el mejor predictor lineal insesgado (BLUP), descomponiéndose la varianza fenotípica en los componentes genético y ambiental. El objetivo es predecir los valores genéticos individuales o el mérito genético de los árboles, que corresponde a la proporción de alelos favorables transferidos desde los parentales a su progenie. Tales valores genéticos consideran la suma de miles de efectos alélicos, localizados físicamente en algún lugar del genoma a través del uso de marcadores moleculares (Zapata-Valenzuela et al. 2012).

La metodología BLUP ha sido ampliamente desarrollada en programas de mejoramiento animal (Henderson 1949; Mrode 2005), y ahora está siendo utilizada en varias áreas de la investigación, principalmente en la selección fenotípica y molecular de plantas. BLUP ha mostrado ser apropiado para la predicción de valores genéticos cuando los caracteres presentan distribución continua (altura, diámetro, rendimiento etc.) (Piepho et al. 2008). Cuando el carácter presenta distribución discreta, por ejemplo presencia/ausencia de una determinada enfermedad, se requieren diferentes métodos de evaluación (Xu y Hu, 2011). Una alternativa es el uso de métodos bayesianos, el cual ha sido destacado por calcular con mayor precisión el mérito genético en caracteres que no presentan distribución normal (Van-Tassell y Van Vleck, 1995; Hadfield et al. 2010). Esta metodología ha sido utilizada para evaluar genéticamente diferentes especies de árboles forestales (Mora et al. 2009), y de acuerdo a Lund et al. (2009), el análisis bayesiano es la metodología más prometedora para su uso en la predicción de valores o mérito genético.

Los marcadores moleculares corresponden a regiones del genoma altamente conservadas. Su utilización en especies forestales ha permitido caracterizar poblaciones naturales y de mejoramiento, y en conjunto con informaciones fenotípicas, ha permitido mapear e identificar loci de características cuantitativas (QTL: del inglés *quantitative trait loci*). En los últimos años, diversos reportes han enfatizado en la relevancia de establecer estudios relacionados al genoma de *Eucalyptus*, incluyendo las aplicaciones al mejoramiento y conservación (Byrne 2008; Grattapaglia y Kirst 2008; Myburg et al. 2007; Poke et al. 2005). De la misma manera, mapas de ligamiento para múltiples caracteres, incluyendo crecimiento, calidad de madera, resistencia a enfermedades y tolerancia a diversos tipos de estrés, se han basado en el uso de cientos de marcadores tales como: RFLP; RAPD; AFLP y microsatélites (Grattapaglia and Sederoff 1994; Verhaegen et al. 1997, Marques et al. 1998; Myburg et al. 2003). Estos

estudios se han desarrollado en base a poblaciones segregantes o biparentales (F₂, retrocruzamiento, líneas endogámicas recombinantes, etc.), las cuales pueden ser una limitante para especies que presentan altos niveles de depresión endogámica y largos periodos de rotación, tal como las especies forestales.

Por otra parte, investigaciones recientes indican que no existe un uso operacional de los QTL descubiertos en eucalipto (Grattapaglia y Resende, 2011; Grattapaglia et al. 2012). Las causas se atribuyen a que la mayoría de los QTL han sido ampliamente sobreestimados (Grattapaglia et al. 1995; Shepherd et al. 1999), porque los procedimientos comúnmente referidos a mapeo por ligamiento presentan limitaciones, las cuales incluyen escasa resolución en la detección de QTLs (Beavis 1994; Stich et al. 2008), y cuya asociación es válida para una estructura poblacional específica (Grattapaglia y Resende, 2011). Por esta razón, diferentes metodologías han sido examinadas para estudiar caracteres cuantitativos, considerando que la eficiencia de mapeo por ligación y la construcción de mapas han sido limitadas en diversos programas de mejoramiento genético de árboles forestales, debido al rápido decaimiento del desequilibrio de ligación (Neale y Savolainen, 2004).

Estudios de asociaciones entre el fenotipo y el genotipo han sido comunes en especies forestales (Brawner et al. 2012; Lepoittevin et al. 2012), considerando que los marcadores moleculares han permitido detectar asociaciones entre la variabilidad fenotípica y regiones genómicas, las cuales pueden ser aplicadas a nivel poblacional (Cumbie et al. 2011). La integración entre la información molecular con un carácter de interés para los programas de mejoramiento provee una conexión con los datos fenotípicos que históricamente han sido colectados para mejorar un carácter (Brawner et al. 2012). En general, existen dos métodos utilizados para evaluar la significancia de relación entre marcadores y fenotipo: uno se basa en el concepto de distancia física o mapeo de ligamiento, el cual ha mostrado ser ineficiente en árboles forestales, y el otro se basa en un modelo probabilístico que establece una asociación entre los marcadores y el fenotipo (Brawner et al. 2012). Cuando el análisis de asociación involucra un amplio número de marcadores distribuidos en todo el genoma (no genes candidatos), es común utilizar los métodos de regresión Ridge (Endelman 2011) o regresión bayesiana. En esta metodología el uso de variaciones de las técnicas de Monte Carlo, vía Cadenas de Markov (MCMC), tal como el algoritmo de Gibbs, ha mostrado ser útil en el

análisis de QTLs (Xu y Hu, 2011). En este contexto, diversos programas estadísticos han implementado el muestreo de gibbs para el análisis bayesiano (SAS Institute Inc, 2009). Recientemente, Hu y Xu (2009) publicaron un procedimiento para el análisis de QTL, en el cual enfatizan las ventajas del análisis bayesiano por sobre el método clásico de máxima verosimilitud.

El efecto de la estructura poblacional es la principal limitante en estudios de genética asociativa (Pritchard et al. 2000). La presencia de grupos a nivel poblacional genera falsas asociaciones, la cual se debe a la relación entre los individuos de una misma especie (Pritchard y Rosenberg, 1999). Con el objetivo de eliminar este sesgo, metodologías basadas en análisis bayesiano han permitido determinar la relación de estructura poblacional entre individuos emparentados. En el caso de *E. cladocalyx*, McDonald et al (2003) utilizando marcadores isoenzimáticos, encontró que la especie presenta un alto grado de estructuración en condiciones naturales, y que además presenta niveles de endogamia más altos en comparación a la mayoría de las otras especies de *Eucalyptus*. Estos resultados ha sido recientemente confirmados por Bush y Thumma (2013), quienes encontraron elevados niveles de endogamia, particularmente en la isla Kangaroo, sumado a un alto grado de estructuración ($F_{st}=0,18$).

A pesar de la distribución natural restringida de *E. cladocalyx*, altos niveles de diferenciación entre poblaciones naturales han sido observadas en los parámetros de crecimiento, sobrevivencia, florecimiento precoz (Mora et al. 2009), y durabilidad de la madera (Bush et al. 2011). En Chile, considerando el carácter de floración temprana, se ha observado que la magnitud de diferencia entre la mejor y la peor procedencia fue muy alta (82%) (Mora et al. 2009). Esta diferencia fue confirmada por Contreras-Soto et al (2011) nueve años después de la plantación. Consecuentemente, Cané-Retamales et al (2011), encontró una alta correlación genética entre floración precoz e intensidad de floración ($\Gamma=0,45$). De acuerdo a las variaciones fenotípicas en procedencias naturales de la especie, y lo encontrado en ensayos establecidos en Chile, el presente trabajo sugiere que existe un alto grado de diferenciación genética (i.e Índice de fijación de Wright o F_{st}), mayor a lo reportado por Bush y Thumma (2013), por ejemplo, quienes utilizaron marcadores moleculares SNP para evaluar árboles de dos de las tres regiones de la distribución natural de la especie (Isla Kangaroo y Peninsula Eyre). Este alto grado de diferenciación, sumado a la amplia variabilidad fenotípica de la

floración reportada en los ensayos de campo permitirían identificar QTLs de efecto mayor, usando los marcadores moleculares ISSR asociados a la característica cuantitativa que controla en parte la floración en la especie.

REVISION BIBLIOGRAFICA

Información general, *Eucalyptus cladocalyx*

Eucalyptus cladocalyx F. Muell, es un árbol endémico del sur de Australia, distribuido en tres regiones fraccionadas en un rango total de menos de 500 km (Bush et al. 2011). Dos centros de origen están ubicados en la parte continental y uno en la Isla Kangaroo (Ruthrof et al. 2003). Por sus características morfológicas y fisiológicas, ha sido catalogada como una especie idónea para zonas áridas y terrenos considerados marginales desde el punto de vista productivo (Hobbs et al. 2009; McMahon et al. 2010; Cané-Retamales et al. 2011). Aunque su utilidad para establecer programas de mejoramiento, fue identificada en la década de 1870, sólo recientemente se han realizado colecciones sistemáticas de semillas de todas las procedencias naturales conocidas de la especie (Bush et al. 2009). Actualmente, diversos programas de mejoramiento genético cooperativos se están llevando a cabo en Australia, posterior a la evaluación de múltiples ensayos genéticos (Callister et al. 2008; Bush et al. 2011).

E. cladocalyx puede adaptarse bien fuera de su área de distribución natural, prefiriendo zonas con precipitación media de 400 a 1010 mm anuales y entre 0 y 600 m de altitud (Jovanovic y Booth 2002). Diversas plantaciones han sido establecidas con éxito en Etiopía y el sur de África (Lesotho, Mozambique, Sudáfrica y Zimbabwe), y en varios países mediterráneos (Argelia, Italia, Marruecos, Grecia, España, Portugal e Israel) (Poynton, 1979; Jacobs, 1981; Turnbull y Pryor, 1984). En la región de Coquimbo, norte de Chile; una zona de clima mediterráneo con potencial para el desarrollo de especies forestales que toleran continuos periodos de sequía (Montenegro et al. 2003), múltiples ensayos de *E. cladocalyx* han sido establecidos como recursos forestales para la rehabilitación de suelos y/o producción agroforestal (Mora et al. 2007). Resultados preliminares han demostrado que las diferencias fenotípicas en crecimiento, altura e intensidad de florecimiento, son consistentes con la variabilidad genética presente en las poblaciones estudiadas, por cuanto se discute la

importancia de una selección que combine el uso de marcadores moleculares (Mora et al. 2009).

La región de Coquimbo tiene una gran diversidad florística, mayor que en otras regiones de Chile (Montenegro et al. 2003), la que es utilizada en parte para producir miel de *Apis mellifera* (Mora et al. 2007). Al respecto, Montenegro et al (2003) sostiene que la producción de miel constituye una fuente de ingreso de comprobada rentabilidad para los agricultores de la zona. Al estudiar la composición del polen en diferentes apiarios de la zona, se encontró la presencia de polen de varias especies de *Eucalyptus* en aproximadamente el 38% de las explotaciones apícolas, siendo *Eucalyptus globulus* la principal fuente de néctar (Montenegro et al. 2003). Sin embargo, árboles de *E. globulus* son inherentemente susceptibles a la sequía, bajas temperaturas y heladas (Coopman et al. 2010), lo que limita el uso potencial de sus flores para la producción de miel. Por otra parte, *E. cladocalyx* ha mostrado buenas tasas de floración (Mora et al. 2007; Cané-Retamales et al. 2011; Contreras-Soto et al. 2011), por lo que sus flores proveerían una excelente fuente de néctar para la producción de mieles de buena calidad (McMahon et al. 2010).

Importancia de la floración y su control genético

E. cladocalyx es una especie heteroblástica, y sólo produce flores cuando los árboles han alcanzado la madurez (Meilan 1997). Presenta una inflorescencia de color blanco, poco vistosa y en grupos de 7 a 11 flores o más. En poblaciones naturales, la floración se produce entre los meses de enero a abril (Blakely 1955; Boomsma 1981), y la polinización es principalmente por insectos y pájaros (Pryor 1976). Los frutos son ovoides con forma de campana con 3-4 válvulas muy cerradas. La semilla se encuentra disponible para la recolección a los 12 meses después de la floración y puede ser retenida en el árbol durante 3-4 años. La recolección de semillas puede ser a lo largo de todo el año (McMahon et al. 2010).

Fenotípicamente los árboles de *Eucalyptus* se caracterizan por presentar variabilidad intraespecífica en sus patrones de floración (Jacobs 1955; Brooker y Kleinig 1990). Con ello, el tiempo de floración puede variar entre especies, individuos de la misma especie y por localidad (Chambers et al. 1997; Potts et al. 2001), y como en la mayoría de los árboles perennes, se presenta un ciclo de vida con una fase juvenil de varios años de duración (no

floreCIMIENTO) seguidos de una fase adulta en la que hay ciclos repetidos entre crecimiento vegetativo y reproductivo (Meilan et al. 1997). Por ejemplo, algunas especies de climas más fríos como *E. globulus*, *E. nitens*, *E. regnans* y *E. dunnii* no florecen hasta que alcanzan los 7 a 10 años de edad (Eldridge et al. 1993). Pero esto es relativamente variable, ya que algunas familias y procedencias de *E. globulus*, rara vez florecen en condiciones normales de plantación, mientras que otros pueden florecer entre los primeros 2-3 años (Chambers et al. 1997; Dutkowski y Potts 1999). Interesantemente, genotipos de esta especie (procedencia Wilsons Promontory) han florecido durante el primer año de vida después de la plantación (Jordan et al. 1999; Potts et al. 2001). Asimismo, árboles establecidos individualmente pueden florecer durante un tiempo aceptable, por ejemplo: *E. globulus*: 37 días promedio (Gore y Potts 1995), *E. regnans*: 43 días promedio (Griffin 1980) y *E. urnigera*: un máximo de 100 días (Savva et al. 1988).

El largo período juvenil descrito en diferentes especies de eucalipto, es un impedimento para el estudio de la biología reproductiva, y una limitante para acelerar las pruebas de progenie en los programas de mejoramiento genético (Jordan et al. 1999). Este largo periodo de espera, representa una pérdida de tiempo invaluable cuando se trata de adquirir semillas mejoradas. En Chile, las especies más importantes para plantación como *E. globulus* y *E. nitens* podrían no florecer hasta los 10 años de edad, lo mismo sucede en *E. dunnii* (Potts et al. 2001). En *E. cladocalyx*, los árboles no florecen hasta que tienen entre 3-6 años después de la plantación (Mora et al. 2009), y a la fecha no existe reporte de la edad del primer florecimiento en las poblaciones naturales de origen (Potts et al. 2001), sin embargo, en ensayos de progenie establecidos en el norte de Chile, Mora et al (2007) ha reportado variabilidad en el florecimiento precoz (medido a los 30 meses de edad) considerando procedencias del rango de distribución natural de la especie. Sanhueza et al (2010), señala que para Chile, la edad del primer florecimiento en diferentes especies del género, tiene mucho que ver con la zona climática donde están establecidas las plantaciones, además del control endógeno y señales ambientales que actúan a través de una vía genética compleja que se ha estudiado ampliamente en plantas anuales como *Arabidopsis* (Yaish et al. 2011; Xu et al. 2012), pero que sólo recientemente ha sido dilucidada en plantas perennes (*Populus trichocarpa*-Böhlenius et al. 2006), incluyendo una especie de *Eucalyptus* (*E. globulus subsp globulus*-Jones et al. 2011).

Además del componente genético, el medioambiente y los tratamientos silviculturales son factores determinantes de la intensidad de floración en eucaliptos (Eldridge et al. 1993; Moncur y Hasan 1994; Moncur et al. 1994). Por cuanto, resultados similares a los de precocidad se ha reportado en diferentes especies del género. Por ejemplo, Law et al (2001) al evaluar 14 especies de *Eucalyptus* (*E. acmenoides*, *E. bancroftii*, *E. grandis*, *E. microcorys*, *E. siderophloia*, *E. pilularis*, *E. propinqua*, *E. resinifera*, *E. robusta*, *E. saligna*, *E. signata*, *E. tereticornis* y *E. tetrapleura*) en el norte de New South Wales (Australia), encontraron que la abundancia floral fue altamente variable entre especies, sitios y años, variando desde floración anual a ninguna floración durante un período de 10 años. En el norte de Chile, se ha reportado alta variabilidad en la intensidad de floración de *E. cladocalyx*, al encontrar que 45% de los árboles de la procedencia de Cowell (Australia) presentaron máxima intensidad de florecimiento, en contraste con el 3% presentado por Flinders Chase NP (Australia) (Contreras-Soto et al. 2011). La amplia variabilidad encontrada en los patrones de precocidad e intensidad de floración, tiene una explicación relacionada a la variación continua que presentan las características controladas por la acción conjunta de varios genes de efecto menor, como es el caso del florecimiento (Missiaggia et al. 2005; Thumma et al. 2010). En este contexto, la característica ha mostrado ser altamente heredable en distintas especies del género (*E. globulus*= 0,47 ó 0,59 Chambers et al. 1997; *E. risdonii*–*E. tenuiramis*= 0,31- 0,41 Wiltshire et al. 1998; *E. cladocalyx*= 0,52 Mora et al. 2009), lo que sugiere que la característica presenta un alto control genético aditivo (Chambers et al. 1997).

En términos generales, los mayores alcances para el estudio de la floración en diversas especies de eucalipto se han focalizado en: promover y evaluar el control genético del primer florecimiento (Gore y Potts 1995; Jordan et al. 1999; Williams et al. 2003), caracterizar la biología reproductiva (Potts et al. 2008), y contribuir a la generación de huertos semilleros (Varghese et al. 2009b). Adicionalmente se han aislado genes ortólogos putativos asociados al florecimiento de *Arabidopsis thaliana* y se ha examinado su expresión en plantas transgénicas de *Eucalyptus* (Southerton et al. 1998; Dornelas et al. 2004; Dornelas y Rodriguez 2005) y más recientemente se han identificado regiones genómicas relacionadas al florecimiento precoz (Farro et al. 2013).

Del punto de vista económico, las plantaciones comerciales de *Eucalyptus* ssp pueden ser importantes para la industria melífera. Por ejemplo, en New South Wales (NSW) Australia, 14 de 26 fuentes principales de miel fueron árboles forestales (Somerville y Nicholson 2005; Birtchnell y Gibson 2006; Birtchnell y Gibson 2008; Law y Chidel 2007), siendo los eucaliptos, las especies con mayor participación en la producción de la miel en la empresa Capilano Pty Ltda (Somerville y Moncur 1997). En las regiones áridas y semiáridas de Chile la producción de miel constituye una importante alternativa para los agricultores locales (Montenegro et al. 2003). En estas regiones, la producción de miel depende, entre otros factores, de la existencia de abundancia floral y de su adecuada distribución en el año (Porter 1978). Observaciones previas realizadas por agricultores de la región de Coquimbo, confirman que *E. cladocalyx* florece aproximadamente entre los meses de noviembre y marzo; un período crítico desde el punto de vista de la disponibilidad floral, ya que se caracteriza por la ausencia total de lluvias, por ello sería ventajoso caracterizar genéticamente individuos que florecen intensamente bajo esas condiciones. Las principales características de la miel monofloral originada de *E. cladocalyx* ha sido reportada previamente en el estudio de Manikis y Thrasivoulou (2001), y fueron indicadas en el directorio de recursos de mieles monoflorales de la Asociación Internacional de Investigación en Apicultura (Crane et al. 1984).

Loci de característica cuantitativa y desafíos para la selección asistida por marcadores

Las regiones dentro del genoma que presentan genes asociados con un carácter cuantitativo son conocidas como *loci* de características cuantitativas (QTL: del inglés *Quantitative trait loci*) (Collard et al. 2005). Características importantes para la industria forestal como crecimiento en volumen, rendimiento pulpable, contenido de lignocelulósicos, resistencia a enfermedades, estrés abiótico y biótico son de tipo cuantitativas (Neale y Kremer 2011), y generalmente son de herencia poligénica (Tanskley 1993). El análisis de tales *loci* permite mejorar la comprensión del control genético de diferentes rasgos, ayudando a identificar regiones genómicas específicas, y en última instancia, la expresión de los genes involucrados (Freeman et al. 2009). En eucalipto, durante los últimos 20 años, un considerable número de estudios han permitido en primera instancia diseccionar caracteres cuantitativos, reportando mapas de ligamiento basados en el uso de cientos de marcadores, tales como: RFLP, RAPD, AFLP y microsatélites (Grattapaglia y Sederoff 1994; Verhaegen et al. 1997; Marques et al.

1999; Myburg et al. 2003). La construcción de tales mapas se ha llevado a cabo utilizando la genealogía (*pedigree*) de árboles biparentales, generalmente padres altamente heterocigotos y a su progenie; ya sean hermanos completos o medios hermanos disponibles de los programas de mejoramiento operacional (Grattapaglia y Sederoff 1994; Grattapaglia et al. 1994; Grattapaglia et al. 1996). Al respecto, predicciones iniciales de que los QTLs detectados mediante esta metodología sólo serían útiles dentro de las mismas familias (Grattapaglia et al. 1994), ha sido confirmado por varias investigaciones que muestran que muchos más QTLs tienen efectos pequeños y variables entre medioambiente y progenies, normalmente subyacen en rasgos complejos (Dillen et al. 2008; Mamani et al. 2010). Por consiguiente, la detección de QTL mediante las metodologías anteriormente descritas, ha tenido una limitada aplicación en el mejoramiento y estudio de genes funcionales (Grattapaglia y Kirst 2008; Resende et al. 2012).

A pesar de que en árboles forestales el mapeo de QTLs se presentó como una opción para desarrollar métodos más eficientes de selección asistida por marcadores (*Marker Assisted Selection* or MAS) (Neale y Savolainen 2004), los resultados experimentales aún no han capturado fracciones satisfactorias de la varianza genética para ser útiles en mejoramiento (Resende et al. 2008; Resende et al. 2012). Y aunque actualmente son muchos los artículos científicos publicados bajo el concepto de MAS en QTL, casi todos se basan en demostrar las aplicaciones potenciales para los programas de mejoramiento, más que su uso práctico (Grattapaglia y Resende 2011; Schuster et al. 2011). Por consiguiente, hasta la fecha, los resultados cuestionan la importancia de los esfuerzos de mapeo para caracteres cuantitativos en los programas de mejoramiento (Resende et al. 2012).

Floración y QTL

La floración ha sido identificada como una característica importante en diversos programas de mejoramiento, incluyendo principalmente especies de larga rotación tal como el caso de *Eucalyptus* (Chambers et al. 1997; Wiltshire et al. 1998; Missiaggia et al. 2005; Jones et al. 2011; Mora et al. 2009; Cané-Retamales et al. 2011). Al ser la floración una característica controlada por múltiples genes, se considera útil el análisis genético a partir de la detección de QTLs. Por ejemplo, Bundock et al (2008) encontró un QTL que afecta la floración temprana en *E. globulus*. Así también, Missiaggia et al (2005) identificó un QTL de efecto mayor en el

fenotipo de floración temprana en *E. grandis*. No obstante, y a pesar de estos descubrimientos, un estudio reciente realizado en *E. nitens*, donde se encontró un QTL para florecimiento, señala la necesidad de otros estudios para determinar el o los QTLs que controlan el rasgo (Thumma et al. 2010), considerando que el porcentaje de variación fenotípica explicada por estos QTL es generalmente baja, como ocurre con la mayoría de los caracteres cuantitativos (Buckler et al. 2009). A pesar de estos resultados, en ninguno de los casos se ha evaluado la ganancia genética de seleccionar en función de marcadores moleculares, esto representa un desafío para la selección asistida por marcadores basada en el análisis de asociación, más aún cuando existe especial interés por mejorar la ganancia genética operacional en especies forestales.

Diversos marcadores moleculares han sido ampliamente utilizados para determinar *loci* de características cuantitativas en *Eucalyptus sp*, incluyendo RAPD (*Random Amplified Polymorphism DNA*) (Propagación vegetativa, Grattapaglia et al. 1995; Thumma et al. 2010); AFLP (*Amplified Fragment Length Polymorphism*) y SSR (*Simple Sequence Repeats*) (Crecimiento y floración, Bundock et al. 2008). La importancia de esta información se relaciona con la buena transferibilidad de marcadores entre las especies del género *Eucalyptus* (Brondani et al. 2006). ISSR (*Inter-Simple Sequence Repeat*), que corresponden a secuencias flanqueadas por marcadores SSR, han mostrado alta aplicabilidad, reproducibilidad y distribución en el genoma. Una gran ventaja frente a SSR, es que su utilización no requiere conocimiento previo de la secuencia de ADN (Bornet y Branchard 2001; Chezian et al. 2010); proveen un enfoque novedoso de fingerprinting aplicable en las comparaciones taxonómicas y filogenéticas y constituye una herramienta de mapeo genético en una amplia gama de organismos (Zietkiewicz et al. 1994). En *Eucalyptus sp* se han utilizado recientemente para evaluar la diversidad y la relación genética dentro y entre poblaciones (Balasaravanan et al. 2005; Okun et al. 2008; Chezian et al. 2010), y en otras especies forestales, como por ejemplo *Pinus radiata*, ha sido propuesto como una herramienta para la evaluación en programas de mejoramiento (Moraga-Suazo et al. 2012). La asociación que pudiera existir entre ISSR- fenotipo es el fundamento que hay detrás de la selección asistida por marcadores, y se relaciona con la existencia de genes con efectos significativos que puedan ser destinados específicamente a la selección. Al respecto, una serie de características

cuantitativas evaluadas en un sinnúmero de cultivos, se han podido seleccionar en función de la MAS incorporando el uso de ISSR (Reddy et al. 2002).

Inferencia Bayesiana en la detección de QTL

Al evaluar rasgos cuantitativos en los programas de mejoramiento genético, usualmente se utilizan modelos lineales mixtos para descomponer la varianza fenotípica en sus componentes genético y ambiental. El objetivo es predecir valores genéticos aditivos (proporción de alelos favorables tras pasados desde los padres hacia su progenie). Estos se tratan como variables aleatorias en el cálculo de (co)varianzas, junto a variables fijas como sitio o bloque. El resultado es el mejor predictor lineal insesgado (del inglés *Best Linear Unbiased Prediction* o BLUP) del valor genético para cada individuo (Henderson 1949; Henderson 1985), valor utilizado por el genetista para seleccionar los mejores fenotipos. Desde hace varios años, BLUP ha sido la metodología más ampliamente utilizada en los programas de selección para el mejoramiento genético animal (Henderson 1949; Mrode 2005), y ahora es utilizado en diversas áreas de la investigación, incluyendo el mejoramiento de plantas, donde ha demostrado tener una precisión predictiva buena en comparación con otros procedimientos (Piepho et al. 2008).

Cuando un carácter presenta una distribución discreta (i.e Resistencia, sobrevivencia, florecimiento precoz), se requiere de métodos alternativos para la estimación de valores genéticos (Van Tassell y Van Vleck 1996; Hadfield et al. 2010). Una opción es el uso de métodos bayesianos, considerados como una metodología precisa para su estimación en caracteres que presentan distribución discreta (Van Tassell y Van Vleck 1996), y que han sido utilizados para evaluar genéticamente una serie de diferentes especies de árboles forestales (Cappa y Cantet 2006; Waldmann y Ericsson 2006; Mora et al. 2009). Es así como, variaciones de cadenas de Markov Monte Carlo (MCMC), y algoritmo de Gibbs Sampling han demostrado ser apropiados para el análisis de datos categóricos en los programas de mejoramiento (Mora et al. 2009).

La metodología bayesiana ha sido ampliamente aceptada por los investigadores (Sillanpää y Corander 2002; Wang et al. 2005; Yi y Xu 2008) y según diversos autores se convertirá en el principal método desarrollado para la búsqueda de QTLs (Xu y Hu 2011). Actualmente

existen variados programas estadísticos que han implementado análisis bayesiano (ie. WinBUGS, JAGS, SAS) utilizando por ejemplo algoritmo de MCMC (Xu y Hu 2011), y algoritmo de *Gibbs*. En el programa SAS (SAS Institute Inc, 2009), por ejemplo, los procedimientos GENMOD, PHREG, LIFEREG y MIXED han sido propuestos para ser utilizados en el análisis bayesiano de búsqueda y mapeo de QTLs (Xu y Hu 2011). Adicionalmente, y considerando el actual uso de marcadores moleculares para los programas de mejora, todos los modelos bayesianos han incorporado el uso de algoritmo de *Gibbs*, en la que los efectos de marcador o QTL se supone que siguen una distribución mixta, y donde se asume que relativamente pocos marcadores explican una varianza grande y un gran número explica una varianza muy pequeña (Lund et al. 2009). Lund et al (2009), al evaluar diferentes modelos para validar la precisión de predicción de valores genéticos utilizando marcadores moleculares, determinó que: el análisis bayesiano corresponde al enfoque más prometedor para la selección basada en el uso de marcadores moleculares, seguido por los modelos BLUP, mientras que los modelos de efectos fijos generalmente presentaron una baja precisión y una alta varianza del error.

Actualmente diferentes metodologías estadísticas han sido propuestas para evaluar la asociación marcadores-fenotipo, con el propósito de disminuir los errores tipo I y II que surgen en el análisis de datos experimentales (Wei et al. 2006; Stich et al. 2008; Brawner et al. 2012; Lepoittevin et al. 2012). El error tipo I o falsos positivos, surge debido a subdivisiones de la población, referido a efectos de estructura poblacional (Pritchard et al. 2000). La presencia de subgrupos de la población en estudio, podría crear covarianzas entre los individuos, y el no incluir tal efecto en el modelo de asociación, genera sesgos en las estimaciones de los efectos alélicos (Kennedy et al. 1992). En el caso del aumento de la tasa de error tipo II, atribuible según Breseghello y Sorrells (2006), a al menos tres factores: (i) baja correlación entre los marcadores y genes, debido al descenso en el desequilibrio de ligamiento LD, (ii) diseño desequilibrado debido a la presencia de alelos en diferentes frecuencias y (iii) un grave problema de múltiples pruebas; se traduce en falsos negativos debido a la multiplicidad de asociaciones en el genoma. Tales evaluaciones se han implementado bajo el paradigma bayesiano, el cual ha demostrado ser una excelente herramienta de inferencia para la selección de animales y plantas (Meuwissen et al. 2001; Resende et al. 2008). De hecho, su uso se ha propuesto como ventajoso en árboles forestales

tras obtener mayores ganancias genéticas al utilizar esta metodología de selección (Grattapaglia y Resende 2011).

Estructura poblacional y detección de *QTLs*

La estructura genética resulta de la acción conjunta de procesos naturales de migración, mutación, selección y deriva génica, que a su vez operan dentro del contexto histórico y biológico de cada especie (Falush et al. 2007). Cuando hay subdivisión presente en la población, es posible encontrar asociaciones estadísticas entre el fenotipo, por ejemplo de una enfermedad, y marcadores arbitrarios que no tienen ninguna conexión física a los *loci* causales (Lander y Schork 1994). Estas asociaciones se producen debido a la subdivisión poblacional (o cualquier otra forma de apareamiento no aleatorio) que modifica las frecuencias alelo-marcador entre los segmentos de la población, como resultado de la deriva genética o efecto fundador (Slatkin 1991). No obstante, corrección de la estructura poblacional es común en los estudios asociativos realizados en árboles forestales que tienen como objetivo identificar marcadores que contribuyen al fenotipo (González-Martínez et al. 2007; Eckert et al. 2009; Eckert et al. 2010), para ello modelos lineales que incorporan el efecto subpoblacional como fijo (Brescghello y Sorrells 2006) o test de regresión logística (Pritchard et al. 2000; Thornsberry et al. 2001) pueden ser empleados. Al respecto, el programa STRUCTURE, propuesto por Pritchard et al (2000) e implementado por Falush et al (2007) para su uso con marcadores moleculares dominantes, utilizando los procedimientos de Monte Carlo, vía Cadenas de Markov, ha permitido agrupar poblaciones basado en datos de genotipos multilocus. Este programa se ha aplicado a múltiples problemas tales como la identificación de la falsa estructura poblacional, la detección de inmigrantes o personas mezcladas, e inferir la mezcla histórica de la población (Rosenberg et al. 2002; Falush et al. 2003; Albert et al. 2006; Lecis et al. 2006), así como en un sinnúmero de investigaciones relacionadas con asociaciones marcador molecular-fenotipo (Thumma et al. 2005; Lepoittevin et al. 2012).

STRUCTURE asume que todo el material genético de las poblaciones muestreadas proviene de una o más poblaciones no observadas (K). El análisis se basa en un modelo de ascendencia que identifica grupos de poblaciones que presentan frecuencias alélicas distintivas, asignando individuos a cada grupo. Cada población es caracterizada por un conjunto de frecuencias alélicas (P) para cada locus; donde P es un vector multidimensional con elementos P_{kj} que

representan la frecuencia del alelo j en el locus l de la población k (Pritchard et al. 2000; Falush et al. 2007). En un modelo que evalúa la relación entre marcador molecular y *loci* de característica cuantitativa, la estructura poblacional se considerada como un efecto fijo, es decir, su efecto representa una función de ascendencia genética, que es la misma para todas las poblaciones, y asociaciones falsas podrían resultar si se modela como un efecto aleatorio basado en la covarianza general (Price et al. 2010).

Las poblaciones naturales de *E. cladocalyx* se caracterizan por presentar un nivel relativamente bajo de diversidad genética y alto nivel de divergencia (McDonald et al. 2003), característico de poblaciones con distribución natural restringida (Eldridge et al. 1993). Esto puede ser debido a que los altos niveles de consanguinidad presentes en poblaciones naturales de la especie, tienen un marcado efecto sobre la estructura genética poblacional; asimismo la endogamia puede reducir la variación genética y el flujo de genes, lo que resulta en una diferenciación interpoblacional mayor (McDonald et al. 2003). Para el caso de los ensayos de *E. cladocalyx* establecidos en el norte de Chile, se espera encontrar estructuración genética poblacional, ya que el material genético introducido presenta una distribución natural fraccionada (McDonald et al. 2003). Consistente con esto, hay algunas diferencias morfológicas entre las poblaciones naturales. Por ejemplo, se ha visto que poblaciones que producen cápsulas más largas se producen en la Isla Kangaroo, y en menor medida, en la Península Eyre; mientras que las poblaciones de la Cordillera Flinders (región Wirrabara) producen cápsulas relativamente pequeñas. Los árboles en la parte meridional de la cordillera Flinders alcanzan hasta 35 m de altura, mientras que en la Península Eyre por lo general presentan menos de 12 m de altura, en Isla Canguro, son intermedio entre estos dos extremos (Boomsma 1981; Boland et al. 1984; McMahan et al. 2010). De la misma manera, variaciones morfológicas significativas han sido reportadas también en Chile (i.e Florecimiento, crecimiento, curvatura, Mora et al. 2009; Cané-Retamales et al. 2011). De acuerdo a las variaciones que presentan las poblaciones naturales de *E. cladocalyx*, y las diferencias establecidas en los ensayos del norte de Chile, en el presente estudio, se espera encontrar un alto grado de diferenciación (i.e Índice de fijación de Wright ó F_{st}), mayor que lo encontrado en anteriores investigaciones (Bush y Thumma 2013) donde se han utilizado marcadores moleculares SNPs para evaluar árboles de dos de las tres regiones de distribución natural de la especie (Isla Kangaroo y Cordillera Flinders). Este alto grado de diferenciación, sumado a la

amplia variabilidad fenotípica encontrada en las poblaciones de *E. cladocalyx* existentes en Chile permitiría determinar marcadores moleculares ISSR asociados a *loci* de características cuantitativas, que controlan en parte el florecimiento en la especie.

HIPOTESIS

El germoplasma de *Eucalyptus cladocalyx* existente en Chile presenta suficiente variabilidad para el descubrimiento de regiones genómicas asociadas al florecimiento.

OBJETIVOS

Objetivo general

Determinar marcadores moleculares ISSR asociados a *loci* de características cuantitativas (QTLs) que controlan en parte el florecimiento en *Eucalyptus cladocalyx* F. Muell.

Objetivos específicos

1. Examinar la estructura genética en subpoblaciones de *Eucalyptus cladocalyx* F. Muell existentes en Chile.
2. Implementar un modelo bayesiano para el análisis de asociación entre valores genéticos y *loci* de característica cuantitativas (ISSR) que incorpore el efecto de la estructura poblacional de la especie

CAPITULO 1

Variabilidad genética del florecimiento en *Eucalyptus cladocalyx* usando marcadores moleculares e informaciones fenotípicas

Rodrigo Contreras-Soto^{a*}, Freddy Mora^c, Glenda Fuentes^b, Eduardo Ruiz^b, Camila Cané-Retamales^a

^aUniversidad de Concepción, Facultad de Ciencias Forestales, Departamento de Silvicultura, Concepción, Chile.

^bUniversidad de Concepción, Facultad de Ciencias Naturales y Oceanográficas, Departamento de botánica, Concepción, Chile.

^cUniversidad de Talca, Instituto de Biología Vegetal y Biotecnología, Talca, Chile

*Autor de correspondencia: rcontrerass@udec.cl

Abstract

Flowering is a complex but important trait in many tree improvement programmes. To understand the genetic control of this trait, we implemented two Bayesian models, one for evaluated population structure of an open-pollinated progeny trial including five region of provenance (ROP) of *Eucalyptus cladocalyx* F. Muell and for discover loci of inter simple sequence repeat molecular markers associated with breeding values of flowering. Population structure analysis indicates that provenances were clustered in two distinct groups. Bayesian regression identified one loci associated with early flowering and three with flowering intensity, which explain between 8-16% of phenotypic variation. Additionally, one locus was reported how false positive, and their removal reduced spurious genetic association. These results suggest that genetic association approaches may be feasible in *E. cladocalyx*, and that the use of genetic diversity of natural populations could provide a useful source of genes for discovery of loci in traits of importance for perennial species.

Keywords: *Sugar gum, early flowering, Bayesian inference, semi-arid region, Eucalyptus.*

Introducción

Las estrategias de mejoramiento genético, como en la mayoría de las especies forestales, son logísticamente complejas y costosas (Zapata-Valenzuela et al. 2012). En este contexto, la floración ha sido considerada un interesante carácter cuantitativo de estudio, ya que permitiría acortar los intervalos de generación en los programas de mejoramiento genético de eucaliptos (Missiaggia et al. 2005). La principal limitante es que muchas especies leñosas exhiben un prolongado periodo de desarrollo, relacionado a la fase juvenil, periodo en que los árboles no

florece bajo condiciones medioambientales favorables (Jaya et al. 2010). Este largo periodo es un problema para la selección y mejoramiento genético de especies con importancia forestal y hortícola. En eucalipto la edad de inicio de floración (Chambers et al. 1997; Jordan et al. 1999), inicio de reproducción (McGowen et al. 2004), y temporada de floración (Gore y Potts 1995; Apiolaza et al. 2001; McGowen 2007) son genéticamente variables tanto entre como dentro de genotipos, adicionalmente el número de loci involucrados en el tiempo de floración ha mostrado depender tanto del genotipo como del medioambiente en que crecen las plantas (Clarke et al. 1995).

Algunas especies de climas templados, tales como *Eucalyptus globulus*, *E. nitens*, *E. regnans* y *E. dunnii* no florecen hasta por un periodo de entre 7 a 10 años (Eldridge et al. 1993). En Chile, por ejemplo, *E. globulus* ssp *globulus*, *E. nitens* y *E. dunnii* presentan un extenso periodo juvenil y no florecen hasta que los árboles tienen 10 años de edad (Potts et al. 2001). Este extenso periodo juvenil descrito para diferentes especies de eucalipto es una limitante que impide acelerar evaluaciones de progenie en programas de mejora, especialmente cuando es necesario adquirir semilla genéticamente superior (Jordan et al. 1999). En muchos programas de mejoramiento genético, el paclobutrazol es rutinariamente aplicado para acelerar la floración (Griffin et al. 1993) y reducir los tiempos entre generaciones (Hasan y Reid 1995). Sin embargo esta hormona es resistente a degradación química y biótica (Jackson et al. 1996), y persiste en el suelo por muchos años (Griffin et al. 1993), por cuanto es una preocupación los efectos que a largo plazo genera este componente en el crecimiento de las plantas. Por esta razón, la selección de árboles que florecen precoz e intensamente es interesante para mejorar la producción de semilla mejorada en los huertos semilleros, en producción de miel, en el desarrollo de híbridos de eucalipto y en la producción de semilla genéticamente mejorada para plantaciones.

Con el advenimiento de los marcadores moleculares, diversos mapas de ligación han sido construidos y utilizados para detectar loci que contribuyen a la varianza de caracteres cuantitativos, también conocidos como loci de características cuantitativas o QTL (Bundock et al. 2008). En eucalipto, un número considerable de investigaciones basados en el uso de marcadores que incluyen: RFLP, RAPD, AFLP y microsatélites (SSR) han informado de mapas de ligamiento y detección de QTL para caracteres complejos tales como: crecimiento,

calidad de madera, enfermedades y resistencia a insectos (Grattapaglia y Sederoff 1994; Verhaegen et al. 1997, Marques et al. 1999; Myburg et al. 2003). Marcadores de tipo ISSR (*Inter Simple Sequence Repeats*), aunque individualmente menos informativos que marcadores altamente multi-alélicos como SSR, otorgan información multilocus útil, ya que han mostrado ser reproducibles, abundantes y polimórficos en genomas de plantas (Bornet y Branchard 2001). Sus partidores están constituidos con un extremo anclado por pocas unidades de SSR, para evitar la hibridación no específica a la región microsatélite (Zietkiewicz et al. 1994). Estos marcadores han sido importantes para la estimación de la diversidad genética en diferentes especies de árboles, como cítricos (Fang y Roose 1997), olivos (Gemas et al. 2004a), casuarina (Yasodha et al. 2004; Kamalakannan et al. 2006) y eucaliptos (Gemas et al. 2004b; Balasaravanan et al. 2005; Chezhan et al. 2010). Adicionalmente, hasta ahora, algunos estudios de asociación han adoptado el uso de ISSR asociados a caracteres cuantitativos (rendimiento (Vijayan et al. 2006), resistencia (Ruan et al. 2009), contenido de proteína (Kar et al. 2008).

Los estudios de QTL dependen del desarrollo de poblaciones segregantes (es decir, F₂, retrocruzamientos o líneas isogénicas), que constituyen una limitante para especies que presentan altos niveles de depresión por endogamia y largos tiempos de generación, tal como ocurre en árboles forestales. Investigaciones recientes indican que no existen reportes sobre el uso real de QTL en los programas de mejoramiento forestales asistidos por marcadores (Resende et al. 2012; Grattapaglia y Resende 2011), las causas tal vez se relacionan con el problema de que la mayoría de los QTL han sido ampliamente sobreestimados (Grattapaglia et al. 1996; Shepherd et al. 1999). Por esta razón, diferentes metodologías se han explorado para estudiar caracteres cuantitativos, teniendo en cuenta, además, que la eficacia de análisis de ligamiento y la construcción de mapas genéticos han sido limitados en muchos programas de mejoramiento forestales debido a la rápida descomposición del desequilibrio de ligamiento, que se observa normalmente en poblaciones domesticadas con sistemas de cría exogámica (Neale y Savolainen 2004).

Los estudios de asociación se han vuelto cada vez más comunes en especies leñosas (Brawner et al. 2012; Lepoittevin et al. 2012), teniendo en cuenta que los marcadores moleculares permiten detectar asociaciones entre la variabilidad fenotípica de muchos individuos con

diferentes regiones del genoma, y que pueden ser aplicados a nivel poblacional (Cumbie et al. 2011). Esta integración de la información molecular para caracteres de interés en los programas de mejoramiento, proporciona una conexión entre los marcadores moleculares y datos fenotípicos que históricamente han sido recopilados para el mejoramiento genético (Brawner et al. 2012). En general, existen dos métodos para evaluar la significancia de la relación entre los marcadores genéticos y rasgos fenotípicos: uno se basa en el concepto de distancia física o análisis por ligamiento, que ha demostrado ser ineficaz en árboles, y la otra implica el uso de modelos probabilísticos para el análisis de asociación (Brawner et al. 2012). Cuando se utiliza un gran número de marcadores distribuidos equitativamente (no hay genes candidatos) en el análisis de asociación, usualmente se aplican métodos estadísticos de regresión contraída (Ridge) (Endelman 2011) o regresión bayesiana. En el enfoque bayesiano, variantes del procedimiento de Cadenas de Markov (MCMC) Monte-Carlo como el algoritmo de muestreo de Gibbs han demostrado ser apropiados para el análisis de QTL (Xu y Hu, 2011). Recientemente, Hu y Xu 2009 publicaron un procedimiento para el análisis de QTL, donde destacan la ventaja del análisis bayesiano con respecto al método clásico de máxima verosimilitud para la detección simultánea de múltiples QTL dentro de un mismo modelo.

Eucalyptus cladocalyx F. Muell, es un árbol endémico del sur de Australia, localizado en tres regiones geográficas disjuntas en un rango total de 500 km, dos de ellas están situadas en la parte continental (Flinders Range y Península de Eyre) y una en la isla Kangaroo (Bush et al. 2011). La especie ha sido considerada en los programas de mejoramiento debido a su aptitud para crecer en sitios con escasa precipitación, donde podrían ser plantados para la producción de madera y postes, así como para la remediación de suelos (Hardwood et al. 2007). En el sur del desierto de Atacama, Región de Coquimbo, Chile, también está siendo genéticamente evaluada para la producción de polines y miel (Mora et al. 2009), teniendo en cuenta que sus flores son una fuente de néctar que complementa la errática floración de especies endémicas y nativas de esa zona (Montenegro et al. 2003). Adicionalmente es preferida por sobre otras especies debido a su moderada tolerancia a salinidad (Marcar et al. 1995) y su capacidad para crecer en ambientes áridos y semiáridos (Gleadow y Woodrow 2000; Woodrow et al. 2002; McDonald et al. 2003).

McDonald et al. (2003), haciendo uso de marcadores isoenzimáticos ha encontrado que los genotipos silvestres de *E. cladocalyx* son menos alógamos y más puros que otras especies de *Eucalyptus*. Este resultado fue confirmado recientemente por Bush y Thumma (2013), quienes encontraron elevados niveles de endogamia, particularmente en la subpoblación de la isla Kangaroo. A pesar de que *E. cladocalyx* presenta una distribución natural relativamente restringida y alta endogamia, se han observado altos niveles de diferenciación genética entre las poblaciones naturales y plantaciones para caracteres tales como crecimiento, supervivencia, floración precoz (Mora et al. 2009), y características de durabilidad natural (Bush et al. 2011), lo que sugiere un amplio margen para el mejoramiento genético. Caracteres como floración precoz e intensidad de floración, por ejemplo, han mostrado diferencias significativas entre procedencias establecidas en Chile (Mora et al. 2007, 2009; Contreras-Soto et al. 2011). La magnitud de diferencia para floración precoz en árboles de 30 meses de edad, teniendo en cuenta la mejor (Mt. Remarkable) y peor (Flinders Chase) procedencia fue de 82% (Mora et al. 2009), y según Contreras-Soto et al. (2011), esta diferencia fue consistente al evaluar el mismo ensayo 6 años más tarde, debido a que la intensidad de floración mostró diferencias altamente contrastantes entre las procedencias de Cowell y Flinders Chase (diferencia de 247%). Consecuentemente, la significativa correlación genética entre intensidad de floración y floración precoz ($\Gamma = 0,45$) encontrada en la especie, sugiere que la selección genética basada en la intensidad de floración tendría un impacto positivo y significativo en floración temprana (y vice versa) (Cané-Retamales et al. 2011).

De acuerdo a variaciones naturales reportadas en las poblaciones de *E. cladocalyx*, se espera que en el presente estudio se encuentre un alto grado de diferenciación genética (Índice de fijación de Wright o F_{st}), superior a lo encontrado por Bush y Thumma (2013), por ejemplo, quienes evaluaron el uso de marcadores basado en el polimorfismo de un solo nucleótido (SNP) para analizar los programas de mejoramiento genético de *E. cladocalyx* F. Muell en dos de las tres regiones de distribución natural de la especie (Isla Kangaroo y Flinders Ranges). Este alto grado de diferenciación, junto con la amplia variabilidad fenotípica encontrada para caracteres de floración en poblaciones de *E. cladocalyx* establecidas en Chile, ayudaría a identificar marcadores moleculares asociados a loci de características cuantitativas que controlan la floración. Por lo tanto, el objetivo del presente estudio fue identificar marcadores

moleculares ISSR asociados a loci de características cuantitativas (QTL) que controlan en parte la floración de *Eucalyptus cladocalyx*.

Materiales y Métodos

Ensayo de progenie y evaluación fenotípica

Un ensayo de progenie de polinización abierta se utilizó para la evaluación de la asociación genética entre marcadores moleculares y la floración precoz (EF) e intensidad de floración (FI) en procedencias de *Eucalyptus cladocalyx* sobre la base de los valores genéticos para ambos caracteres. El estudio se localiza en el sur del desierto de Atacama, en la Región de Coquimbo, provincia de Choapa (31°38' S, 71°19' O; 297m de altitud). Las semillas se obtuvieron en el año 2000, desde la colección base del centro de semillas CSIRO (Commonwealth Scientific and Industrial Research Organization), Australia, y las subpoblaciones seleccionadas son representativas del área de distribución natural de *E. cladocalyx*. El ensayo se estableció bajo los supuestos de un diseño de bloques completos al azar, con 30 bloques representados en un esquema de un árbol por parcela con 44 familias de medios hermanos que representan cinco procedencias de la distribución natural de la especie (Tabla 1). Los árboles se plantaron cada 2m X 3m entre cada fila, dando una densidad aproximada de 1.667 árboles ha⁻¹.

Tabla 1: Procedencias Australianas y Chilenas de *Eucalyptus cladocalyx* establecidas en el norte de Chile: Distribución natural de origen, número de familias representadas y precipitación media anual.

Origen	Familias (N°)	Latitud Sur	Longitud Oeste	Precipitación (mm·año⁻¹)
Flinders Chase NP (Kangaroo Island)	8	35°57'	136°42'	637.9
Marble Range (Eyre Peninsula)	4	34°30'	135°30'	485.1
Mt. Remarkable (Flinders Ranges)	16	32°43'	138°06'	242.8
Cowell (Eyre Peninsula)	10	33°38'	136°40'	405.0
Wirrabara St. Forest (Flinders Ranges)	9	33°06'	138°14'	256.6

La floración precoz se determinó mediante la evaluación de árboles de 30 meses de edad de acuerdo a la presencia (1) o ausencia (0) de cápsulas y/o brotes de flores en junio de 2004 (Mora et al. 2007, 2009). Otros estudios de precocidad en eucalipto han evaluado la floración precoz de forma similar, como una respuesta binaria (Chambers et al. 1997).

La estrategia de evaluación para la intensidad de floración incluyó una escala de graduación medido al comienzo de la estación de crecimiento 2010/2011, en concordancia con Cané-Retamales et al. (2011), similar a Keatley y Hudson (2007) y Hudson et al. (2009). La floración se cuantificó utilizando una escala de 4 categorías: 0 ausencia de flores; 1: menos del 50% de cápsulas, yemas y/o flores a lo largo del fuste del árbol; 2: más del 50% de cápsulas, yemas y/o flores a lo largo del fuste del árbol, y 3 para el caso de numerosas cápsulas, yemas y/o flores (Cané-Retamales et al. 2011). La puntuación de la intensidad de floración asume igual ponderación (importancia) de sus componentes (cantidad y distribución).

Datos genotípicos

ADN genómico de las 44 familias de *E. cladocalyx*, fue extraído a partir de tejido foliar juvenil fresco o congelado utilizando el protocolo de extracción CTAB, tal como lo describe Doyle y Doyle (1987), con modificaciones menores. La integridad y concentración del ADN se visualizó en geles de agarosa al 0,8%, y se comparó con un estándar de marcador Promega ® 50-8000 kb lambda (Promega, Madison, WI, EE.UU).

En el presente estudio se realizaron amplificaciones ISSR-PCR en un volumen de reacción de 20µl que contiene 2µL de ADN, 1X GoTaq reacción de PCR Buffer (Promega, Madison, WI, EE.UU), 2,5 mM de MgCl₂, 0,4 mM de dNTPs, 100 nM de cada partidor y 0.3U Taq ADN polimerasa (Promega, Madison, WI, EE.UU). Se utilizaron ocho partidores (Tabla 2), que se describieron previamente en otras especies de eucaliptos (Balasavaraban et al. 2005; Okun et al. 2008; Chezhan et al. 2010). En *E. cladocalyx* estos partidores se utilizaron recientemente para analizar la diversidad genética en 49 familias de medios hermanos plantados en el norte de Chile (Ballesta et al. 2013). Para la amplificación por PCR, dos programas diferentes fueron diseñados, y para cada partidor se utilizaron diferentes temperaturas de alineamiento (Tabla 2). Para los partidores (RCA)₇, T(GT)₉, TA(CAG)₄, RA(GCT)₆, GA₈, y UBC810 se utilizó el siguiente programa de amplificación de PCR: denaturación inicial (3 min, 94°C), seguido por 35 ciclos que consiste en la denaturación (30s, 94°C), alineamiento (30s),

extensión (1min, 72°C) y una extensión final (10min, 72°C). En el caso de ISO1 e ISO2 el termociclador se programó como sigue: denaturación inicial (7min, 94°C), seguido por 45 ciclos de denaturación (30s, 94°C), alineamiento (45s), extensión (2min, 72°C) y una extensión final (7min, 72°C). Los productos de amplificación se resolvieron en gel de agarosa al 2 % con tampón TAE 1X a 70V durante 1h con marcador de tamaño de 1 kb (Axygen®). Los productos de PCR se visualizaron en un transiluminador UV Vilbert Loumart después de la tinción con bromuro de etidio, y se fotografiaron con cámara Nikon Coopix 4300.

Tabla 2: Secuencias de partidores inter-microsatelites (ISSR) utilizados en el presente estudio.

Nombre	Secuencia partidor
¹ R(CA)7	5'-GATACTGATACACACACACACA-3'
¹ T(GT)9	5'-CATATGTGTGTGTGTGTGTGTGTGT-3'
¹ TA(CAG)4	5'-AGATTCAGCAGCAGCAG-3 • '
¹ RA(GCT)6	5'-ATAAAGCTGCTGCTGCTGCTGCT-3'
¹ (GA)8 R	5'-GAGAGAGAGAGAGAGAAGT-3 • '
¹ UBC810	5 • '-GAGAGAGAGAGAGAGAT-3 • '
² ISO1	5'-CACACACACACAGG-3'
² ISO2	5'-CTCTCTCTCTCTCTAAC-3'

Referencia partidores previamente utilizados en diferentes especies de *Eucalyptus*. 1: *E. tereticornis* (Chezhian et al. 2010); *E. camaldulensis*, *E. citriodora*, *E. grandis*, *E. pellita*, *E. tereticornis* y *E. urophylla* (Balasaravanan et al. 2005); *E. camaldulensis*, *E. citriodora*, *E. grandis*, *E. pellita*, *E. tereticornis* y *E. urophylla* (Balasaravanan et al. 2006); 2 *E. grandis* (Okun et al. 2008).

Análisis de datos moleculares y estructura poblacional

Para el análisis de las bandas, a cada amplicón ISSR se le otorgó un valor 1 en caso de presencia y 0 en caso de ausencia. Debido a que los marcadores ISSR son dominantes, la varianza en un solo *loci* se incrementa y sesga las medidas de frecuencias alélicas (Lynch y Milligan 1994; Zhivotovsky 1999), por lo tanto, los análisis se desarrollaron de acuerdo al método de Lynch y Milligan (1994), después de estimar las frecuencias de los alelos responsables de la presencia o ausencia de bandas. Análisis de varianza molecular (AMOVA)

se realizó utilizando Arlequin, V3.5 (Excoffier et al. 2010) para estimar la variación existente dentro de cada población y entre las poblaciones. Cabe destacar que para obtener tal resultado, se eliminaron los *outlier* de acuerdo a Excoffier et al (2009).

La detección de la estructura poblacional se examinó con el programa STRUCTURE V2.2 (Falush et al. 2007), utilizando la matriz de marcadores moleculares que no incluye los *outliers*. El análisis se basó en un modelo bayesiano implementado por medio del algoritmo de *Cadenas de Markov Monte Carlo* (MCMC), que de acuerdo al genotipo multilocus de cada población, se asignaron en *clusters* (K, el cual puede ser desconocido) sobre la base de una matriz de pertenencias (Matriz Q) generada por el programa. Como el valor K puede variar en función de eventos de estructuración individual, se determinó un valor óptimo de K según lo propuesto por Evanno et al (2005).

El modelo poblacional que implementará el algoritmo para el análisis de la matriz Q correspondió a un modelo admixture con frecuencia de alelos independiente, propuesto por Pritchard et al (2000), cuya base constituye el análisis de genotipos multilocus (Falush et al. 2007). Este modelo supone que las k poblaciones ancestrales desconocidas se encuentran en equilibrio de Hardy-Weinberg y en equilibrio de ligamiento (Pritchard et al. 2000).

La convergencia de las Cadenas de Markov, que corresponden a los Fst (Índice de diferenciación) *a posteriori* se analizaron a partir del test Heidelberger y Welch (1983) del programa R implementando en la rutina CODA (*Convergence Diagnosis and Output Analysis*).

Análisis de modelo bayesiano

Los valores genéticos para los caracteres de floración (EF y FI) fueron predichos para todos los árboles (N=1320) por medio de un enfoque bayesiano utilizando el programa MTGSAM (*Multi Trait Gibbs Sampling for Animal Model*) propuesto por Van-Tassell y Van Vleck (1996). Este programa resuelve ecuaciones del modelo mixto del Henderson (2) (Henderson 1986) e implementa un análisis bayesiano basado en el algoritmo de muestreo de Gibbs. La notación del modelo lineal mixto utilizado por Van Tassel y Van Vleck (1996) se utiliza para describir los supuestos del modelo y los resultados. El siguiente modelo umbral representa los datos de floración para predecir los valores genéticos:

$$\eta = X\beta + Za + \varepsilon \quad (1)$$

Donde, η es el vector de la variable aleatoria continua, normalmente distribuida, no observable (desconocida) y referida al carácter de interés (EF y FI); β es el vector de los efectos de procedencia y bloque; a es el vector del efecto aditivo de los árboles; X y Z son las matrices de incidencia de β y a respectivamente, mientras que ε es el vector de efecto residual.

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}Z' \\ Z'R^{-1}X & Z'R^{-1}Z + G^{-1} \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{u} \end{bmatrix} = \begin{bmatrix} X'R^{-1}y \\ Z'R^{-1}y \end{bmatrix} \quad (2)$$

Los efectos debidos a procedencia y bloque se asumieron con una distribución a priori plana (uniforme), el efecto genético aditivo se asumió con una distribución normal, mientras que los componentes de varianza se asumieron con distribución invertida Chi-cuadrado, según lo informado por Mora et al. (2009) para *E. cladocalyx*.

Para generar las MCMC, se han ejecutado 250.000 iteraciones correspondientes a cadenas de Gibbs utilizadas para el muestreo de la distribución marginal *a posteriori* de los parámetros de interés. Las primeras 50.000 iteraciones fueron descartadas como burn-in. La convergencia de las varianzas producidas por las iteraciones MCMC se han probado utilizando el test Heidelberger y Welch (1983) implementado en CODA (Diagnóstico y Análisis de Convergencia de salida) del programa R (Plummer et. Al 2006).

Los valores genéticos predichos por el modelo Bayesiano (1) se usaron en un modelo de asociación Bayesiano, valor genético-marcador:

$$y_{ijk} = X_j + Z_k + \varepsilon_{ijk} \quad (3)$$

Donde y_{ijk} corresponde al i -ésimo valor genético con $y_{ijk} \sim N(0; \sigma^2)$; X_j es el efecto del j -ésimo marcador molecular (ISSR); Z_k es el efecto de la k -ésima subpoblación determinada con el programa STRUCTURE, mientras que ε_{ijk} corresponde al efecto residual de la ijk -ésima asociación que depende de la k -ésima subpoblación y del j -ésimo ISSR con $\varepsilon \sim N(0;$

σ_e^2). X y Z son las matrices de incidencia de los efectos marcador molecular y estructura poblacional, respectivamente.

Para evaluar la asociación entre los caracteres de floración (EF y FI) con los marcadores moleculares ISSR, se utilizó el procedimiento GENMOD con la opción BAYES, ya que tiene la capacidad de implementar un análisis bayesiano a través del muestreo de Gibbs (Xu y Hu, 2011).

El porcentaje de varianza fenotípica explicado por cada loci ISSR fue calculado a través de un modelo de regresión lineal entre el fenotipo y los ISSR, utilizando el procedimiento GLM implementado en SAS.

Resultados y discusión

Variabilidad de la floración

En la presente investigación, las procedencias evidenciaron diferencias estadísticas significativas para la intensidad de floración (Tabla 3). Este resultado es importante ya que la mejora genética de especies forestales, comienza normalmente con la evaluación fenotípica de ensayos de procedencia-progenie para diversos caracteres de interés, como el caso de *E. cladocalyx* en Chile (Mora et al. 2009; Bush y Thumma 2013). Las procedencias de Cowell y Wirrabara no mostraron diferencias significativas, sin embargo se diferenciaron significativamente de Flinders, una población que ha mostrado bajos niveles de floración en ésta y otras investigaciones. Este resultado confirma lo encontrado por Contreras-Soto et al. (2011) y Cané-Retamales et al. (2011) quienes reportaron que el porcentaje de intensidad de floración fue significativamente diferente entre procedencias de la especie. La variabilidad del comportamiento de la floración en distintas especies del género *Eucalyptus* ha sido mencionada desde el año 1893 (Chambers 1893), y en todas las especies evaluadas se destaca la alta variabilidad de la característica (Birtchnell y Gibson 2008). Esta variabilidad se ha reportado a distintos niveles: variación debida a región geográfica, especie y procedencias (Chambers et al 1997; Butcher et al 2009; Jones et al 2011). En un estudio realizado en Australia, por ejemplo, donde se evaluó la intensidad de floración en 13 especies de eucaliptos, se encontró una variabilidad significativa entre especies y sitios, con resultados que fueron desde ocurrencia de floración anual a ninguna floración durante un período de 10 años

(Law et al. 2001). Estos resultados son coincidentes con el presente estudio, donde la floración inexistente, en un periodo de 10 años, varió de 14 a 50%, para las procedencias de Cowell y Flinders, respectivamente.

En el caso de *E. cladocalyx* se han encontrado diferencias a nivel de población y región de procedencia (Bush et al. 2011). En el presente estudio se encontraron diferencias fenotípicas contrastantes entre individuos de la procedencia de isla Kangaroo (Flinders) y península Eyre (Cowell y Marble) (Tabla 3). Por ejemplo, Cowell en la clase 3 de floración presenta 45%, mientras que Flinders tan solo 7%. En la clase 0 sucede lo contrario, Cowell presenta 14% y Flinders 50%. Evaluaciones de la floración a lo largo del tiempo han sido comunes en diversas especies del género *Eucalyptus* (Law et al. 2001; Birtchnell y Gibson 2006; Jones et al. 2011). En *E. globulus*, por ejemplo, Jones et al. (2011) monitoreó durante 5 años la iniciación floral en 63 clones que representan cuatro regiones de la distribución natural de la especie, y encontró diferencias significativas entre regiones cada año. Este resultado es similar a lo reportado en *E. cladocalyx* para las diferentes clases de floración medidas en los periodos 2009-2012, y aunque los resultados mostraron leves diferencias cada año, al igual que lo reportado por Jones et al. (2011), el ranking a nivel de procedencia fue similar en los diferentes años.

La alta variabilidad encontrada en los distintos niveles de intensidad de floración, corrobora el hecho de que la especie ha sido insuficientemente domesticada, y esto representa una oportunidad para el estudio de genes candidatos que otorgan variación genética natural en los distintos genotipos de la especie. Esta variación, ha sido informada en varios caracteres (crecimiento y durabilidad de la madera-Bush et al. 2011; altura y diámetro-Mora et al. 2009) y ello resulta relevante para considerar el establecimiento de ensayos de progenie con otras procedencias, a fin de evaluar otros rasgos no estudiados en *E. cladocalyx*, tales como: tolerancia a distintos tipos de estrés como sequía, sales y metales pesados o la resistencia a patógenos (*Phoracantha semipunctata*). Esto sobre la base, de que estudios anteriores destacan su moderada tolerancia a salinidad (Marcar et al. 1995), capacidad para crecer en zonas áridas (Woodrow et al. 2002; Mora et al. 2009) y resistencia a patógenos (Hanks et al. 1994).

Tabla 3: Medias poblacionales para la intensidad de floración medida a los 10 años de edad en 44 familias de medios hermanos de *Eucalyptus cladocalyx* establecidos en un ensayo en el Sur del desierto de Atacama.

Poblaciones	Arboles* evaluados	Familias evaluadas	% de floración en clases				Diferencias significativas floración
			0	1	2	3	
Cowell	300	10	14	16	10	45	1.69 a
Wirrabara	270	8	21	19	11	44	1.75 a
Mt. Remarkable	420	14	23	20	12	35	1.50 ab
Marble Range	120	4	26	32	8	21	1.09 b
Flinders Chase	240	8	50	21	8	7	0.58 c

* 120 Árboles muertos. La suma de los arboles evaluados fue 1.470.

Diferenciación genética y estructura poblacional

En el presente estudio, se encontraron altos niveles de diferenciación genética entre procedencias ($F_{st}= 0,38$; Tabla 4). Según AMOVA, esta diferenciación se dividió en 38,47% entre las procedencias y el 61,53% entre los individuos dentro de las procedencias. Estos valores fueron mayores a los reportados por Bush y Thumma (2013), y McDonald et al. (2003). Los resultados del presente estudio fueron mayormente similares a los obtenidos por McDonald et al. (2003). Estos últimos consideraron poblaciones de todo el rango de distribución natural de la especie, por lo tanto esto indica que las poblaciones de *E. cladocalyx* establecidas en Chile son una buena representación de la variabilidad presente en Australia.

Tabla 4: Análisis de varianza molecular (AMOVA) implementada en Arlequin.

Fuente de variación	DF	Suma de cuadrados	Componentes de varianza	Porcentaje de variación
Dentro Poblaciones	4	315,854	7,8183	38.47
Entre Poblaciones	39	487,782	12,5072	61.53
Total	43	803,636	20,3256	

Índice de fijación de Wright $F_{st}= 0,3847$

Ciento cuarenta y un locus ISSR fueron utilizados en el análisis de asociación marcador-fenotipo, de los cuales 45 fueron discriminados del análisis de estructura genética debido a que no resultaron ser loci neutrales. De acuerdo con el criterio de Evanno et al. (2005), el mejor

valor de K fue dos (Figura 1). Las cinco procedencias (Cowell, Marble Range, Mt Remarkable, Flinders y Wirrabara) fueron agrupadas en dos diferentes clusters. Las procedencias Cowell y Marble Range fueron agrupadas en un mismo grupo (65% y 95% de pertenencia, respectivamente; Figura 2), mientras que Mt. Remarkable, Wirrabara y Flinders Chase fueron asignadas a un grupo diferente (98%, 76% y 99% de pertenencia, respectivamente; Figura 2). Los resultados de diferenciación genética encontrados en el presente estudio son coherentes con los obtenidos por McDonald et al (2003), y Bush y Thumma (2013). El análisis de estructura genética entregó evidencias de la separación entre EP y el resto de las ROP. Anteriormente, McDonald et al. (2003) determinaron tres grupos distintos para agrupar a las procedencias, y recientemente Bush y Thumma (2013) encontraron que las poblaciones segregan en dos regiones de procedencia (KI y SFR). La incorporación de EP en nuestro análisis pudo generar diferencias en los resultados, debido a que estudios previos indican que *E. cladocalyx* se constituye de dos grupos principales (subpoblaciones), los cuales corresponden a SFR y KI.

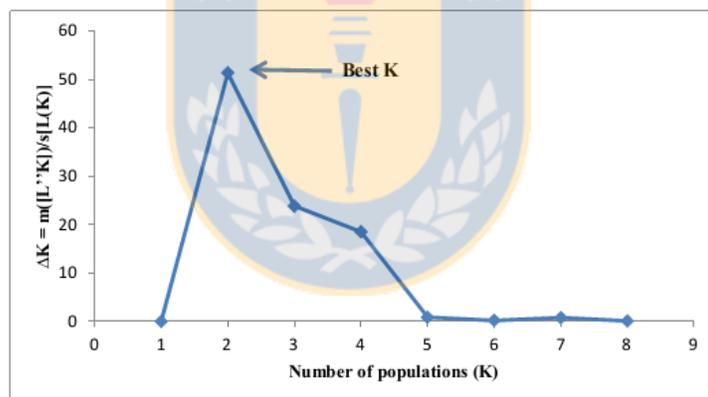


Figura 1: K estimado de acuerdo al logaritmo de máxima verosimilitud de ΔK .

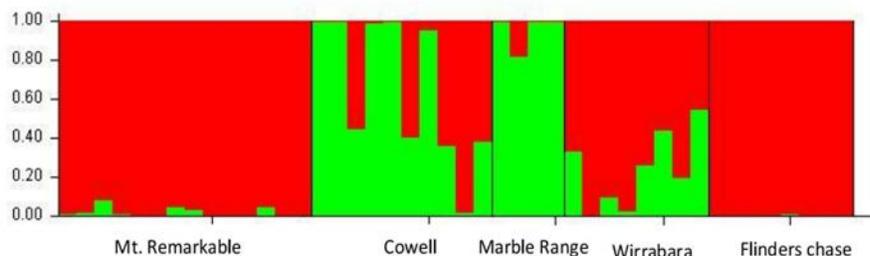


Figura 2: Estructura poblacional de 5 procedencias de *E. cladocalyx* evaluada con 96 locus ISSR con valor $K=2$. Color verde representa el porcentaje de pertenencia en el cluster 1 de las poblaciones Cowell y Marble Range, y el color rojo que representa al cluster 2 con las poblaciones Mt. Remarkable, Wirrabara y Flinders.

Asociación bayesiana entre marcadores ISSR y floración

El tiempo de floración es una característica fuertemente influenciada por el efecto de la estructura poblacional (Zhu et al. 2008), la cual aumenta la probabilidad de obtener falsas asociaciones marcador-fenotipo. En el presente estudio, se implementaron dos modelos bayesianos para detectar falsas asociaciones fenotipo-genotipo debidas al efecto de estructura genética. El primero, consideró el efecto de la estructura genética como un factor que genera variación en el fenotipo de los individuos, y el segundo sólo consideró a los loci como causales de variación de las características fenotípicas. Los resultados revelaron la presencia de un locus falso positivo para la asociación marcador y florecimiento precoz. Thornsberry et al. (2001) reportaron una significativa asociación entre polimorfismos del gen Dwarf8 y la variación en el tiempo de floración en maíz. En estos resultados, los autores encontraron hasta un 80% de falsas asociaciones debido al efecto de estructura poblacional. De acuerdo con Zhao et al. (2007), los análisis basados en métodos de regresión lineal estándar que no corrigen el efecto de estructura poblacional, frecuentemente identifican asociaciones significativas.

Mediante el modelo bayesiano de asociación, que consideró el efecto de la estructura poblacional, se identificaron loci ISSR asociados a los valores genéticos de EF y FI. Un locus fue significativamente asociado a EF, el cual explicó un 17% de la variación fenotípica del carácter, mientras que tres loci explicaron entre un 8-16% de la variación fenotípica de FI (Tabla 5). La heredabilidad de EF y FI en *E. cladocalyx* se ha considerado de moderada a alta (Mora et al. 2009; Cané-Retamales et al. 2011). De acuerdo a Thumma et al. (2010), el número de QTLs detectados puede estar directamente relacionado con la heredabilidad del carácter. En adición, la heredabilidad y el porcentaje de variación fenotípica debida a un QTL parecen estar altamente relacionadas en especies endogámicas. Unos pocos QTLs explican una alto porcentaje de la variación fenotípica del tiempo de floración en arabidopsis y arroz (El-Lithy et al. 2006; Izawa et al. 2003). McDonald et al (2003) han determinado altos índices de

endogamia en *E. cladocalyx*, lo cual puede explicar los importantes efectos que poseen los loci sobre las características de EF y FI.

Tabla 5: Marcadores moleculares ISSR asociados con floración precoz e intensidad de floración, de acuerdo a un modelo bayesiano con un 99% de intervalo de credibilidad.

Floración precoz					Intensidad de floración				
ISSR	TM*(pb)	Intervalo credibilidad 99%		R ²	ISSR	TM(pb)	Intervalo credibilidad 99%		R ²
		Inferior	superior				Inferior	Superior	
RCA7	890	0,021	0,269	0,174	ISO1	500	0,026	0,489	0,152
					GA8	950	0,015	0,563	0,081
					RCA7	1110	0,032	0,482	0,159

*TM: Tamaño de fragmento

Previamente, marcadores SSR han sido relacionados a floración precoz (EF), tales como EMBRA02 en *E. globulus* (Thamarus et al. 2002; Southerton et al. 1998; Farro et al. 2013), y EMBRA27 y EMBRA164 en *E. grandis* (Missiaggia et al. 2005). Missiaggia et al. (2005) identificaron un QTL asociado a floración precoz (Eef1) en mutantes de *E. grandis*. No obstante, el marcador EMBRA27 (asociado al locus Eef1) no mostró estar ligado al locus PFEg (EMBRA02), el cual previamente había sido reportado por Farro et al (2013) para florecimiento precoz en *E. grandis*. Esto sugiere la existencia de múltiples regiones genómicas relacionadas al inicio de floración en *Eucalyptus*. Entre los loci asociados a EF y FI encontrados en el presente estudio, destacan los amplificadas a partir de R(CA)7 y (GA)8R. La combinación de pares de partidores ISSR han permitido identificar dos marcadores ligados a un locus que controla la temporada de floración (Seasonal Flowering Locus, SFL) en *Fragaria vesca* (Cekic et al. 2001). Adicionalmente, Albani et al (2004) desarrollaron marcadores SCAR (Sequence-Characterized Amplified Region SCAR) ligados a SFL a través de la combinación de pares de partidores ISSR en *F. vesca*. Los genotipos que llevan el alelo dominante del gen SFL muestran un período restringido de la floración, mientras que los genotipos que son homocigotos para el alelo recesivo florecen continuamente durante todo el año (Albani et al. 2004). En *E. cladocalyx*, la procedencia Cowell no ha presentado periodos restringidos de floración en al menos cinco años, por lo tanto podría ser mayormente frecuente un alelo no-restringido de SFL en esta procedencia, al contrario de árboles de Flinders Chase, los cuales presentan escasa floración en el mismo periodo (Mora et al. 2009; Cané-Retamales

et al. 2011; Contreras-Soto et al. 2011). Aunque la fenología ha sido ampliamente estudiada en *Eucalyptus*, los mecanismos moleculares asociados al gen SFL han sido escasamente descritos. Esto se debe a la factores como la prolongada juvenilidad en especies perennes, y la complejidad de la genética de la floración (Albani et al. 2004).

De acuerdo a nuestros resultados, el uso de partidores ISSR podría ser una ventajosa metodología para localizar regiones genómicas asociadas a floración en *E. cladocalyx*. La detección de loci asociados a floración en *E. cladocalyx*, es importante para los programas de mejoramiento genético en la especie, ya que la selección de individuos que florecen en forma precoz y abundante permite desarrollar progenies elite en menor tiempo (Farro et al. 2013). Por otro lado, en el norte de Chile estos resultados adquieren mayor relevancia, debido a que *E. cladocalyx* florece entre diciembre y abril, el cual es un período crítico desde el punto de vista floral. Otras especies exhiben escasos niveles de floración en este periodo debido a la escasez de precipitaciones. La selección previa asistida por marcadores de genotipos que florecen en forma precoz e intensa, puede contribuir para fortalecer la producción de miel de *Apis mellifera*, la cual es un recurso que genera ingresos económicos extras en plantaciones en condiciones de déficit hídrico.

CONCLUSIONES

Las regiones amplificadas por los partidores ISSR agruparon las distintas procedencias de *E. cladocalyx* de manera consistente con estudios previos. En adición, el alto grado de diferenciación entre procedencias encontrado puede estar asociado con la fuerte estructuración genética que presenta la especie. El análisis de asociación bayesiana implementado confirma el efecto de la estructura genética en la detección de loci relacionados a características de floración, y revela la existencia de múltiples regiones genómicas asociadas a floración en *Eucalyptus*, tal como se ha descrito previamente en otros estudios. Los caracteres de FI y EF fueron asociados a pocos loci, lo cual es esperado en caracteres de heredabilidad relativamente alta.

REFERENCIAS

- Albani MC, Battey NH, Wilkinson MJ (2004) The development of ISSR-derived SCAR markers around the SEASONAL FLOWERING LOCUS (SFL) in *Fragaria vesca*. *Theoretical and Applied Genetics* 109: 571–579.
- Albert V, Jonsson B, Bernatchez L (2006) Natural hybrids in Atlantic eels (*Anguilla anguilla*, *A. rostrata*): evidence for successful reproduction and fluctuating abundance in space and time. *Molecular Ecology* 15:1903–1191.
- Apiolaza LA, Potts BM, Gore PL (2001) Genetic control of peak flowering time of *Eucalyptus globulus*. In: *Developing the eucalypt of the future: IUFRO international symposium, 10–15 September 2001, Valdivia, Chile*. p 155
- Balasaravanan T, Chezian P, Kamalakannan R, Ghosh M, Yasodha R, Varghese M, Gurumurthi K (2005) Determination of inter and intra-species genetic relationships among six *Eucalyptus* species based on inter-simple sequence repeats (ISSR). *Tree Physiology* 25:1295–1302.
- Balasaravanan T, Chezian P, Kamalakannan R, Yasodha R, Varghese M, Gurumurthi K, Ghosh M (2006) Identification of Species-Diagnostic ISSR Markers for Six *Eucalyptus* Species *Silvae Genetica* 55-3
- Ballesta P, Mora F, Contreras-Soto R, Ruiz E, Perret S (2013) Genetic diversity analysis in Australian accessions of sugar gum (*Eucalyptus cladocalyx*) using ISSR markers. *Acta Scientiarum Agronomy* (Accepted).
- Birtchnell MJ y M Gibson (2006) Long-term flowering patterns of melliferous *Eucalyptus* (Myrtaceae) species. *Australian Journal of Botany* 54: 745–754.
- Birtchnell MJ, Gibson M (2008) Flowering ecology of honey-producing flora in South-East Australia. In: *RIRDC Report 08/098, Rural Industries Research and Development Corporation, Canberra, 77 pp.*
- Blakely WF (1955) *A key to the eucalypts, with descriptions of 522 species and 150 varieties* (2nd edn). (Forestry and Timber Bureau: Canberra)

- Boland DJ, Brooker MIH, Chippendale GM, Hall N, Hyland BPM, Johnston RD, Kleinig DA, Turner JD (1984) 'Forest trees of Australia.' (Nelson-CSIRO: Melbourne).
- Boomsma CD. 1981. The forest vegetation of South Australia. Woods and Forests Department Bulletin No 19. (Woods and Forests Department: Adelaide)
- Bornet B, Branchard M (2001) Non-anchored inter-simple sequence repeat (ISSR) markers: reproducible and specific tools for genome fingerprinting. *Plant Molecular Biology Report* 19:209–215
- Brawner JT, Dillon SK, Lee DJ, Meder AR, Dieters MJ, Southerton SG (2012) The use of genetic correlations to evaluate associations between SNP markers and quantitative traits. *Tree Genetics and Genomes* 8(6): 1423-1435.
- Breseghele F, Sorrells ME (2006) Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* L.) cultivars. *Genetics* 172:1165-1177.
- Brondani R, Williams E, Brondani C, Grattapaglia D (2006) A microsatellite-based consensus linkage map for species of *Eucalyptus* and a novel set of 230 microsatellite markers for the genus. *BMC Plant Biol* 6:20.
- Brooker MIH, DA Kleinig (1990) *Field Guide to Eucalypts*. Volume 2 South-western and southern Australia. Inkata Press, Melbourne, Australia.
- Bundock PC, Potts BM, Vaillancourt RE (2008) Detection and stability of quantitative trait loci (QTL) in *Eucalyptus globulus*. *Tree Genetics and Genomes* 4 (1): 85-95.
- Bush D, McCarthy K, Meder R (2011) Genetic variation of natural durability traits in *Eucalyptus cladocalyx* (sugar gum). *Annals of Forestry Science* 68:1057–1066
- Bush D, Thumma B (2013) Characterizing a *Eucalyptus cladocalyx* breeding population using SNP markers. *Tree Genetics and Genomes* DOI 10.1007/s11295-012-0589-1.
- Butcher PA, McDonald MW, Bell JC (2009) Congruence between environmental parameters, morphology and genetic structure in Australia's most widely distributed eucalypt, *Eucalyptus camaldulensis* *Tree Genetics and Genomes* 5:189-210.

Callister A, Bush D, Collins S, Davis W (2008) Prospects for genetic improvement of *Eucalyptus cladocalyx* in Western Australia. *New Zeal J For Sci* 38:211–226.

Cané-Retamales C, Mora F, Vargas-Reeve F, Perret S, Contreras-Soto R (2011) Bayesian threshold analysis of breeding values, genetic correlation and heritability of flowering intensity in *Eucalyptus cladocalyx* under arid conditions. *Euphytica* 178:177–183.

Cappa EP, Cantet RJC (2006) Bayesian inference for normal multiple-trait individual-tree models with missing records via full conjugate Gibbs. *Can J For Res* 36:1276–1285.

Cekic C, Battey NH, Wilkinson MJ (2001) The potential of ISSR-PCR primer-pair combinations for genetic linkage analysis using the seasonal flowering locus in *Fragaria* as a model. *Theoretical and Applied Genetics* 103:540–546.

Contreras-Soto R, Mora F, Perret S, Vargas-Reeve F, Cané-Retamales C (2011) Predicción bayesiana del comportamiento poblacional de *Eucalyptus cladocalyx* para características binarias de componentes de florecimiento y supervivencia en las zonas áridas de Chile. *Interciencia* 36 (9): 644-649.

Coopman RE, Jara JC, Escobar R, Corcuera LJ, Bravo LA (2010) Genotypic variation in morphology and freezing resistance of *Eucalyptus globulus* seedlings subjected to drought hardening in nursery. *Electron. J. Biotechnol* 13 (1):1-9.

Chambers PGS, Potts BM, Tilyard PA (1997) The genetic control of flowering precocity in *Eucalyptus globulus* ssp. *globulus*. *Silvae Genetica* 46:207-214

Chezian P, Yasodha R, Ghosh M (2010) Genetic diversity analysis in a seed orchard of *Eucalyptus tereticornis*. *New Forests* 40:85–99.

Clarke JH, Mithen R, Brown JKM, Dean C (1995) QTL analysis of flowering time in *Arabidopsis*. *Molecular and General Genetics* 248: 278–286.

Crane E, Walter P, Day R (1984) *Directory of Important World Honey Sources*. International Bee Research Association, London, UK, 384 pp.

Cumbie WP, Eckert A, Wegrzyn J, Whetten R, Neale D, Goldfarb B (2011) Association genetics of carbon isotope discrimination, height and foliar nitrogen in a natural population of *Pinus taeda* L. *Heredity* 107: 105–114.

Dillen S, Storme V, Marron N, Bastien C, Neyrinck S, Steenackers M, Ceulemans R, Boerjan W (2008) Genomic regions involved in productivity of two interspecific poplar families in Europe. 1. Stem height, circumference and volume. *Tree Genet Genomes* 5:147–164.

Domingues DS, Farro AP, Coscrato VE, de Mello EJ, Oda S, Marino CL (2006) Identification of RAPD and SCAR marker related to early flowering in *Eucalyptus grandis*. *Ciência Florestal* 16 (3): 251-260.

Dornelas MC, Neves do Amaral W, Rodríguez Martinelli AP (2004) EgLFY, the *Eucalyptus grandis* homolog of the Arabidopsis gene LEAFY is expressed in reproductive and vegetative tissues. *Braz. J. Plant Physiol.* 16 (2): 105-114.

Dornelas MC, Rodriguez Martinelli AP (2005) Identifying *Eucalyptus* expressed sequence tags related to Arabidopsis flowering-time pathway genes. *Braz. J. Plant Physiol* 17(2): 255-266.

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Photochemistry Bulletin* 19: 11–15.

Dutkowski GW, BM Potts (1999) Geographical patterns of genetic variation in *Eucalyptus globulus* ssp. *globulus* and a revised racial classification. *Australian Journal of Botany* 46: 237-263.

Eckert AJ, Bower AD, Pande B, Jermstad KD, Krutovsky KV, St. Clair JB, Neale DB (2009) Association genetics of coastal Douglas fir (*Pseudotsuga menziesii* var. *menziesii*, Pinaceae). I. Cold-hardiness related traits. *Genetics* 182: 1289-1302.

Eckert AJ, van Heerwaarden J, Wegrzyn JL, Nelson CD, Ross-Ibarra J, González-Martínez SC, Neale DB (2010) Patterns of population structure and environmental associations to aridity across the range of loblolly pine (*Pinus taeda* L., Pinaceae). *Genetics* 185: 969-982.

Eldridge K, Davidson J, Harwood C, Van Wyk G (1993) *Eucalypt* domestication and breeding. Clarendon Press, Oxford.

El-Lithy ME, Bentsink L, Hanhart CJ, Ruys GJ, Rovito D, Broekhof JLM, van der Poel HJA, van Eijk MJT, Vreugdenhil D, Koornneef M (2006) New *Arabidopsis* Recombinant Inbred Line Populations Genotyped Using SNP Wave and Their Use for Mapping Flowering-Time Quantitative Trait Loci. *Genetics* 172:1867–1876.

Endelman JB (2011) Ridge regression and other kernels for genomic selection with R package rrBLUP. *Plant Genetics* 4(3):250–255.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. *Mol Ecol* 14(8):2611–2620.

Excoffier L, Hofer T, Foll M (2009) Detecting loci under selection in a hierarchically structured population. *103*: 285-298.

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* 10: 564-567.

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164: 1567–1587.

Falush D, Stephens M, Pritchard JK (2007) Inference of population structure using multilocus genotype data: dominant markers and null allele. *Molecular Ecology Notes* 7: 574–578.

Fang DQ, Roose ML (1997) Identification of closely related citrus cultivars with inter-simple sequence repeat markers. *Theoretical and Applied Genetics* 95:408–417.

Farro AP, Bortoloto TM, Oda S, Marino CL (2013) Identification of molecular marker linked to early flowering in *Eucalyptus grandis*. *The Revista do Instituto Florestal* 24(2): 149-157.

Freeman JS, PW Simon, BM Potts, RE Vaillancourt (2009) QTL influencing growth and wood properties in *Eucalyptus globulus*. *Tree Genetics & Genomes* 5:713-722.

Gemas VJV, Almadanim MC, Tenreiro R, Martins A, Fevereiro P (2004a) Genetic diversity in the Olive tree (*Olea europaea* L. subsp. *europaea*) cultivated in Portugal revealed by RAPD and ISSR markers. *Genetic Resource Crop Evolution* 51:501–511.

Gemas VJV, Neves LO, Araujo C, Fevereiro P (2004b) How can ISSR markers be useful in the management of genetic resources in a *Eucalyptus globulus* ssp. *globulus* baseline collection. In: Borralho N (ed) *Eucalyptus* in a changing world. Aveiro, pp 1–7.

Gleadow RM, Woodrow IE (2000) Polymorphism in cyanogenic glycoside content and cyanogenic beta-glucosidase activity in natural populations of *Eucalyptus cladocalyx*. *Australian Journal of Plant Physiology* 27:693–699.

González-Martínez SC, NC Wheeler, EE C, D Nelson, DB Neale (2007) Association Genetics in *Pinus taeda* L. I. Wood Property Traits. *Genetics* 175 (1):399-409.

Gore PL and Potts BM (1995) The genetic control of flowering time in *Eucalyptus globulus*, *E. nitens* and their F1 hybrids. In: Potts BM, Borralho NMG, Reid JB, Cromer RN, Tibbits WN and Raymond CA (eds) *Eucalypt plantations: improving fibre yield and quality*. Cooperative Research Centre for Temperate Hardwood Forestry, Hobart. pp 241–242.

Grattapaglia D, Sederoff R (1994) Genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-testcross: mapping strategy and RAPD Markers. *Genetics* 137: 1121–1137.

Grattapaglia D, Bertolucci FL, Penchel R, Sederoff RR (1996) Genetic mapping of quantitative trait loci controlling growth and wood quality traits in *Eucalyptus grandis* using a maternal half-sib family and RAPD markers. *Genetics* 144: 1205–1214

Grattapaglia D, Kirst M (2008) *Eucalyptus* applied genomics: from gene sequences to breeding tools. *New Phytol* 179:911–929.

Grattapaglia D, Resende MDV. 2011. Genomic selection in forest tree breeding. *Tree Genetics and Genomes* 7: 241–255

- Grattapaglia D, Vaillancourt RE, Shepherd M, Thumma BR, Foley W, Külheim C, Potts BM, Myburg AA (2012) Progress in Myrtaceae genetics and genomics: *Eucalyptus* as the pivotal genus. *Tree Genetics and Genomes* DOI 10.1007/s11295-012-0491-x.
- Griffin AR, P Whiteman, T Rudge, IP Burgess, M Moncur (1993) Effect of paclobutrazol on flower-bud production and vegetative growth in two species of *Eucalyptus*. *Canadian Journal Research* 23:640–647.
- Griffin AR (1980). Floral phenology of a stand of mountain ash (*Eucalyptus regnans* F. Muell) in Gippsland, Victoria. *Aust J Bot* 28: 393-404.
- Hadfield JE (2010) MCMC Methods for Multi-Response Generalized Linear Mixed Models: The MCMCglmm R Package. *Journal of Statistical Software* 2(3): 1-22.
- Hanks LM, Paine TD, Millar JG, Hom JL (1994) Variation among *Eucalyptus* species in resistance to eucalyptus long horned borer in Southern California. *Entomol Exp Appl* 74:185-194.
- Hardwood CE, Bush DJ, Butcher T, Bird R, Henson ML, Lott R, Shaw S (2007) Achievements in forest tree genetic improvement in Australia and New Zealand. 4. Tree improvement for low-rainfall farm forestry. *Australian Forestry* 70:23–27
- Hasan O, Reid JB (1995) Reduction of generation time in *Eucalyptus globulus*. *Plant Growth Regulation* 17:53–60.
- Henderson CR (1949) Estimation of changes in herd environment. *Journal of Dairy Science* 32: 709 (Abstract).
- Henderson CR (1985) Best linear unbiased prediction of non-additive genetic merits in non-inbred populations. *Journal of Animal Science* 60: 111–117.
- Hobbs TJ, M Bennell, D Huxtable, J Bartle, C Nuemann, N George, W O'Sullivan, D McKenna (2009) Potential agroforestry species and regional industries for lower rainfall southern Australia: Florasearch 2. Report to the Joint Venture Agroforestry Program (JVAP) and the Future Farm Industries CRC. RIRDC Publication 07/082, RIRDC, Canberra.

Hu Z, Xu S (2009). PROC QTL-A SAS procedure for mapping quantitative trait loci. *International Journal of Plant Genomics* 2009: 141-234.

Hudson IL, Kim SK, Keatley MR (2009) Climatic influences on the flowering phenology of four Eucalypts: a GAMLSS approach. 18th World IMACS / MODSIM Congress, Cairns, Australia 13-17 July 2009.

Izawa T, Takahashi Y, Yano M (2003) Comparative biology comes into bloom: genomic and genetic comparison of flowering pathways in rice and Arabidopsis. *Curr Opin Plant Biol* 6:113–120.

Jackson MJ, Line MA, Hasan O (1996) Microbial degradation of a recalcitrant plant growth retardant-paclobutrazol (PP333). *Soil Biology and Biochemistry* 28:1265–1267.

Jacobs MR (1955) *Growth Habits of the eucalypts*. 262 pp. Commonwealth Forestry and Timber Bureau, Government Printer, Canberra, Australia.

Jacobs MR (1981) *Eucalypts for Planting, Edition 2*. FAO Forestry, Food and Agriculture Organization of the United Nations, Rome, Italy.

Jaya ESKD, Clemens J, Song J, Zhang H, Jameson PE (2010) Quantitative expression analysis of meristem identity genes in *Eucalyptus occidentalis*: AP1 is an expression marker for flowering. *Tree Physiology* 30: 304-312.

Jones R, Hecht VFG, Potts BM, Vaillancourt R, Weller JL (2011) Expression of a FLOWERING LOCUS T homologue is temporally associated with annual flower bud initiation in *Eucalyptus globulus* subsp. *globulus* (Myrtaceae). *Australian Journal of Botany* 59: 756–769.

Jordan GJ, Potts BM, Wiltshire RJE (1999) Strong, independent, quantitative genetic control of the timing of vegetative phase change and first flowering in *Eucalyptus globulus* ssp. *globulus* (Tasmanian Blue Gum). *Heredity* 83:179–187.

Jovanovic T, Booth T (2002) Improved species climatic profiles. RIRDC Publication No 02/095, RIRDC, Canberra

Kamalakannan R, Barthwal S, Chezian P, Balasaravanan T, Yasodha R, Gurumurthi K, Ghosh M (2006) Morphological and Molecular diversity among *Casuarina* and *Allocasuarina* species. *Biotechnology* 5:301–307.

Kar PK, Srivastava PP, Awasthi AK, Urs SR (2008) Genetic variability and association of ISSR markers with some biochemical traits in mulberry (*Morus sp.*) genetic resources available in India. *Tree Genetics and Genomes* 4 (1): 75– 83.

Keatley MR, Hudson IL (2007) A Comparison of Long-Term Flowering Patterns of Box-Ironbark Species in Havelock and Rushworth Forests. *Journal of Environmental Modelling and Assessment* 12: 279-292.

Kennedy BW, Quinton M, Vanarendonk JAM (1992) Estimation of effects of single genes on quantitative traits. *J. Anim. Sci* 70:2000–2012.

Lander ES, Schork NJ (1994) Genetic dissection of complex traits. *Science* 265: 2037–2048.

Law B, Mackowski C, Schoer L, Tweedie T (2001) Flowering phenology of Myrtaceous trees and their relation to climatic, environmental and disturbance variables in northern New South Wales. *Austral Ecology* 25: 160–178.

Law B y M Chidel (2007) Effects of logging on nectar-producing eucalypts – Spotted Gum and Grey Ironbark, Publication No. 07/138. Rural Industries and Development Corporation, Canberra, pp.1-69.

Lecis R, Pierpaoli M, Biro ZS (2006) Bayesian analyses of admixture in wild and domestic cats (*Felis silvestris*) using linked microsatellite loci. *Molecular Ecology*. 15: 119–131.

Lepoittevin C, Harvengt L, Plomion C, Garnier-Géré P (2012) Association mapping for growth, straightness and wood chemistry traits in the *Pinus pinaster* Aquitaine breeding population. *Tree Genetics and Genomes* 8 (1):113–126.

Lund MS, Sahana G, de Koning D, Su G, Carlborg O (2009) Comparison of analyses of the QTLMAS XII common dataset. I: Genomic selection. *BMC Proceedings* 2009, 3(Suppl 1): S1.

Lynch M, Milligan BG (1994) Analysis of population genetic structure with RAPD markers. *Mol Ecol* 3: 91–99.

Mamani EMC, Bueno NW, Faria DA, Guimarães LMS, Lau D, Alfenas ACA, Grattapaglia D (2010) Positioning of the major locus for *Puccinia psidii* rust resistance (Ppr1) on the *Eucalyptus* reference map and its validation across unrelated pedigrees. *Tree Genet Genomes* 6:953–962.

Manikis I, Thrasivoulou A (2001) The relation of physicochemical characteristics of honey and the crystallization sensitive parameters. *Apiacta* 36:106-112.

McMahon L, G Brendan, R Hean (2010) *Eucalyptus cladocalyx*. *Primefact* 1077:1-5.

Marques CM, Vasquez-Kool J, Carocha VJ, Ferreira JG, O'Malley DM, Liu BH, Sederoff R (1999) Genetic dissection of vegetative propagation traits in *Eucalyptus tereticornis* and *E. globulus*. *Theoretical and Applied Genetics* 99:936–946

Marcar NE, Crawford DF, Leppert PM, Jovanovic T, Floyd R, Farrow R (1995) *Trees for saltland: a guide to selecting native species for Australia*. CSIRO, Melbourne, p 72.

McDonald MW, Rawlings M, Butcher PA, Bell JC (2003) Regional divergence and inbreeding in *Eucalyptus cladocalyx* (Myrtaceae). *Australian Journal of Botany* 51:393–403.

McGowen MH, Potts BM, Vaillancourt RE, Gore P, Williams DR, Pilbeam DJ (2004) The genetic control of sexual reproduction in *Eucalyptus globulus*. In: Borralho NMG, Pereira JS, Marques C, Coutinho J, Madeira M and Tomé M (eds) *Eucalyptus in a changing world: proceedings of the IUFRO conference, Aveiro, 11–15 October 2004*. RAIZ, Instituto Investigação de Floresta e Papel, Aveiro. pp 104–108

McGowen MH (2007) Genetic control of reproductive traits in *Eucalyptus globulus*. PhD thesis, University of Tasmania, Hobart.

Meilan R (1997) Floral induction in woody angiosperms. *New Forest* 14:179–202.

Meuwissen TH, Hayes BJ, Goddard ME. 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157: 1819–1829.

Missiaggia AA, Piacuzzi AL, Grattapaglia D (2005) Genetic mapping of Eef1, a major effect QTL for early flowering in *Eucalyptus grandis*. *Tree Genetics and Genomes* 1(2): 79–84.

Moncur MW, Hand FC, Ramsden NG (1994). Environmental and cultural effects on flowering and seed production of plantation grown *Eucalyptus nitens*. Report for the Tasmanian Forest Research Council, Inc.(Division of Forestry, CSIRO Canberra).

Moncur MW, Hasan O (1994) Floral induction in *Eucalyptus nitens* (Deane and Maiden) Maiden. *Tree Physiology* 14: 1303-1312.

Montenegro G, Pizarro R, Avila G, Castro R, Ríos C, Muñoz O, Bas F, Gómez M (2003) Origen botánico y propiedades químicas de las mieles de la región mediterránea árida de Chile. *Ciencia e Investigación Agraria* 30:161–174.

Moraga-Suazo P, Hasbún R, Balocchi C, Valenzuela S (2012) Establishment and optimization of ISSR and SAMPL molecular markers as a tool for breeding programs of *Pinus radiata*. *Bosque* 33(1): 93-98.

Mora F, Perret S, Scapim CA, Martins EN, Molina MP (2007) Source-dependent blooming variability of *Eucalyptus cladocalyx* in the region of Coquimbo, Chile. *Ciencia e Investigación Agraria* 34(2):99–106.

Mora F, Gleadow R, Perret S, Scapim CA (2009) Genetic variation for early flowering, survival and growth in sugar gum (*Eucalyptus cladocalyx* F. Muell) in southern Atacama Desert. *Euphytica* 169:335–344.

Mrode RA (2005) Linear models for the prediction of animal breeding values. 2nd edition 2005, Technology & Engineering, 344 pp.

Myburg AA, Griffin RA, Sederoff RR, Whetten RW (2003) Comparative genetic linkage maps of *Eucalyptus grandis*, *Eucalyptus globulus* and their F1 hybrid based on a double pseudo-backcross mapping approach. *Theoretical and Applied Genetics* 107:1028–1042.

Myburg AA, Potts BM, Marques CM, Kirst M, Gion JM, Grattapaglia D, Grima-Pettenati J (2007) *Eucalyptus*. In: Kole C (ed) *Genome mapping and molecular breeding in plants* Springer. NY, USA, New York pp 115–160

Neale DB, Savolainen O (2004) Association genetics of complex traits in conifers. *Trends Plant Science* 9: 325–330.

Neale DB, Kremer A (2011) Forest tree genomics: growing resources and applications. *Nat. Rev. Genet.* 12: 111–122.

Okun DO, Kenya EU, Oballa PO, Odee DW, Muluvi GM (2008) Analysis of genetic diversity in *Eucalyptus grandis* (Hill ex Maiden) seed sources using inter simple sequence repeats (ISSR) molecular markers. *African Journal of Biotechnology* 7:2119–2123

Piepho HP, Mohring J, Melchinger AE, Buchse A (2008) BLUP for phenotypic selection in plant breeding and variety testing. *Euphytica* 161:209–228.

Poke FS, Vaillancourt RE, Potts BM, Reid JB (2005) Genomic research in *Eucalyptus*. *Genetica* 125:79–101

Plummer M, Best N, Cowles K, Vines K (2006) CODA: convergence diagnosis and output analysis for MCMC. *R News* 6(1):7–11.

Porter JW (1978) Relationships between flowering and honey production of red ironbark, *Eucalyptus sideroxylon* (A. Cunn.) Benth and climate in the Bendigo district of Victoria. *Australian Journal of Agricultural Research* 29:815-829.

Potts BM, Barbour RC, Hingston AB (2001) Genetic Pollution from Farm Forestry using eucalypt species and hybrids. RIRDC Publication No 01/114.

Potts BM (2004) Genetic improvement of eucalypts. In: Burley J, Evans J, Youngquist JA (eds) *Encyclopedia of forest science*. Elsevier Science, Oxford, pp1480–1490.

Price AL, Zaitlen NA, Reich D, Patterson N (2010) New approaches to population stratification in genome-wide association studies. *Nat Rev Genet* 11: 459–463.

Pritchard JK, Rosenberg NA (1999) Use of unlinked genetic markers to detect population stratification in association studies. *Am. J. Hum. Genet.*65:220–228.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945–959.

- Pryor LD (1976) *The biology of eucalypts* (Edward Arnold: Southampton, London).
- Poynton RJ (1979) *Tree Planting in Southern Africa: The Eucalypts*. Department of Forestry, Pretoria, South Africa.
- Reddy M, Sarla N, Siddiq E (2002) Inter simple sequence repeat (ISSR) polymorphism and its application in plant breeding. *Euphytica* 128: 9–17.
- Resende MDV, Lopes PS, Silva RL, Pires IL (2008) Seleção genômica ampla (GWS) e maximização da eficiência do melhoramento genético. *Pesqui Florestal Bras* 56:63–77.
- Resende MDV, Resende Jr MFR, Sansaloni CP, Petroli CD, Missiaggia AA, Aguiar AM, Abad JM, Takahashi EK, Rosado AM, Faria DA, Pappas Jr GJ, Kilian A, Grattapaglia D (2012) Genomic selection for growth and wood quality in *Eucalyptus*: capturing the missing heritability and accelerating breeding for complex traits in forest trees. *New Phytologist* 194 (1): 116-128.
- Rosenberg NA, Pritchard JK, Weber JL (2002) Genetic structure of human populations. *Science* 298: 2381–2385.
- Ruan CH, Li H, Mopper S (2009) Characterization and identification of ISSR markers associated with resistance to dried-shrink disease in sea buckthorn. *Molecular Breeding* 24:255–268.
- Ruthrof KX, Loneragan WA, Yates CJ (2003) Comparative population dynamics of *Eucalyptus cladocalyx* in its native habitat and as an invasive species in an urban bushland in south-western Australia. *Diversity & Distributions* 9(6):469-483.
- Savva M, Potts BM, Reid JB (1988) The breeding system and gene flow in *Eucalyptus urnigera*. In: *Pollination '88*. (Ed. R.B. Knox, M.B. Sing and L.F. Troiani.) pp. 176-182.
- Schuster I (2011) Marker-assisted selection for quantitative traits. *Crop Breeding and Applied Biotechnology* S1: 50-55.
- Shepherd M, Chaparro JX, Teasdale R (1999) Genetic mapping of monoterpene composition in an interspecific eucalypt hybrid. *Theoretical and Applied Genetics* 99:1207–1215.

Sillanpää MJ, Corander J (2002) Model choice in gene mapping: what and why. *Trends in Genetics* 18 (6): 302-307.

Somerville DC, Moncur MW (1997) The importance of eucalypt species for honey production in NSW, Australia, Proceedings of the 35th International Apicultural Congress of Apimodia, 1-6 September 1997, Antwerp, Belgium. pp. 327–331.

Somerville DC, Nicholson D (2005) The primary melliferous flora and other aspects associated with beekeeping within State forests of New South Wales as determined by surveys of beekeepers. *Australian Forestry* 68 (1) 9 - 16.

Southerton SG, Strauss SH, Olive MR, Harcourt RL, Decroocq V, Zhu X, Llewellyn DJ, Peacock WJ, Dennis ES (1998) *Eucalyptus* has a functional equivalent of the *Arabidopsis* floral meristem identity gene LEAFY. *Plant Molecular Biology* 37:897–910

Steane DA, Nicolle D, Sansaloni CP, Petroli CD, Carling J, Kilian A, Myburg AA, Grattapaglia D, Vaillancourt RE (2011) Population genetic analysis and phylogeny reconstruction in *Eucalyptus* (Myrtaceae) using high-throughput, genome-wide genotyping. *Molecular Phylogenetics Evolution* 59:206-224.

Slatkin M (1991) Inbreeding coefficients and coalescence times. *Genet Res* 58:167–175.

Stich B, J Möhring, H Piepho, M Heckenberger, E Buckler, A Melchinger (2008) Comparison of Mixed- Model Approaches for Association Mapping. *Genetics* 178: 1745–1754.

Thamarus K, Groom K, Murrell J, Byrne M, Moran G (2002) A genetic linkage map for *Eucalyptus globulus* with candidate loci for wood, fibre and floral traits. *Theoretical and Applied Genetics* 104:379–387.

Thornsberry JM, Goodman MM, Doebley J, Kresovich S, Nielsen S, Buckler ES (2001) Dwarf8 polymorphisms associate with variation in flowering time. *Nat. Genet.* 28:286–289.

Thumma BR, Nolan MR, Evans R, Moran GF (2005) Polymorphisms in cinnamoyl CoA reductase (CCR) are associated with variation in microfibril angle in *Eucalyptus sp.* *Genetics* 171:1257–1265.

Thumma BR, Baltunis BS, Bell JC, Emebiri LC, Moran GF, Southerton SG (2010) Quantitative trait locus (QTL) analysis of growth and vegetative propagation traits in *Eucalyptus nitens* full-sib families. *Tree Genet Genomes* 6(6):877–889.

Turnbull JW, Pryor LD (1984) Choice of species and seed sources. In *Eucalypts for wood production*. (Eds WE Hillis, AC Brown) CSIRO, Academic Press: Melbourne, Australia.

Van-Tassell CP, Van-Vleck LD (1996) Multiple-trait Gibbs sampler for animal models: flexible programs for Bayesian and likelihood-based (co)variance component inference. *Journal of Animal Science* 74:2586–2597.

Verhaegen D, C Plomion, JM Gion, M Poitel, P Costa, A Kremer (1997) Quantitative trait dissection analysis in *Eucalyptus* using RAPD markers: 1. Detection of QTL in interspecific hybrid progeny, stability of QTL expression across different ages *Theoretical and Applied Genetics* 95 (4): 597-608.

Vijayan K, Srivatsava PP, Nair CV, Awasthi AK, Tikader A, Sreenivasa B, Urs SR (2006) Molecular characterization and identification of markers associated with yield traits in mulberry using ISSR markers. *Plant Breeding* 125:298–301.

Waldmann P, Ericsson T (2006) Comparison of REML and Gibbs sampling estimates of multi-trait genetic parameters in Scots pine. *Theor Appl Genet* 112:1441–1451

Wang H, Zhang Y, Li X, Masinde GL, Mohan S, Baylink DJ (2005) Bayesian shrinkage estimation of quantitative trait loci parameters. *Genetics* 170: 465–480.

Wei X, Jackson PA, McIntyre CL, Aitken KS, Croft B (2006) Associations between DNA markers and resistance to diseases in sugarcane and effects of population substructure. *Theor Appl Genet* 114:155-164.

Williams DR, Potts MB, Philip JS (2003) Promotion of flowering in *Eucalyptus nitens* by paclobutrazol was enhanced by nitrogen fertilizer. *Can. J. For. Res.* 33: 74–81.

Wiltshire RJE, J Reid, BM Potts (1998) Genetic control of reproductive and vegetative phase change in the *Eucalyptus risdonii* – *E. tenuiramis* complex. Australian Journal of Botany 46:45-63.

Woodrow IE, Slocum D, Gleadow RM (2002) Influence of water stress on cyanogenic capacity in *Eucalyptus cladocalyx*. Australian Journal of Plant Physiology 29:103–110.

Xu S, Hu Z (2011) Mapping quantitative trait loci using the MCMC procedure in SAS. Heredity 106: 357-369.

Xu F, Rong X, Huang X, Cheng S (2012) Recent Advances of Flowering Locus T Gene in Higher Plants. Int. J. Mol. Sci. 13: 3773-3781.

Yaish MW, Colasanti J, Rothstein SJ (2011) The role of epigenetic processes in controlling flowering time in plants exposed to stress. Journal Experimental Botany 62(11):3727-3735.

Yasodha R, Kathirvel M, Sumathi R, Gurumurthi K, Sunil A, Nagaraju J (2004) Genetic analyses of casuarinas using ISSR and FISSR markers. Genetica 122:161–172.

Zapata-Valenzuela J, Isik F, Maltecca C, Wegrzyn J, Neale D, McKeand S, Whetten R (2012) SNP markers trace familial linkages in a cloned population of *Pinus taeda*—prospects for genomic selection. Tree Genetics and Genomes 8 (6): 1307-1318.

Zhao K, Aranzana MJ, Kim S, Lister C, Shindo C, Tang C, Toomajian C, Zheng H, Dean C, Marjoram Z, Nordborg M (2007) An Arabidopsis Example of Association Mapping in Structured Samples PLoS Genet 3(1).

Zhivotovsky LA (1999) Estimating population structure in diploids with multilocus dominant DNA markers. Mol Ecol 8:907–913.

Zhu C, Gore M, Buckler ES, Yu J (2008) Status and prospects of association mapping in plants. The Plant Genome 1:5–20.

Zietkiewicz E, Rafalski A, Labuda D (1994) Genome fingerprinting by simple sequence repeat (SSR)-anchored polymorphism chain reaction amplification. Genomics 20 (2): 176-183