



Universidad de Concepción
Dirección de Postgrado
Facultad de Ciencias Forestales

PROGRAMA DE DOCTORADO EN CIENCIAS FORESTALES

Título en español:

“UTILIZANDO HERRAMIENTAS DE GENÓMICA DEL PAISAJE PARA
MODELAR LA ADAPTACIÓN LOCAL EN *Araucaria araucana* (Molina) K. Koch”

Título en inglés:

“USING LANDSCAPE GENOMICS TOOLS TO MODEL LOCAL ADAPTATION IN
Araucaria Araucana (Molina) K. Koch”

Tesis para optar al grado de Doctor en Ciencias Forestales

ANTONIO ALBERTO VARAS MYRIK

Concepción-Chile

2022

Profesor Guía: Rodrigo Hasbún Z.

Dpto. de Silvicultura

Facultad de Ciencias Forestales

Universidad de Concepción

UTILIZANDO HERRAMIENTAS DE GENÓMICA DEL PAISAJE PARA MODELAR LA ADAPTACIÓN LOCAL en *Araucaria araucana* (Molina) K. Koch

Comisión Evaluadora:

Rodrigo Hasbún Z. (Profesor Guía) _____
Ingeniero Forestal; Dr.

Óscar Toro N. (Comisión Evaluadora) _____
Biólogo; Dr.

Eugenio Sanfuentes V. (Comisión Evaluadora) _____
Ingeniero Forestal; Dr.

Diego Alarcón. (Asesor externo) _____
Ingeniero Forestal; Dr.



Álex Fajardo Yáñez (Asesor externo) _____
Ingeniero Forestal; Dr.

Directora de Postgrado:

Darcy Ríos Leal _____
Bióloga; Dr.

Decano de Facultad de Ciencias Forestales:

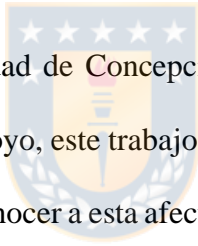
Manuel Sánchez Olate _____
Ingeniero Forestal; Dr.

AGRADECIMIENTOS

Esta tesis está dedicada a Juanita, Pamela, Javiera, Paola y Francisca. Todas grandes mujeres cuyo amor incondicional e infinita paciencia hicieron posible sortear los momentos difíciles que se presentaron. A todas ellas mi más sincero y profundo agradecimiento.

Quiero agradecer a mi profesor Guía, Rodrigo Hasbún, por su gran apoyo, disposición, comprensión y trato cordial. También mis agradecimientos para el resto de los profesores de la comisión, quienes siempre mostraron su gran voluntad por apoyar, enseñar y también por aprender con este trabajo.

Finalmente quiero agradecer a la Universidad de Concepción, la Facultad de Ciencias Forestales y su Dirección de Posgrado. Sin su irrestricto apoyo, este trabajo hubiera sido imposible. Siempre me hicieron sentir como en casa. Tuve el privilegio de conocer a esta afectuosa familia, en cuyo seno me sentí valorado, querido y respetado.



INDICE GENERAL

AGRADECIMIENTOS	iii
INDICE DE FIGURAS.....	vii
INDICE DE TABLAS	ix
RESUMEN.....	1
ABSTRACT	2
CAPÍTULO I.....	3
INTRODUCCIÓN GENERAL.....	3
1.1 <i>La adaptación de las plantas al ambiente</i>	3
1.2 <i>Cambio climático, desadaptación y conservación</i>	6
1.3 <i>Los marcadores moleculares y la divergencia adaptativa</i>	9
1.3.1 <i>RADseq</i>	11
1.4 <i>La genómica del Paisaje y el modelamiento de la Divergencia Adaptativa</i>	13
1.5 <i>Especie de estudio: Araucaria araucana</i>	17
1.5.1 <i>Estudios genéticos en Araucaria araucana.</i>	19
1.5.2 <i>La muerte apical en Araucaria araucana.</i>	20
1.6 <i>Aporte Científico de la Tesis</i>	22
BIBLIOGRAFÍA.....	23
CAPÍTULO II.....	32
Predicting climate change-related genetic offset for the endangered southern South American conifer <i>Araucaria araucana</i>	32
ABSTRACT	33
2.1 INTRODUCTION	34
2.2 MATERIALS AND METHODS	36
2.2.1 <i>Sampling, RAD-Seq library preparation and sequencing</i>	36

2.2.2	<i>Environmental predictors</i>	38
2.2.3	<i>Building the adaptive divergence model</i>	40
2.2.4	<i>Distance-based redundancy analysis</i>	41
2.2.5	<i>Genetic offset predictions</i>	42
2.3	RESULTS	43
2.3.1	<i>Adaptive divergence model</i>	43
2.3.2	<i>Distance based redundancy analysis</i>	47
2.3.3	<i>Genetic offset</i>	48
2.4	DISCUSSION	50
2.4.1	<i>Population Structure</i>	50
2.4.2	<i>Local Adaptation</i>	50
2.4.3	<i>Genetic Offset and Conservation Strategy</i>	52
2.4.4	<i>Methodological Challenges</i>	52
	LITERATURE CITED	53
	SUPPORTING INFORMATION	60
	CAPÍTULO III	65
	Searching for Pre-Adapted Genotypes Using an Adaptive Divergence Model for the iconic South American conifer <i>Araucaria araucana</i>	65
	ABSTRACT	66
3.1	INTRODUCTION	67
3.2	MATERIALS AND METHODS	69
3.3	RESULTS	71
3.3.1	<i>Four Zones Resolution</i>	71
3.3.2	<i>Eight Zones Resolution</i>	73
3.4	DISUSSSION	75
3.4.1	<i>In-situ conservation</i>	77

LITERATURE CITED	78
CAPÍTULO IV	81
Aplicando el modelo de divergencia adaptativa para investigar zonas de acogida potenciales fuera de la distribución natural de <i>Araucaria araucana</i>	81
4.5 Metodología	81
4.6 Resultados	82
4.7 DISCUSIÓN	85
5 CONCLUSION GENERAL	87



INDICE DE FIGURAS

Figure 1. Sampling locations map of <i>A. araucana</i> . Panel A): regional map. The red box indicates sampling area. Panel B): detailed map. Red dots show sampling locations and its code.....	37
Figure 2. Population memberships ($K = 2$) of the 134 individuals of <i>A. Araucana</i> , based on sNMF algorithm. Vertical bars represent the percentage of individual's genome that belonged to each of the two ancestral groups (G1 and G2). Note that individuals from the VA location represent a distinctive class of admixture, showing a high presence of G1. Individuals from TR and PN constituted a second class of admixture, with a relatively balanced presence of G1 and G2, and the remaining individuals from Andes Mountain Range made up a third admixture class, with a low presence of G1.....	43
Figure 3. GDM I-splines for each environmental predictor and geographic distance. The maximum height of each curve indicates the total amount of change in allele frequencies associated with that predictor (variable importance). The shape of each curve indicates how the rate of change in allele frequencies varies along the predictor gradient.....	45
Figure 4. Map showing six zones of similar adaptive allelic compositions for <i>A. araucana</i> , built with the K-means clustering algorithm applied to GDM transformed predictor's data frame. These zones were used as genetic offset quantification units.....	46
Figure 5. Genetic offset predictions per year and adaptive zone, for each RCP. Bars represent an area-weighted average of each of 6 adaptive zones overlaid onto 4 genetic offset raster (2 RCPs and 2 Years). Genetic offset values (axis x) correspond to individual adaptive genomic distance computed as Jaccard index.....	48
Figure 6. Continuous genetic offset maps of <i>A. araucana</i> for the years 2050 and 2070 under the RCP4.5 (up) and RCP8.5 (down) scenarios. Black and yellow represent high and low offset levels, respectively.	49
Figure 7. Illustration of ex-situ genetic offset calculation of a pixel regarding the mean future environmental conditions of a recipient adaptive zone. Offset is then evaluated between the present environmental data of a given pixel and the mean future conditions of an adaptive recipient zone. In this case, the recipient zone may be assumed as a single pixel, since its future environmental conditions (Worldclim 1.4, CMIP5, year 2070, RCP4.5) were averaged. On the other hand, in the candidate donor zone, each pixel is considered in the evaluation of genetic offset.	71
Figure 8. Map showing four zones of similar adaptive allelic compositions for <i>A. araucana</i> , built using the K-means clustering algorithm applied to GDM transformed predictor's data frame.	72

Figure 9. Genetic offset continuous map for each four adaptive zone. Each pixel represents genetic offset among current environmental conditions of that pixel and average future environmental conditions of the adaptive zone. 73

Figure 10. Map showing eight zones of similar adaptive allelic compositions for *A. araucana*, built with the K-means clustering algorithm applied to GDM transformed predictor's data frame. Black arrows represent AM from zone 5 to zone 7. 74

Figure 11. Offset genético de cada una de las 4 zonas adaptativas respecto del área de acogida potencial. Zonas más claras representan un potencial mayor para la colonización asistida de *A. araucana*, para la zona adaptativa correspondiente 83

Figure 12. Zona candidata de acogida para la zona adaptativa 4 (4 zonas). Se muestran los lugares donde el offset es menor a 0.63, que es el offset in-situ promedio de la zona 4. Colores más claros representan un mayor potencial de acogida. 84



INDICE DE TABLAS

Table 1. Sampling locations, geographic coordinates (degrees of south latitude and west longitude), sample size (N) and Mountain Range of the 12 *A. araucana* sampled sites. 37

Table 2. Distance-based redundancy analysis (db-RDA). The amount of variance explained by each component or its combination is expressed as adjusted R^2 , in addition to its significance level. **Note:** F: Dependent matrix of individual genetic distances; the db-RDA tests are in the form of independent matrices $F \sim |$ covariate matrices. env.: environment; geo.: geographic distance; anc.: ancestry. Total Explained: Adjusted R^2 total of individual fractions. Abbreviation: ns: not significant. * $p < .05$; ** $p < .01$; *** $p < .001$ 46

Table 3. Genetic offset values for each adaptive zone, according to RCP and year of projection. Values correspond to area-weighted average of each of 6 adaptive zones overlaid onto 4 genetic offset raster (2 RCPs and 2 Years). Genetic offset values correspond to individual adaptive genomic distance computed as Jaccard index 47

Table 4. Pairwise genetic offset values of four adaptive zones. Diagonal values in bold are in-situ genetic offset. If in a row there are a lower value than bold number, the zone matching to the column of this cell was assumed as a candidate donor zone. 72

Table 5. Pairwise genetic offset values of eight adaptive zones. Diagonal values in bold are mean in-situ genetic offset. If in a row there are a lower value than bold number (*), the zone matching to the column of this cell was assumed as a candidate donor zone. 74

Table 6. Rangos climáticos utilizados para construir la máscara de proyección de zonas de acogida potenciales para migración asistida de *A. araucana*. Los rangos utilizados fueron Min-B y Max-B, que corresponden al rango actual del nicho más y menos 10%. 82

RESUMEN

Comprender la variación genética adaptativa es clave para predecir la respuesta evolutiva de especies y poblaciones al cambio climático, lo que influye de manera decisiva en la gestión para la conservación y el manejo de los bosques en el mundo. En la búsqueda de herramientas científicas para cuantificar, comprender y modelar dicha variación, ha emergido la genómica del paisaje como una disciplina sintética de desarrollo reciente, que combina elementos de genómica de poblaciones, ecología del paisaje y estadística espacial, permitiendo la cuantificación de los efectos de la selección natural, la geografía y la historia demográfica en la variación genómica, a través del análisis de las relaciones entre el genotipo y los elementos del paisaje. Estas herramientas permiten, entre otras aplicaciones, construir modelos de divergencia adaptativa espacialmente explícitos, que pueden emplearse, por ejemplo, para identificar aquellas poblaciones sujetas a un mayor riesgo de mala adaptación debido al cambio climático antropogénico, y también para informar programas de migración y colonización asistida.

En este trabajo, utilizando secuenciación de ADN asociada al sitio de restricción (RAD-Seq) y más de 49.000 polimorfismos de un solo nucleótido obtenidos de poblaciones de *Araucaria araucana* en Chile, se construyó un modelo de divergencia adaptativa, que luego fue utilizado para responder las siguientes preguntas: ¿Qué poblaciones de la distribución natural de *A. araucana* estarán expuestas a un riesgo más alto de desadaptación producido por el cambio climático?; ¿existen genotipos pre-adaptados a las condiciones futuras que permitan implementar un programa de migración asistida para disminuir el riesgo de desadaptación?; y ¿es posible implementar un programa de colonización asistida fuera de la distribución de *A. araucana*?. Los resultados muestran que las poblaciones ubicadas en el piedemonte del sector Andino sur de la distribución de *A. araucana*, presentan el mayor riesgo de desadaptación, mientras que las poblaciones que viven en las zonas altas de la Cordillera de los Andes presentan el riesgo más bajo. También se concluyó que un programa de migración asistida para mejorar la aptitud promedio de las poblaciones frente al cambio climático será difícil de implementar, debido a la falta de poblaciones que contengan genotipos pre-adaptados a las condiciones climáticas futuras. Finalmente, se concluye que la colonización asistida sólo podría ser factible para las poblaciones con mayor riesgo de mala adaptación. Los hábitats adecuados para la traslocación se identificaron en las zonas altas de la región andina, desde Futrono en Chile (-40,00° latitud sur), hasta Río Pico en Argentina (-44,27° sur). latitud), predominantemente en las regiones chilenas de Los Ríos y Los Lagos.

ABSTRACT

Understanding adaptive genetic variation is key to predicting the evolutionary response of species and populations to climate change, which decisively influences forest management and conservation efforts in the world. In the searching for scientific tools to quantify, understand and modelling such variation, landscape genomics has emerged as a synthetic discipline of recent development, which combines elements of population genomics, landscape ecology and spatial statistics, allowing the quantification of the effects of natural selection, geography and demographic history in genomic variation, through the analysis of relationships between genotypes and landscape elements. These tools allow, among other applications, to build spatially explicit adaptive divergence models, which can be used, for example, to identify those populations subject to a higher risk of maladaptation due to anthropogenic climate change. In addition, these skills can be helpful to inform assisted migration and colonization programs.

In this work, using restriction site-associated DNA sequencing (RAD-Seq) and more than 49,000 single nucleotide polymorphisms obtained from populations of *Araucaria araucana* in Chile, an adaptive divergence was built, which was later used to answer the following questions: Which populations of the natural distribution of *A. araucana* will be exposed to a higher risk of maladaptation due to climate change?; Is possible to find pre-adapted genotypes to future conditions that allow implementing an assisted migration program to reduce the risk of maladaptation?; Is it possible to implement an assisted colonization program outside the range of *A. araucana*? The results show that the populations located in the foothills of the southern Andean zone of *A. araucana* range present the greatest risk of maladaptation, while the populations living in the highlands of the Andes Mountains present the lowest risk. It was also concluded that an assisted migration program to improve the average fitness of populations against climate change will be difficult to implement, due to the lack of populations containing pre-adapted genotypes to future climatic conditions. Finally, it was concluded that assisted colonization could only be feasible for populations showing a higher risk of maladaptation. Suitable habitats for translocation were identified in the highlands of the Andean region, from Futrono in Chile (-40.00° south latitude) to Río Pico in Argentina (-44.27° south latitude), predominantly in the Chilean regions of Los Ríos and Los Lagos.

CAPÍTULO I

INTRODUCCIÓN GENERAL

1.1 *La adaptación de las plantas al ambiente*

La aptitud, adecuación biológica o “fitness”, hace referencia a un factor evolutivo que mide la contribución relativa de alelos de un individuo a la próxima generación respecto de otros de la misma población, y está estrechamente relacionada a su capacidad para crecer y reproducirse en un ambiente dado (Eriksson *et al.*, 2006). Si existen diferencias entre genotipos que afecten a la aptitud, entonces las frecuencias de los genotipos cambiarán a lo largo de las generaciones, es decir, los genotipos con mayor aptitud se harán más frecuentes. Así, las poblaciones compuestas preferentemente por individuos con menor aptitud irán desapareciendo en el tiempo y las poblaciones dominadas por individuos con mayor aptitud prevalecerán, en un proceso llamado selección natural (Orr, 2009).

Estrechamente ligada a la aptitud, la adaptación local se presenta cuando los individuos de un sitio tienen una aptitud promedio más alta que individuos provenientes de otros lugares (Hoban *et al.*, 2016). Ésta puede entenderse como un proceso donde la variación genética se crea mediante la recombinación de alelos y mutaciones, y luego, producto de la selección natural a lo largo de gradientes ecológicos, las frecuencias alélicas poblacionales cambian. En los períodos de relativa estabilidad ambiental, la selección natural no genera cambios en las frecuencias alélicas y la adaptación local se mantiene. Por el contrario, el flujo genético aumenta la diversidad alélica y disminuye la adaptación local, siendo la interacción entre estos dos factores evolutivos lo que determina, en gran medida, el grado de adaptación local de las poblaciones Leimu and Fischer (2008).

En un meta-análisis para determinar el grado de adaptación local en plantas, Leimu and Fischer (2008) analizaron el crecimiento de plantas locales y foráneas en ensayos de trasplante recíproco. Ellos demostraron que, en general, las plantas locales se desempeñaron significativamente mejor que las plantas foráneas en su sitio de origen en el 71.0% de los sitios estudiados. Sin embargo, las plantas locales se desempeñaron mejor que las plantas foráneas solo en el 45,3% de los 1032 pares de población comparados (definición estricta de adaptación local). Además, encontraron que la adaptación local es mucho más común para grandes poblaciones (más de 1000 individuos con flores) que para poblaciones pequeñas, para las cuales la adaptación local fue poco observada. Por otro lado, el grado de adaptación local fue independiente de la historia de vida de las plantas, la heterogeneidad espacial o temporal del hábitat y la escala geográfica. Estos resultados sugieren que la adaptación local es menos común en las poblaciones de plantas de lo que generalmente se supone. Además, estos hallazgos muestran el claro efecto positivo del tamaño efectivo poblacional sobre la capacidad de las poblaciones para hacer frente a entornos cambiantes.



La adaptación local en plantas puede ser afectada por múltiples factores, tanto abióticos como bióticos, los que pueden actuar a distintos niveles. Dentro de los factores abióticos, el sitio juega un papel preponderante (Eriksson *et al.*, 2006). La temperatura del suelo y la atmósfera, el potencial hídrico del suelo, la cantidad y frecuencia de la precipitación, el fotoperíodo y la radiación solar, son algunas de las variables de gran importancia en procesos fisiológicos básicos como la fotosíntesis, respiración, el ritmo de crecimiento, la dormancia y la reproducción (Taiz *et al.*, 2015). Por ejemplo, la sincronización entre las fases crecimiento-latencia y el clima local, se reconoció desde el principio como requisito para una forestación exitosa, y condujo a una práctica general de utilizar semillas forestales colectadas localmente (Aitken and Bemmels, 2016). Recientemente, el descubrimiento del reloj circadiano, un reloj molecular

de control genético, que permite a las plantas anticipar eventos de clima favorable para la fotosíntesis, entre otros mecanismos, ha entregado más evidencias de la adaptación local de las plantas al clima (de Dios *et al.*, 2017). Sin embargo, existen otros factores que influyen en la adaptación local. Como se dijo anteriormente, el flujo genético desde poblaciones adaptadas a climas divergentes, reduce la adaptación local, por lo tanto, los elementos del paisaje que imponen limitaciones a la dispersión de polen y propágulos, como la distancia geográfica, deben ser considerados al momento de modelar la divergencia adaptativa (Orsini *et al.*, 2013).

Otro factor que debe ser considerado por su indiscutible influencia en la divergencia adaptativa es la demografía y la estructura poblacional. Esto se refiere a que, en una avanzada de colonización, los fundadores son objeto de posteriores efectos de la adaptación local, disminuyendo el flujo genético desde otras poblaciones, lo que se traduce en una monopolización de estos fundadores (De Meester *et al.*, 2002). La monopolización supone que la adaptación local, combinada con una ventaja numérica de los primeros migrantes, contribuye a un fuerte efecto de prioridad de los individuos locales sobre los inmigrantes. En la medida en que los primeros fundadores posean suficiente variación genética para fomentar una rápida adaptación local, esto puede resultar en una estructura genética del paisaje que es, en gran medida, dictada por eventos de colonización (Orsini *et al.*, 2013). El efecto de los fundadores también puede ser causado por factores puramente ecológicos, principalmente por la ventaja numérica de los primeros migrantes (Boileau *et al.*, 1992). En este caso, cuando las poblaciones crecen rápidamente en tamaño después de la fundación de unos pocos individuos, la divergencia genética establecida durante la colonización es resistente al decaimiento por intercambio genético.

1.2 Cambio climático, desadaptación y conservación

El cambio climático antropogénico es definido como un cambio de clima que se atribuye directa o indirectamente a la actividad humana, que altera la composición de la atmósfera global y que se suma a la variabilidad natural del clima durante períodos de tiempo comparables (Pielke Jr, 2004). Es considerado una de las mayores amenazas contra la biodiversidad, y se espera que produzca cambios en la distribución de las especies, disminución de poblaciones y extinción, debido a que las comunidades biológicas no serán capaces de enfrentar estos cambios ambientales que ocurren a una velocidad nunca antes vista (Loarie *et al.*, 2009). La rápida tasa actual de cambio climático y el aumento de la variabilidad ambiental están alejando a las poblaciones locales de las condiciones climáticas a las que se adaptaron en el pasado (Jump and Penuelas, 2005; Vázquez *et al.*, 2017). Así, las relaciones genotipo-ambiente establecidas a lo largo de prolongados períodos de relativa estabilidad ambiental, se verán fracturadas debido a que las especies deberán afrontar cambios ambientales que acontecerán a una velocidad sin precedentes, lo que limita seriamente su capacidad de respuesta evolutiva.

El potencial de estas poblaciones para responder y adaptarse a nuevas condiciones ambientales impactará directamente en los procesos y servicios ecosistémicos a escala local y global, haciendo que algunos ecosistemas sean vulnerables a la transformación. Si bien algunas poblaciones pueden tolerar estas condiciones rápidamente cambiantes a través de la variación genética permanente (Brennan *et al.*, 2019), plasticidad fenotípica (Gárate-Escamilla *et al.*, 2019), procesos epigenéticos (Weigel and Colot, 2012), rescate evolutivo a través de la adaptación a nuevas condiciones (Gonzalez and Bell, 2013) o migración a ubicaciones dentro de sus nichos climáticos basados en modelos de nicho predictivo (Loarie *et al.*, 2009). Otras poblaciones no podrán superar las cambiantes condiciones locales y disminuirán en abundancia y/o se extinguirán, como consecuencia de la mala adaptación.

Los antecedentes previamente expuestos han generado una urgente necesidad de evaluar intervenciones humanas apropiadas, como la translocación de individuos con alelos adaptados al clima futuro, dentro y fuera de su distribución natural (Chen *et al.*, 2021). En este sentido, la migración asistida (MA) es una opción de conservación en respuesta al cambio climático, que consiste en el movimiento asistido de individuos adaptados al clima dentro o fuera del rango de las especies (Ste-Marie *et al.*, 2011). Los enfoques para la MA incluyen a) el movimiento de individuos resilientes dentro del rango de especies a áreas que han sido duramente afectadas por el cambio climático, conocido como flujo de genes asistido, o b) el movimiento de individuos fuera de los rangos actuales a nuevas áreas que se han convertido o se prevé que sean adecuadas, conocido como colonización asistida. El flujo genético asistido a veces se denomina una forma de evolución asistida, ya que podría acelerar la adaptación. Estas dos subcategorías de MA (es decir, flujo genético asistido y colonización asistida) pueden ser usadas indistintamente para referirse al movimiento controlado de individuos en respuesta a cambios ambientales (Aitken and Whitlock, 2013; Aitken and Bemmels, 2016). Teniendo en cuenta que la MA se centra explícitamente en el movimiento de individuos para ayudar en la respuesta al cambio climático, las aplicaciones de MA tienen un propósito distinto de introducir diversidad genética adaptativa a las poblaciones receptoras en contraste con otras translocaciones de conservación como el rescate genético que tiene como objetivo reducir la carga genética en pequeñas poblaciones afectadas por la endogamia (Hoffmann *et al.*, 2021). Sin embargo, el cambio climático puede causar una reducción en el tamaño de la población y escenarios complicados adicionales con respecto a la variación genética, lo que sugiere que las opciones de manejo no son mutuamente excluyentes y pueden necesitar integrar varias estrategias de mezcla genética para lograr los resultados de conservación previstos. En este ámbito es donde las herramientas de la genómica tienen mucho que aportar, dada su capacidad de informar y modelar la adaptación local.

El desarrollo de programas de conservación es un proceso complejo que requiere de un marco conceptual para su desarrollo. El primer componente en MA es establecer criterios para la caracterización de las posibles poblaciones donantes y receptoras. Existen diferentes perspectivas sobre la selección de las poblaciones de origen preferidas para la MA: una que maximiza la diversidad genética general y la otra que se centra en los alelos adaptados al clima para mejorar la capacidad de adaptación en las poblaciones receptoras (Chen *et al.*, 2021). La perspectiva para maximizar la diversidad es que los intentos de emparejar la fuente y el receptor deben evaluar cuidadosamente la variación genética en las poblaciones fuente para garantizar una mayor capacidad de adaptación de las poblaciones establecidas con MA (Hoffmann *et al.*, 2021); esta capacidad de adaptación puede ayudar a las poblaciones establecidas por MA frente a nuevos factores estresantes. Algunos investigadores abogan por el uso de fuentes mixtas para maximizar la variación genética y mejorar la capacidad de adaptación (Burton and Burton, 2002), y otros han demostrado que los individuos con mayor heterocigosidad tienen una mayor probabilidad de supervivencia después de la translocación (Scott *et al.*, 2020). Sin embargo, la selección de los individuos fundadores únicamente para maximizar la diversidad genética podría reducir la aptitud promedio de la población fundadora (Edmands, 2007). Por esta razón, la adaptación local y la diferenciación regional, como las distintas unidades biogeográficas, son consideraciones importantes para la selección de las poblaciones donantes (Weeks *et al.*, 2011).

Los criterios primarios de las poblaciones donantes para MA deben incluir evidencia de adaptación local a climas fuera del rango preferido (p. ej., temperatura extrema o poca precipitación), alta frecuencia de alelos adaptados al clima (p. ej., > 50% de frecuencia de alelos para alta tolerancia térmica; establecido a través de jardines comunes o genes candidatos de genotipificación), un tamaño de población relativamente moderado a grande para evitar la endogamia, y datos de tendencias que indiquen un censo (N_c) y un

tamaño efectivo (N_e) estables. Los criterios para las poblaciones receptoras potenciales incluyen poblaciones locales que experimentan niveles de desadaptación que pueden conducir a la extinción, frecuencia alélica adaptativa actualmente baja (p. Ej., <10%), tamaños de población relativamente pequeños a moderados y tendencias a la disminución de N_c y N_e (Chen *et al.*, 2021).

Una vez que las poblaciones potenciales se han caracterizado con los criterios anteriores, es necesario hacer coincidir la fuente con las poblaciones receptoras que serían las más apropiadas para las translocaciones. En particular, las poblaciones bien emparejadas deberían tener niveles actuales de flujo de genes que sean relativamente modestos, ya que las poblaciones con un flujo de genes bajo pueden estar en riesgo de padecer depresión exógena cuando los individuos se mueven. Por el contrario, es probable que las poblaciones con altos niveles de flujo génico incorporen naturalmente alelos adaptados localmente y, por lo tanto, la intervención es innecesaria. Para facilitar la predicción de la depresión por exogamia, los marcos de decisión desarrollados previamente (por ejemplo, (Frankham *et al.*, 2011; Hoffmann *et al.*, 2021) pueden usarse para incluir parámetros adicionales de estado taxonómico, intensidad de selección, diversidad genética y tiempo de aislamiento. Además, el clima histórico y actual de la población de origen debe ser similar al nuevo clima que experimentará actualmente o se prevé que experimente la población receptora. Por ejemplo, se esperaría que una población de *Araucaria araucana* que normalmente experimenta un clima con grandes temperaturas extremas sea un donante adecuada para las poblaciones que tienen un clima proyectado en el futuro de estas características (Varas-Myrik *et al.*, 2022).

1.3 Los marcadores moleculares y la divergencia adaptativa

El concepto básico expuesto anteriormente establece que, para poder desarrollar una metodología de conservación basada en el conocimiento de la divergencia adaptativa, debemos establecer una relación

entre genotipo (adaptativo) y ambiente, para lo cual el proceso de genotipado es crucial. Esto permitiría transferir de manera efectiva los alelos adaptados al clima desde la población donante hacia las poblaciones receptoras. Por lo tanto, la identificación de loci involucrados en la selección natural es un requisito para los estudios genómicos orientados a conservación adaptativa (Hoban *et al.*, 2016)

Para poder detectar loci selectivos, es necesario contar con una matriz de marcadores moleculares de tamaño lo suficientemente grande y que estén dispersos por todo el genoma (Balkenhol *et al.*, 2009). Sin embargo, la mayoría de las especies no tienen información genómica previa para localizar adecuadamente marcadores en todo el genoma (especies no modelo). Por lo tanto, la ausencia del requisito de conocimiento previo del genoma y una alta densidad de cobertura, son características indispensables de los marcadores moleculares a utilizar en la genómica aplicada a conservación y cambio climático (Rajora, 2019).



Dos tipos de marcadores moleculares son adecuados estudios genómicos de la adaptación local. Los marcadores de tipo I, que no tienen información de secuencia de ADN, tales como los polimorfismos de longitud de fragmentos amplificados (AFLP, por sus siglas en inglés) y repeticiones de secuencia simple (SSR, por sus siglas en inglés). Los marcadores de tipo II, al contrario, contienen información sobre la secuencia de ADN, tales como los polimorfismos de un solo nucleótido (SNP). Aunque los marcadores de tipo I pueden permitir detectar loci potencialmente responsables de la adaptación mediante la detección de loci típicos (outliers) y el análisis de asociación ambiental (EAA, por sus siglas en inglés), la función genética de esos loci no se puede validar fácilmente y, por lo tanto, podrían ser falsos positivos. Los marcadores de tipo II presentan varias ventajas, ya que contienen información sobre la secuencia de ADN, lo que permite anotar y mapear estos marcadores en el genoma. Enfoques de secuenciación de

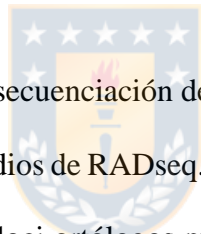
representación reducida (RRS), por ejemplo, RAD-seq (Bay *et al.*, 2018; Varas-Myrik *et al.*, 2022) y ARN-seq (Bay *et al.*, 2017), se han aplicado recientemente para detectar loci putativamente bajo selección y aplicaciones para la gestión de la conservación. RRS reduce el costo de secuenciación, mantiene una alta cobertura del genoma y no requiere información genómica a priori. Por lo tanto, el uso de RRS es beneficioso en la investigación genómica orientada a la conservación (Brauer *et al.*, 2016). La mayoría de los métodos RRS se basan actualmente en las plataformas de secuenciación de *Illumina*, que tienen la ventaja de una alta precisión y rendimiento, y la desventaja de que las lecturas son cortas. En resumen, el uso de marcadores de tipo II para realizar estudios genómicos puede facilitar la validación indirecta de los loci potencialmente responsables de la adaptación.

1.3.1 RADseq

La secuenciación de ADN asociada a sitio de restricción (RADseq, Baird *et al.* (2008)) consiste en un submuestreo de regiones homólogas putativas del genoma de varios individuos con el objetivo de descubrir y genotipar miles de SNPs que pueden usarse para estudios evolutivos, filogenómicos y de estructura poblacional, entre otros (Andrews *et al.*, 2016). RADseq es particularmente informativo para estudios enfocados en especies no modelo, ya que permite descubrir de manera costo efectiva miles de SNPs de todo el genoma en cientos de individuos mediante un ensamble *de novo* de las lecturas (Davey *et al.*, 2011).

El término RADseq se utilizó originalmente para describir un método en particular (Baird *et al.*, 2008), pero posteriormente se adoptó para referirse a una variedad de técnicas relacionadas que se basan en enzimas de restricción para determinar el conjunto de loci a secuenciar. Estos métodos también se agrupan a veces bajo el término técnicas de "genotipificación por secuenciación" (GBS, Baird *et al.* (2008)). Al

igual que con RADseq, el término GBS se utilizó originalmente para describir un método específico (Elshire *et al.*, 2011); sin embargo, este término es menos descriptivo que RADseq, el cual captura la característica definitoria de estos métodos, es decir, el uso de enzimas de restricción para obtener la secuencia de ADN en un conjunto de loci de todo el genoma. Las enzimas de restricción se han utilizado durante mucho tiempo para muestrear loci en todo el genoma y para generar información sobre la variación a nivel de población (ej. AFLP; Vos *et al.* (1995)). La característica principal de RADseq es que puede generar datos de secuencia adyacentes a un gran número de sitios de corte de restricción (Baird *et al.*, 2008). Los loci RADseq pueden ocurrir en cualquier área del genoma (es decir, regiones codificantes y no codificantes), y los individuos dentro o entre especies estrechamente relacionadas generalmente comparten la mayoría de los loci debido a la conservación de los sitios de corte (Andrews *et al.*, 2016).



Como en otros enfoques que se basan en la secuenciación de alto rendimiento, el procesamiento de datos es uno de los principales desafíos de los estudios de RADseq. Los cientos de miles de lecturas cortas deben ensamblarse en alelos putativos y luego en loci ortólogos putativos, para lo cual se deben hacer algunas suposiciones. Se han desarrollado varios paquetes de software para ensamblar loci ortólogos y escribir posiciones variantes a partir de datos de secuenciación de representación reducida, por ejemplo ipyrad (Eaton and Overcast, 2020) y Stacks (Catchen *et al.*, 2013). ipyrad es más rápido que su sucesor pyRAD y se puede ejecutar a través de un solo comando (Thallinger, 2017). Este pipeline posee un algoritmo general que consta de siete pasos: (i) demultiplexar lecturas sin procesar en muestras o, alternativamente, importar datos que ya han sido demultiplexados; (ii) control de calidad, filtrado y recorte por contaminación del adaptador; (iii) identificar copias leídas del mismo locus dentro de las muestras utilizando agrupación *de novo* o mapeo de referencia. Para los datos “*paired-end*”, el método *de novo* primero fusiona pares de lectura con VSEARCH (Rognes *et al.*, 2016) antes de agruparlos, y luego se

imputan los *indels* durante un proceso de alineación con aberturas que realiza MUSCLE (Edgar, 2004). Para los ensamblajes con genoma de referencia, las lecturas “*paired-end*” se asignan a la referencia para producir alineaciones con espacios, y los pares de coincidencias que se asignan con una orientación incorrecta o en múltiples ubicaciones (es decir, como parálogos) se descartan; (iv) estimación conjunta de la tasa de error de secuenciación y heterocigosidad de la muestra; (v) realizar llamadas de base de consenso y llamadas de haplotipos dentro de las muestras; (vi) identificar la ortología en las muestras mediante agrupaciones *de novo* o posiciones mapeadas de referencia; y (vii) aplicar una ronda final de filtrado y recorte a los loci ensamblados, generando estadísticas de ensamblaje informativas y escribiendo archivos de salida en numerosos formatos útiles para análisis posteriores (Eaton and Overcast, 2020).

Finalmente, es posible señalar que las técnicas RADseq tienen un enorme poder y versatilidad para el genotipado en los estudios de conservación adaptativa. Los protocolos a aplicar, sin embargo, diferirán en su idoneidad dependiendo de las preguntas de investigación, los sistemas de estudio y el presupuesto, y deben elegirse cuidadosamente. Cuando se implementan de forma correcta, los enfoques RADseq brindan vías eficientes, flexibles y rentables para aplicar el poder de las tecnologías de secuenciación de última generación y así obtener nuevos conocimientos sobre preguntas ecológicas, evolutivas y relacionadas con la conservación (Andrews *et al.*, 2016).

1.4 La genómica del Paisaje y el modelamiento de la Divergencia Adaptativa

Tradicionalmente, la adaptación local en distintas especies se ha evaluado y modelado a través de experimentos de jardín común o ensayos de trasplante recíproco, correlacionando la variación en los rasgos cuantitativos con las características climáticas de las poblaciones de origen (Wang and Summers, 2010). Sin embargo, estos estudios podrían tomar mucho tiempo y ser muy costosos para el desarrollo de

planes de conservación, particularmente en especies arbóreas longevas, ya que los rasgos a considerar deben incluir, además del crecimiento, el éxito reproductivo de los individuos (Anderson and Wadgyar, 2020). Un enfoque alternativo es modelar la adaptación local a través de marcadores moleculares, que con frecuencia se corresponden con rasgos cuantitativos, medidos en jardín común (De Kort *et al.*, 2014; Steane *et al.*, 2014). Este enfoque sin precedentes, denominado **Genómica del Paisaje**, tiene como objetivo revelar la relación entre la variación genética adaptativa (variación entre genomas individuales producto de la selección natural) y la variación ambiental, y permite dilucidar los efectos generalmente confundidos de los distintos factores que influyen en la divergencia adaptativa, estudiando los procesos evolutivos y las bases moleculares de la adaptación (Li *et al.*, 2017).

La genómica del paisaje se caracteriza por no requerir datos fenotípicos sobre rasgos adaptativos, los que pueden ser laboriosos y costosos de recopilar. Por lo tanto, difiere de otras estrategias clásicas destinadas a descubrir las bases genéticas de la adaptación, como el análisis de QTL, el mapeo de asociación, o los estudios de genética cuantitativa (Stinchcombe and Hoekstra, 2008). Un prerrequisito de la genómica del paisaje es examinar muchos loci (cientos o miles), dispersos en el genoma de muchos individuos para descubrir regiones genómicas bajo selección, basado en asociaciones entre marcadores moleculares del tipo SNP (Single Nucleotide Polymorphism) y variables ambientales putativamente influyendo en el proceso de selección natural (Storz, 2005). La base científica de esta relación puede entenderse como una línea de tiempo de 200 años. En esta progresión, un paisaje recientemente deforestado se regenera naturalmente con un grupo genéticamente diverso de plántulas de una o más especies de árboles forestales. Al principio, los árboles con varias combinaciones de genes que confieren más o menos tolerancia al frío, se distribuyen de manera relativamente uniforme, como cabría esperar de las semillas que vuelan hacia el valle con fuertes vientos de las poblaciones cercanas. Los árboles con combinaciones alélicas que

confieren tolerancia al frío, eventualmente se vuelven dominantes en las elevaciones más altas, mientras que los árboles que carecen de tolerancia al frío, se tornan dominantes en las elevaciones más bajas. La genómica del paisaje puede identificar estas relaciones, determinando factores ambientales que impulsan la selección natural (Forester *et al.*, 2021).

Los métodos estadísticos para la detección de loci potencialmente adaptativos, se dividen en dos grupos principales: análisis de asociación ambiental (EAA, por sus siglas en inglés) y los métodos de diferenciación de loci atípicos (outlier loci). Los EAA buscan identificar alelos cuyas frecuencias tienen correlaciones inusualmente altas con ciertas variables ambientales, lo que sugiere que esos loci están involucrados en la adaptación local. Se han desarrollado varios métodos para buscar tales asociaciones (Hoban *et al.*, 2016), sin embargo, el desafío principal de estos métodos es despejar los efectos puramente selectivos de los efectos demográficos, para lo cual se han creado metodologías basadas en modelos mixtos, que consideran los efectos de estructura poblacional como efectos aleatorios, denominados factores latentes (Frichot *et al.*, 2013). Por otro lado, las pruebas de loci atípicos no utilizan variables ambientales, bajo el supuesto de que los alelos involucrados en la adaptación local ocurren con mayor frecuencia cuando mejoran el fitness y con menor frecuencia cuando la disminuyen. Por lo tanto, es posible identificar los loci involucrados en la adaptación local mediante la detección de alelos que muestran una diferenciación genética mayor que la media entre las poblaciones, para lo cual debe construirse primeramente una distribución de frecuencias nula o neutral, con la cual detectar frecuencias inusualmente altas o inusualmente bajas (Foll and Gaggiotti, 2008).

Una vez obtenida la matriz de loci selectivos, traducir esta información en predicciones espacialmente explícitas de la variación adaptativa, es un desafío que ha traído el desarrollo de nuevos métodos de

modelado espacial (Schoville *et al.*, 2012). Idealmente, tales métodos deben: (1) vincular datos genéticos y ambientales para caracterizar cómo las frecuencias alélicas adaptativas varían a lo largo de múltiples gradientes ambientales, a menudo correlacionados, y (2) proyectar estas relaciones genotipo-ambiente en el espacio y en el tiempo y predecir cómo la adaptación local puede verse alterada en escenarios de cambio ambiental (Fitzpatrick and Keller, 2015; Shryock *et al.*, 2015; Gugger *et al.*, 2018; Jia *et al.*, 2020; Varas-Myrik *et al.*, 2022). La genómica del paisaje ofrece dos enfoques para modelar espacialmente la divergencia adaptativa utilizando variables ambientales: un modelo no lineal basado en la distancia, el Modelo Generalizado de Disimilaridad (GDM; (Ferrier *et al.*, 2007)) y un enfoque de aprendizaje automático: Gradient Forests (GF; Ellis *et al.* (2012)) Ambos enfoques se pueden aplicar de manera sólida para analizar y mapear la variación genómica adaptativa cuyas aplicaciones van desde la delimitación de zonas de utilización de semillas (Shryock *et al.*, 2017) pasando por la predicción del riesgo de desadaptación producto del cambio climático (genetic offset, (Jia *et al.*, 2020; Varas-Myrik *et al.*, 2022)) hasta la conservación adaptativa (Chen *et al.*, 2021).

Finalmente, otro desafío de la genómica del paisaje es estimar el grado en que la variación genómica está influenciada por variables ambientales, geográficas, o demográficas. Esto se responde, entre otros métodos, mediante la aplicación de una técnica proveniente del análisis multivariado llamada análisis de redundancia basado en distancia (db-RDA). Esta técnica de ordenación sujeta a restricción, consiste básicamente en una regresión lineal múltiple entre una matriz de variables respuesta y una (o más) matrices de variables explicativas, seguido de un análisis de componentes principales en la matriz de valores estimados por la regresión (Legendre and Anderson, 1999). Generalmente, la matriz dependiente contiene frecuencias alélicas poblacionales, y las matrices independientes o explicativas pueden ser: (1) variables climáticas y de suelo; (2) distancia geográfica; y (3) la matriz de ancestría (Nadeau *et al.*, 2016). El

resultado, expresado a través del parámetro R² ajustado, entrega el porcentaje de varianza en la matriz de respuesta, explicada por cada matriz independiente (Legendre et al, 1999), además de sus distintas interacciones. Este tipo de análisis multivariado es más apropiado que las pruebas de Mantel (Orsini *et al.*, 2013), ya que permite captar la naturaleza compleja de los sistemas naturales midiendo la variación en múltiples variables bióticas y abióticas que interactúan a nivel de paisaje. La significancia de la proporción de varianza explicada, se determina mediante una prueba de permutación ejecutada sobre el db-RDA (Legendre and Legendre, 2012), de forma similar a un análisis de varianza.

1.5 Especie de estudio: *Araucaria araucana*

El género *Araucaria* incluye diecinueve especies, todas pertenecientes al hemisferio sur en distribuciones marcadamente disjuntas, desde Oceanía hasta América del Sur. La variedad de ambientes que habita implica que este género de coníferas ha tenido éxito desde un ambiente de selva tropical, donde generalmente se cree que las angiospermas han reemplazado a las coníferas, hasta un ambiente de tundra. Se estima que el éxito de este género en Australasia puede estar gobernado más por su capacidad para adaptarse a las condiciones ambientales que por su capacidad para competir eficazmente con las angiospermas (Armstrong and This, 2001). Incluso se ha señalado que los patrones relictuales de muchas araucarias en las masas terrestres Gondwánicas pueden no reflejar su rango climático potencial completo (Kershaw and Wagstaff, 2001).

Araucaria araucana (conocida comúnmente como Pehuén o Monkey Puzzle) es una icónica conífera milenaria del género *Araucaria*, que habita en Chile tanto en la Cordillera de la Costa como en la Cordillera de los Andes (37,5° a 38,5° y 37,5° a 39,7° de latitud sur, respectivamente, Varas-Myrik *et al.* (2022); y en Argentina (37 ° 45 'a 40 ° 20' latitud sur). Esta especie ha sido catalogada como "En peligro" en la Lista

de la UICN (Premoli *et al.*, 2013), principalmente debido a la deforestación histórica, invasiones de otras especies de coníferas introducidas, cosecha de semillas ilegal, e introducción de ganado. (Cóbar-Carranza *et al.* 2014; Premoli 2015; Tella *et al.* 2016). Su nicho ecológico actual, aunque relativamente pequeño en extensión, abarca una gran amplitud ambiental, tanto en altitud (664 - 1227 m), precipitación anual (1100 - 2219 mm) como en temperatura (6,1 - 9,6 ° C de temperatura media anual, McIntosh *et al.* (2022)). *A. araucana* es una especie dioica y polinizada por el viento, con marcada presencia de pulsos de semillación, conocido como “*masting*” (Sanguinetti and Kitzberger, 2008). Tanto la semilla como el polen son relativamente pesados y pueden no dispersarse a grandes distancias, aunque la dispersión de semillas puede ser asistida por loros Trichahue y otros animales (Tella *et al.*, 2016). Se ha informado de la reproducción vegetativa por brotes desde la raíz (Schilling y Donoso, 1976), pero se desconoce la importancia de este proceso para el mantenimiento y la expansión de las poblaciones (Veblen *et al.*, 2003).



Respecto de su historia biogeográfica, en los últimos millones de años, el ciclo continuo de expansión y contracción de la cobertura de hielo y el consiguiente cambio ambiental, han modificado intensamente el paisaje patagónico donde habitaba *Araucaria araucana* (Sersic *et al.*, 2011). Los Bosques Subantárticos, dominados por especies de *Nothofagus* y coníferas, asociados con las comunidades de Tundras magallánicas, habrían tenido un área de distribución más extendida que la actual durante la última edad glacial, en los sitios de baja altitud correspondientes al valle central, desde la Región de Los Lagos hasta la Región de la Araucanía (Villagrán, 2018). La hipótesis propuesta es que la fragmentación actualmente observada de esta comunidad estaría determinada por: (i) la capa de hielo en el borde este y sur del área durante el Último Máximo Glacial; (ii) la ocupación gradual y creciente de sitios de baja altitud por elementos de los bosques “nordpatagónicos” y “valdivianos”, en una expansión creciente durante el tardiglacial y el Holoceno temprano. Este proceso de recolonización habría causado el aumento altitudinal

de la comunidad glacial y el aislamiento en las cumbres de las montañas que se observa actualmente en *A. araucana*. Lo anteriormente descrito hace suponer una historia demográfica compleja, más cercana a un modelo de refugio glacial que al tradicional modelo continente-isla (Sersic *et al.*, 2011).

1.5.1 Estudios genéticos en *Araucaria araucana*.

Varios estudios sobre diversidad genética y estructura poblacional han sido publicados para esta especie. Su frecuencia ha aumentado en los últimos años debido al incremento en la accesibilidad tanto tecnológica como económica de la utilización de marcadores moleculares y secuenciación de ADN. Uno de los primeros estudios utilizó alcanos de cadena larga como marcadores genéticos. En este trabajo, Rafii and Dodd (1998), tomando muestras de cuatro poblaciones, encontraron una separación genética de la mayoría de los individuos de la población andina de la ladera occidental de alta elevación, de las otras 3 poblaciones, y en particular de la población andina de sombra de lluvia. Los datos indican que el aislamiento geográfico de la población costera de Nahuelbuta no ha resultado en la diferenciación genética esperada debido a la deriva genética. También observaron una correlación significativa entre la longitud de la cadena de los alcanos, y variables ambientales asociadas a disponibilidad hídrica, en uno de los primeros indicios de la presencia de un gradiente adaptativo de aridez. En un estudio posterior, Bekessy *et al.* (2002), utilizando marcadores del tipo RAPDs en 13 poblaciones de Chile y Argentina, encontraron que la mayor parte de la variación genética (87.2%) se concentró dentro de las poblaciones, y una menor pero significativa proporción, entre las poblaciones de las zonas costera (11.1%) y andina (1.77%). Las poblaciones más septentrionales de los Andes se agruparon y fueron diferentes de las poblaciones más australes de Chile y Argentina. En un estudio posterior, Martín *et al.* (2014), utilizando microsatélites, tomaron muestras de 12 poblaciones a lo largo de la distribución natural de *Araucaria araucana*, y encontraron una diferencia significativa entre los Andes y la Costa, detectando una barrera que separaba

las poblaciones de ambas cadenas montañosas como áreas máximamente diferenciadas y con un flujo genético mínimo entre ellas. A nivel local, la variación genética fue alta dentro de las poblaciones y baja entre ellas, concordante con estudios anteriores. En un estudio muy reciente, (Fuentes *et al.*, 2021), basados en 286 AFLP polimórficos, realizaron un análisis de estructura poblacional y buscaron loci putativamente bajo selección. Sus conclusiones principales fueron que la distribución actual de *Araucaria araucana* podría ser un remanente de una distribución pasada más extensa, reducida no solo debido a eventos glaciares y postglaciales, sino también debido a otros factores como el vulcanismo y la influencia humana. Además, encontraron correlación significativa en 41 loci atípicos con las variables ambientales de precipitación y temperatura, confirmando firmas de selección natural. Respecto de la estructura poblacional, encontraron 3 grupos ancestrales, al contrario del estudio anteriormente citado de Martin et al, quienes encontraron solamente 2 grupos.



Como acabamos de ver, los estudios genéticos desarrollados hasta ahora en *A. araucana* solamente han considerado la utilización de marcadores tipo I, con las desventajas propias de este tipo de marcadores, y sin las ventajas propias de los marcadores tipo II, que provienen de técnicas de secuenciación masiva. En este estudio se utilizarán marcadores de secuenciación RADseq, por lo que esperamos responder una mayor cantidad de preguntas de investigación y con una mayor precisión.

1.5.2 La muerte apical en *Araucaria araucana*.

En la actualidad, esta especie está siendo objeto de una muerte apical, muerte regresiva o “dieback”, que ha encendido las alertas del mundo científico y organismos públicos atingentes al mundo forestal y la conservación, debido a la severidad con que se ha presentado en algunas zonas de su distribución. Esta

situación fue oficialmente comunicada por CONAF en octubre del año 2015, quien compartió la información con su par de Argentina. Dentro de las causas posibles ha tomado fuerza la teoría de la infección provocada por un hongo endófito propio de la especie, que, dadas las condiciones de stress hídrico a las que actualmente está sometida la especie, se transforma en un agente patógeno.

Dentro de las pocas publicaciones que buscan científicamente el agente causante de esta enfermedad, se encuentra el de Velez *et al.* (2018), cuya conclusión principal fue de que la mortalidad de Araucarias reportada por Conaf el año 2015 se trató de un evento inusual dado que no hubo nada similar al menos en los últimos 25 años. También calificaron el evento como de baja intensidad con escasa mortalidad y que afectó principalmente a la regeneración. La causa más probable habría sido las altas temperaturas estivales y sequías más extremas de los últimos 1000 años, que facilitaron el desarrollo y extensión de hongos e insectos, lo que, sumado al stress hídrico, produjeron la mortalidad de los individuos. En el contexto del Cambio Climático, donde las anomalías climáticas con extremos de temperatura y sequía persistirá y se acentuará en las próximas décadas, es esperable que estos eventos vuelvan a ocurrir y con mayor frecuencia.

Otro estudio reciente, Puchi *et al.* (2021) realizaron un análisis retrospectivo para el período 1800-2017. Ellos midieron el crecimiento radial de individuos creciendo en distintos ambientes, además de rasgos anatómicos de la madera e isótopos estables $\delta^{18}O$, con el objetivo de evaluar las causas de muerte regresiva. Descubrieron que la severidad de la sequía, expresada como SPEI (índice estandarizado de precipitación y evapotranspiración), era el principal impulsor climático del crecimiento de los árboles. Por lo tanto, el aumento de la severidad e intensidad de la sequía tendrá un primer efecto perjudicial en la reducción del área foliar, lo que reducirá la conductividad hidráulica y limitará la capacidad de transporte

de agua y el crecimiento, y pondrá a los individuos de *A. araucana* en una situación cada vez más estresante y posiblemente mortal.

1.6 Aporte Científico de la Tesis

Las preguntas de investigación que se pretenden responder en este trabajo son a) ¿Existe una correlación significativa entre la variación genética adaptativa y la variación ambiental en *Araucaria araucana*?; b) ¿Cuáles poblaciones presentan mayor riesgo de desadaptación producto del cambio climático?; c) ¿Existen poblaciones que contengan genotipos pre-adaptados a condiciones climáticas futuras que hagan posible el flujo genético asistido hacia las poblaciones de mayor riesgo de desadaptación?

Las Hipótesis propuestas son que, al crecer esta especie en un gradiente ambiental amplio, existe una correlación significativa entre divergencia ambiental y divergencia adaptativa. Esto nos lleva a hipotetizar también que sus poblaciones sufrirán un gran nivel de desadaptación producto del cambio climático.

El principal objetivo de esta tesis es modelar la adaptación local en *Araucaria araucana*, de forma de evaluar el impacto del cambio climático en el riesgo de desadaptación de sus distintas poblaciones, además de aportar información científica para sustentar una estrategia de conservación a través de migración asistida, tanto en la forma de flujo genético asistido como de colonización asistida fuera de su hábitat natural. Dentro de los objetivos específicos están:

- Determina la estructura genética poblacional de *A. araucana*
- Detectar loci putativamente bajo selección en una matriz de SNPs (Marcadores tipo II)

- Modelar la Divergencia adaptativa de *A. araucana*
- Evaluar el efecto de los distintos factores evolutivos (ambiente, geografía y demografía) en la variación genética total y adaptativa de *A. araucana*
- Estimar el impacto del cambio climático en la adaptación local de la especie
- Aplicar el modelo ajustado para informar estrategias de Migración Asistida y Colonización asistida

BIBLIOGRAFÍA

- Aitken, S.N., Bemmels, J.B., 2016. Time to get moving: assisted gene flow of forest trees. *Evolutionary Applications* 9, 271-290.
- Aitken, S.N., Whitlock, M.C., 2013. Assisted gene flow to facilitate local adaptation to climate change. *Annual Review of Ecology, Evolution, and Systematics* 44, 367-388.
- Anderson, J.T., Wadgymar, S.M., 2020. Climate change disrupts local adaptation and favours upslope migration. *Ecology letters* 23, 181-192.
- Andrews, K.R., Good, J.M., Miller, M.R., Luikart, G., Hohenlohe, P.A., 2016. Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics* 17, 81-92.
- Armstrong, W.P., This, M.F., 2001. *The Araucaria Family: Past & Present*.
- Baird, N.A., Etter, P.D., Atwood, T.S., Currey, M.C., Shiver, A.L., Lewis, Z.A., Selker, E.U., Cresko, W.A., Johnson, E.A., 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PloS one* 3, e3376.
- Balkenhol, N., Waits, L.P., Dezzani, R.J., 2009. Statistical approaches in landscape genetics: an evaluation of methods for linking landscape and genetic data. *Ecography* 32, 818-830.
- Bay, R.A., Harrigan, R.J., Underwood, V.L., Gibbs, H.L., Smith, T.B., Ruegg, K., 2018. Genomic signals of selection predict climate-driven population declines in a migratory bird. *Science* 359, 83-86.
- Bay, R.A., Rose, N.H., Logan, C.A., Palumbi, S.R., 2017. Genomic models predict successful coral adaptation if future ocean warming rates are reduced. *Science Advances* 3, e1701413.

- Bekessy, S.A., Allnutt, T., Premoli, A., Lara, A., Ennos, R., Burgman, M., Cortes, M., Newton, A., 2002. Genetic variation in the vulnerable and endemic Monkey Puzzle tree, detected using RAPDs. *Heredity* 88, 243-249.
- Boileau, M.G., Hebert, P.D., Schwartz, S.S., 1992. Non-equilibrium gene frequency divergence: persistent founder effects in natural populations. *Journal of Evolutionary Biology* 5, 25-39.
- Brauer, C.J., Hammer, M.P., Beheregaray, L.B., 2016. Riverscape genomics of a threatened fish across a hydroclimatically heterogeneous river basin. *Molecular Ecology* 25, 5093-5113.
- Brennan, R.S., Garrett, A.D., Huber, K.E., Hargarten, H., Pespeni, M.H., 2019. Rare genetic variation and balanced polymorphisms are important for survival in global change conditions. *Proceedings of the Royal Society B* 286, 20190943.
- Burton, P.J., Burton, C.M., 2002. Promoting genetic diversity in the production of large quantities of native plant seed. *Ecological Restoration* 20, 117-123.
- Catchen, J., Hohenlohe, P.A., Bassham, S., Amores, A., Cresko, W.A., 2013. Stacks: an analysis tool set for population genomics. *Molecular ecology* 22, 3124-3140.
- Chen, Z., Grossfurthner, L., Loxterman, J.L., Masingale, J., Richardson, B.A., Seaborn, T., Smith, B., Waits, L.P., Narum, S.R., 2021. Applying genomics in assisted migration under climate change: framework, empirical applications, and case studies. *Evolutionary Applications* n/a.
- Davey, J.W., Hohenlohe, P.A., Etter, P.D., Boone, J.Q., Catchen, J.M., Blaxter, M.L., 2011. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics* 12, 499-510.
- de Dios, V.R., Gessler, A., Ferrio, J.P., Alday, J.G., Bahn, M., Del Castillo, J., Devidal, S., García-Muñoz, S., Kayler, Z., Landais, D., 2017. Circadian rhythms regulate the environmental responses of net CO₂ exchange in bean and cotton canopies. *Agricultural and Forest Meteorology* 239, 185-191.
- De Kort, H., Mergeay, J., Vander Mijnsbrugge, K., Decocq, G., Maccherini, S., Kehlet Bruun, H.H., Honnay, O., Vandepitte, K., 2014. An evaluation of seed zone delineation using phenotypic and population genomic data on black alder *Alnus glutinosa*. *Journal of applied ecology* 51, 1218-1227.
- De Meester, L., Gómez, A., Okamura, B., Schwenk, K., 2002. The Monopolization Hypothesis and the dispersal–gene flow paradox in aquatic organisms. *Acta oecologica* 23, 121-135.
- Eaton, D.A., Overcast, I., 2020. ipyrad: Interactive assembly and analysis of RADseq datasets. *Bioinformatics* 36, 2592-2594.

- Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research* 32, 1792-1797.
- Edmands, S., 2007. Between a rock and a hard place: evaluating the relative risks of inbreeding and outbreeding for conservation and management. *Molecular ecology* 16, 463-475.
- Ellis, N., Smith, S.J., Pitcher, C.R., 2012. Gradient forests: calculating importance gradients on physical predictors. *Ecology* 93, 156-168.
- Elshire, R.J., Glaubitz, J.C., Sun, Q., Poland, J.A., Kawamoto, K., Buckler, E.S., Mitchell, S.E., 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PloS one* 6, e19379.
- Eriksson, G., Ekberg, I., Clapham, D., 2006. *An Introduction to Forest Genetics. Second Edition.* Genetic Center. Department of Plant Biology and Forest Genetics, SLU. Box 7070, 750 07 Uppsala. Sweden. ISBN 91-576-7190-7.
- Ferrier, S., Manion, G., Elith, J., Richardson, K., 2007. Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. *Diversity and Distributions* 13, 252-264.
- Fitzpatrick, M.C., Keller, S.R., 2015. Ecological genomics meets community-level modelling of biodiversity: Mapping the genomic landscape of current and future environmental adaptation. *Ecology Letters* 18, 1-16.
- Foll, M., Gaggiotti, O., 2008. A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics* 180, 977-993.
- Forester, B.R., Landguth, E.L., Hand, B.K., Balkenhol, N., 2021. Landscape genomics for wildlife research. *Population genomics: wildlife*, 145-184.
- Frankham, R., Ballou, J.D., Eldridge, M.D., Lacy, R.C., Ralls, K., Dudash, M.R., Fenster, C.B., 2011. Predicting the probability of outbreeding depression. *Conservation Biology* 25, 465-475.
- Frichot, E., Schoville, S.D., Bouchard, G., François, O., 2013. Testing for associations between *loci* and environmental gradients using latent factor mixed models. *Molecular Biology and Evolution* 30, 1687-1699.
- Fuentes, G., González, F., Saavedra, J., López-Sepúlveda, P., Victoriano, P.F., Stuessy, T.F., Ruiz-Ponce, E., 2021. Assessing signals of selection and historical demography to develop conservation strategies in the Chilean emblematic *Araucaria araucana*. *Scientific reports* 11, 1-11.

- Gárate-Escamilla, H., Hampe, A., Vizcaíno-Palomar, N., Robson, T.M., Benito Garzón, M., 2019. Range-wide variation in local adaptation and phenotypic plasticity of fitness-related traits in *Fagus sylvatica* and their implications under climate change. *Global Ecology and Biogeography* 28, 1336-1350.
- Gonzalez, A., Bell, G., 2013. Evolutionary rescue and adaptation to abrupt environmental change depends upon the history of stress. *Philosophical Transactions of the Royal Society B: Biological Sciences* 368, 20120079.
- Gugger, P.F., Liang, C.T., Sork, V.L., Hodgskiss, P., Wright, J.W., 2018. Applying landscape genomic tools to forest management and restoration of Hawaiian koa (*Acacia koa*) in a changing environment. *Evolutionary Applications* 11, 231-242.
- Hoban, S., Kelley, J.L., Lotterhos, K.E., Antolin, M.F., Bradburd, G., Lowry, D.B., Poss, M.L., Reed, L.K., Storfer, A., Whitlock, M.C., 2016. Finding the genomic basis of local adaptation: pitfalls, practical solutions, and future directions. *The American Naturalist* 188, 379-397.
- Hoffmann, A.A., Miller, A.D., Weeks, A.R., 2021. Genetic mixing for population management: From genetic rescue to provenancing. *Evolutionary applications* 14, 634-652.
- Jia, K.H., Zhao, W., Maier, P.A., Hu, X.G., Jin, Y., Zhou, S.S., Jiao, S.Q., El-Kassaby, Y.A., Wang, T., Wang, X.R., 2020. Landscape genomics predicts climate change-related genetic offset for the widespread *Platycladus orientalis* (Cupressaceae). *Evolutionary Applications* 13, 665-676.
- Jump, A.S., Penuelas, J., 2005. Running to stand still: adaptation and the response of plants to rapid climate change. *Ecology Letters* 8, 1010-1020.
- Kershaw, P., Wagstaff, B., 2001. The southern conifer family Araucariaceae: history, status, and value for paleoenvironmental reconstruction. *Annual Review of Ecology and Systematics* 32, 397-414.
- Legendre, P., Anderson, M.J., 1999. Distance-based redundancy analysis: testing multispecies responses in multifactorial ecological experiments. *Ecological Monographs* 69, 1-24.
- Legendre, P., Legendre, L., 2012. *Numerical ecology*. Elsevier.
- Leimu, R., Fischer, M., 2008. A meta-analysis of local adaptation in plants. *PloS one* 3, e4010.
- Li, Y., Zhang, X.-X., Mao, R.-L., Yang, J., Miao, C.-Y., Li, Z., Qiu, Y.-X., 2017. Ten years of landscape genomics: challenges and opportunities. *Frontiers in Plant Science* 8, 2136.
- Loarie, S.R., Duffy, P.B., Hamilton, H., Asner, G.P., Field, C.B., Ackerly, D.D., 2009. The velocity of climate change. *Nature* 462, 1052-1055.

- Martín, M., Mattioni, C., Lusini, I., Molina, J., Cherubini, M., Drake, F., Herrera, M., Villani, F., Martín, L., 2014. New insights into the genetic structure of *Araucaria araucana* forests based on molecular and historic evidences. *Tree Genetics & Genomes* 10, 839-851.
- McIntosh, M., González-Campos, J., Demaree, P., Toro-Salamanca, O., Ipinza, R., Bustamante-Sánchez, M.A., Hasbún, R., Nelson, C.R., 2022. Trait variation between and within Andes and coastal mountain ranges in the iconic South American tree *Araucaria araucana* in Chile. *bioRxiv*, 2022.2001.2004.474828.
- Nadeau, S., Meirmans, P.G., Aitken, S.N., Ritland, K., Isabel, N., 2016. The challenge of separating signatures of local adaptation from those of isolation by distance and colonization history: The case of two white pines. *Ecology and Evolution* 6, 8649-8664.
- Orr, H.A., 2009. Fitness and its role in evolutionary genetics. *Nature Reviews Genetics* 10, 531-539.
- Orsini, L., Vanoverbeke, J., Swillen, I., Mergeay, J., De Meester, L., 2013. Drivers of population genetic differentiation in the wild: isolation by dispersal limitation, isolation by adaptation and isolation by colonization. *Molecular Ecology* 22, 5983-5999.
- Pielke Jr, R.A., 2004. What is climate change? *Energy & environment* 15, 515-520.
- Premoli, A., Quiroga, P., Gardner, M., 2013. Premoli, A., Quiroga, P. & Gardner, M. 2013. *Araucaria araucana*. The IUCN Red List of Threatened Species 2013: e.T31355A2805113. <https://dx.doi.org/10.2305/IUCN.UK.2013-1.RLTS.T31355A2805113.en>. Downloaded on 11 August 2021. In.
- Puchi, P.F., Camarero, J.J., Battipaglia, G., Carrer, M., 2021. Retrospective analysis of wood anatomical traits and tree-ring isotopes suggests site-specific mechanisms triggering *Araucaria araucana* drought-induced dieback. *Global Change Biology*.
- Rafii, Z.A., Dodd, R.S., 1998. Genetic diversity among coastal and Andean natural populations of *Araucaria araucana* (Molina) K. Koch. *Biochemical Systematics and Ecology* 26, 441-451.
- Rajora, O.P., 2019. *Population Genomics. Concepts, Approaches and Applications*. Cham: Springer.
- Rognes, T., Flouri, T., Nichols, B., Quince, C., Mahé, F., 2016. VSEARCH: a versatile open source tool for metagenomics. *PeerJ* 4, e2584.
- Sanguinetti, J., Kitzberger, T., 2008. Patterns and mechanisms of masting in the large-seeded southern hemisphere conifer *Araucaria araucana*. *Austral Ecology* 33, 78-87.

- Schoville, S.D., Bonin, A., François, O., Lobreaux, S., Melodelima, C., Manel, S., 2012. Adaptive genetic variation on the landscape: methods and cases. *Annual Review of Ecology, Evolution, and Systematics* 43, 23-43.
- Scott, P.A., Allison, L.J., Field, K.J., Averill-Murray, R.C., Shaffer, H.B., 2020. Individual heterozygosity predicts translocation success in threatened desert tortoises. *Science* 370, 1086-1089.
- Sersic, A.N., Cosacov, A., Cocucci, A.A., Johnson, L.A., Pozner, R., Avila, L.J., Sites Jr, J.W., Morando, M., 2011. Emerging phylogeographical patterns of plants and terrestrial vertebrates from Patagonia. *Biological Journal of the Linnean Society* 103, 475-494.
- Shryock, D.F., Havrilla, C.A., DeFalco, L.A., Esque, T.C., Custer, N.A., Wood, T.E., 2015. Landscape genomics of *Sphaeralcea ambigua* in the Mojave Desert: a multivariate, spatially-explicit approach to guide ecological restoration. *Conservation Genetics* 16, 1303-1317.
- Shryock, D.F., Havrilla, C.A., DeFalco, L.A., Esque, T.C., Custer, N.A., Wood, T.E., 2017. Landscape genetic approaches to guide native plant restoration in the Mojave Desert. *Ecological Applications* 27, 429-445.
- Ste-Marie, C., A. Nelson, E., Dabros, A., Bonneau, M.-E., 2011. Assisted migration: Introduction to a multifaceted concept. *The Forestry Chronicle* 87, 724-730.
- Steane, D.A., Potts, B.M., McLean, E., Prober, S.M., Stock, W.D., Vaillancourt, R.E., Byrne, M., 2014. Genome-wide scans detect adaptation to aridity in a widespread forest tree species. *Molecular Ecology* 23, 2500-2513.
- Stinchcombe, J.R., Hoekstra, H.E., 2008. Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. *Heredity* 100, 158-170.
- Storz, J.F., 2005. INVITED REVIEW: Using genome scans of DNA polymorphism to infer adaptive population divergence. *Molecular ecology* 14, 671-688.
- Taiz, L., Zeiger, E., Møller, I.M., Murphy, A., 2015. *Plant physiology and development*. Sinauer Associates Incorporated.
- Tella, J.L., Lambertucci, S.A., Speziale, K.L., Hiraldo, F., 2016. Large-scale impacts of multiple co-occurring invaders on monkey puzzle forest regeneration, native seed predators and their ecological interactions. *Global Ecology and Conservation* 6, 1-15.
- Thallinger, G., 2017. Comparison of ddRAD Analysis Pipelines. In. Graz University of Technology.

- Varas-Myrik, A., Sepúlveda-Espinoza, F., Fajardo, A., Alarcón, D., Toro-Núñez, Ó., Castro-Nallar, E., Hasbún, R., 2022. Predicting climate change-related genetic offset for the endangered southern South American conifer *Araucaria araucana*. *Forest Ecology and Management* 504, 119856.
- Vázquez, D.P., Gianoli, E., Morris, W.F., Bozinovic, F., 2017. Ecological and evolutionary impacts of changing climatic variability. *Biological Reviews* 92, 22-42.
- Veblen, T.T., Kitzberger, T., Raffaele, E., Lorenz, D.C., 2003. Fire history and vegetation changes in northern Patagonia, Argentina. In, *Fire and climatic change in temperate ecosystems of the western Americas*. Springer, pp. 265-295.
- Velez, M., Salgado Salomón, M., Marfetan, A., Tirante, S., Mattes Fernández, H., Avila, M., Szychowski, A., Viviani, D., Sanguinetti, J., 2018. Caracterización de la desecación del dosel y sanidad de *Araucaria araucana* en Argentina.
- Villagrán, C., 2018. Biogeografía de los bosques subtropical-templados del sur de Sudamérica. Hipótesis históricas. *Magallania (Punta Arenas)* 46, 27-48.
- Vos, P., Hogers, R., Bleeker, M., Reijans, M., Lee, T.v.d., Hornes, M., Friters, A., Pot, J., Paleman, J., Kuiper, M., 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic acids research* 23, 4407-4414.
- Wang, I.J., Summers, K., 2010. Genetic structure is correlated with phenotypic divergence rather than geographic isolation in the highly polymorphic strawberry poison-dart frog. *Molecular ecology* 19, 447-458.
- Weeks, A.R., Sgro, C.M., Young, A.G., Frankham, R., Mitchell, N.J., Miller, K.A., Byrne, M., Coates, D.J., Eldridge, M.D., Sunnucks, P., 2011. Assessing the benefits and risks of translocations in changing environments: a genetic perspective. *Evolutionary Applications* 4, 709-725.
- Weigel, D., Colot, V., 2012. Epialleles in plant evolution. *Genome biology* 13, 1-6.

CAPÍTULO II

Predicting climate change-related genetic offset for the endangered southern South American conifer *Araucaria araucana*

Running head: genetic offset of a southern South American tree species

Antonio Varas-Myrik¹, Francisco Sepúlveda-Espinoza¹, Alex Fajardo², Diego Alarcón³, Óscar Toro-Núñez⁴, Eduardo Castro-Nallar⁵ and Rodrigo Hasbún¹.

¹Laboratorio de Epigenética Vegetal, Departamento de Silvicultura, Facultad de Ciencias Forestales, Universidad de Concepción, Victoria 631, Concepción, Chile.

²Instituto de Investigación Interdisciplinario (I³), Universidad de Talca, Campus Lircay, Talca 3460000, Chile.

³Departamento de Ciencias Ecológicas, Universidad de Chile. Instituto de Ecología y Biodiversidad. Las Palmeras 3425, Ñuñoa, Santiago, Chile.

⁴Departamento de Botánica, Facultad de Ciencias Naturales y Oceanográficas, Universidad de Concepcion, Victoria 631, Concepción, Chile.

⁵Universidad Andres Bello, Center for Bioinformatics and Integrative Biology, Facultad de Ciencias de la Vida, Av. República 330, Santiago, Chile.

Type of contribution: Article

*Corresponding author: Antonio Varas-Myrik; email: avaras@udec.cl

ABSTRACT

Understanding adaptive genetic variation is key for predicting the evolutionary response of species and populations to climate change, decisively influencing management and conservation decisions. Landscape genomics provides a framework to disentangle the effects of local adaptation from those of geographic distance and demographic history, through genomic analysis and the modeling of genotype-environment relationships. This approach can inform how evolutionary forces shape the neutral and adaptive genetic structure, helping to identify those populations subject to a greater risk of maladaptation due to anthropogenic climate change, i.e., the “genetic offset”. Using restriction-site associated DNA sequencing (RAD-Seq) and more than 49,000 single nucleotide polymorphisms screened from locations of *Araucaria araucana* in Chile, we assessed the genetic structure and predicted the genetic offset of this emblematic tree species under two future climate scenarios. Using generalized dissimilarity modeling (GDM) we found that the temperature annual range was the most important variable shaping the observed patterns of adaptive divergence. Our results show that populations living in the piedmont of the southern Andes Mountain range are at the greatest risk of maladaptation, while populations living in the high elevation zones in the Andes Mountain range are at the lowest risk. This study constitutes an important tool for forestry management and conservation of *A. araucana* forests.

Key words: local adaptation, climate change, landscape genomics, population structure, conservation genetics, RAD-Seq.

2.1 INTRODUCTION

Anthropogenic climate change is expected to cause shifts in the distribution of species, population declines and extinctions. Species can cope the threats that arise from new environmental conditions in current habitats either migrating, adapting in situ or through phenotypic plasticity (Sultan, 2000; Hoffmann and Sgrò, 2011; Gougherty *et al.*, 2021). During adaptive responses, genetic variation increases, causing maladaptation in the short-term, but higher adaptability in the long-term (Urban, 2015; Wiens, 2016; Derry *et al.*, 2019). While maladaptation causes suboptimal population fitness, adaptation can usually restore populations to optimal fitness levels in the long-term. Nonetheless, the normal continuum from maladaptation to adaptation can be disrupted by rapid climate change, causing a breakdown in the genetic-environmental relationship faster than the species' capability to migrate or adapt *in situ*, with a consequent loss of fitness (Jump and Penuelas, 2005; Aitken *et al.*, 2008). The measure of the amount of adaptive genetic change between present and future climate conditions required for to maintain species current genetic-environmental association has been coined as the species' "genetic offset" index (Fitzpatrick and Keller, 2015). Among other applications, this risk index can improve conservation strategies by gathering spatially explicit genetic information that helps to build more resilient and better adapted populations (Fitzpatrick and Keller, 2015; Jia *et al.*, 2020). For example, the genetic offset can guide assisted gene flow programs, supporting the use of seed lots composed by local and non-local—but pre-adapted—genotypes, thus increasing adaptive diversity and resilience (Aitken and Whitlock, 2013; Aitken and Bemmels, 2016). The implementation of genetic offset indexes could represent a modern solution to support the local adaptation of natural populations under rapidly changing environments.

The first step to assessing genetic offset is to determine the strength and structure of adaptive genetic variation, i.e. the variation found between the genomes of individuals due to natural selection (Schoville *et al.*, 2012). Landscape elements that impose dispersal limitations (e.g., geographic distance, physical barriers, etc.) and modulate ecological processes (e.g., colonization) act concomitantly with natural selection to shape adaptive divergence (Hoban *et al.*, 2016). Elucidate the singular and combined effects of each of these factors in adaptive divergence imposes an analytical challenge, since the influence of neutral and Adaptive processes are usually confounded in natural populations (Ahrens *et al.*, 2018). Landscape genomics—a synthetic discipline that combines concepts and tools from population genetics, landscape ecology, geography and spatial statistics (Manel *et al.*, 2003; Holderegger and Wagner, 2008)—provides an robust theoretical framework to face this challenge through the identification of loci

putatively under selection, which are based on associations between genetic data and environmental variables hypothesized to drive selection, while controlling for neutral genetic structure (Rellstab *et al.*, 2015). To gain insight into the processes that shape adaptive divergence, redundancy analysis (RDA) can aid in the quantification of the specific and combined effects of landscape elements on adaptive divergence, partitioning genomic variation into components explained by geographic distance, ecological gradients, and their combinations (Capblancq and Forester, 2021). Thus, adaptive divergence can be modeled using this subset of adaptive loci as response matrix and landscape elements, mainly environment and geography, as predictors. These models, which we will call from now on "adaptive divergence models", can be harnessed to evaluate genetic offset under future conditions by measuring the distance between current and future predictions (Fitzpatrick and Keller, 2015; Holliday *et al.*, 2017).

While studies dealing with genetic offset have initially been implemented for different species across several ecosystems (Fitzpatrick and Keller, 2015; Bay *et al.*, 2018; Gugger *et al.*, 2018; Jia *et al.*, 2020), its implementation has not yet achieved a comprehensive use in several areas that require immediate assistance for conservation. One case is the temperate forests of southern South America, characterized by their complex evolutionary and biogeographic history, sustaining large ecosystems with high endemism, which are of great biological conservation value (Sersic *et al.*, 2011; Villagrán, 2018). While some interest has been placed on the use of genetic tools in conservation studies of Patagonian tree species, (Premoli *et al.*, 2000; Marchelli *et al.*, 2010; Souto *et al.*, 2015; Mattera *et al.*, 2020), landscape genomics has not been widely used to identify adaptive loci and to model adaptive divergence, (Martín *et al.*, 2014; Hasbún *et al.*, 2016), leaving this useful tools and its applications (e.g. genetic offset prediction) unexplored for species growing in temperate forests of southern South America. Given its potential to usefully inform conservation and forest management programs, the prediction of climate change-related genetic offset could aid in defining empirical criteria for conservation of valuable and threatened local tree species.

An emblematic and endangered tree species living in southern temperate forests is *Araucaria araucana* (Molina) K. Koch, commonly referred to as “Pewen” (indigenous name) or “Monkey Puzzle Tree” (English name) (Premoli *et al.*, 2013), which is distributed in restricted and fragmented patches of Chile and Argentina. In Chile, this species occurs in both the Costa and Andes Mountain ranges (37.5° to 38.5° and 37.5° to 39.7° south latitude, respectively). Recent studies in population genetics of *A. araucana* have described weak patterns of population structure, most likely as a result of bird-mediated seed dispersal

and the human consumption of its starch-rich seeds (Martín *et al.*, 2014; Tella *et al.*, 2016). Currently, this species is being increasingly affected by a dieback called “*Araucaria leaf damage*” (Velez *et al.*, 2018; Puchi *et al.*, 2021). Provided that we are dealing with an iconic species, this dieback has prompted public and private conservation efforts, which have included assisted migration. However, there is a dearth of knowledge regarding the genetic-environmental relationship, which could scientifically inform, for example, the most suitable reception areas for assisted migration (or assisted relocation) programs and assisted gene flow programs, reducing the risk of maladaptation in novel and changing environments. With the purpose to provide empirical information for conservation and forest management programs, the main goal of this study was to predict the genetic offset of populations of *A. Araucana* due to climate change. Besides, two complementary objectives were addressed: (a) to assess *Araucaria araucana*’s population genetic structure, and (b) to evaluate the effect of ancestry, environment and geographic distance on whole and adaptive genetic variation.

2.2 MATERIALS AND METHODS

2.2.1 Sampling, RAD-Seq library preparation and sequencing

Tissue samples of *A. araucana* were collected from 12 different locations in both the Coastal and Andes Mountain Ranges in Chile (Fig. 1, Table 1), to cover the range of environmental variation of its habitat. For each location, we traversed a 2 km transect where sampled trees were separated by at least 100 m. Eleven to 20 trees were sampled per transect; we selected fresh, healthy-looking leaves from the tips of the branches. Trees 3 to 5 m in height were selected to study trees of roughly the same age group. Samples were maintained in sterile conical tubes stored in coolers with ice packs until arriving to the laboratory, where they were stored at -80 °C until DNA extraction. Tissue samples were washed sequentially with 1.5 g/liter Captan (PubChem CID 8606), 70% ethanol, 1% sodium hypochlorite, and sterile water to remove epiphytic microbes. Then, plant material was ground manually before being flash frozen for 1 min (with liquid nitrogen). After one cycle of tissue disruption in a TissueLyser II for 1 min at 30 hz, samples were frozen in liquid nitrogen again to repeat the disruption step. Fifty milligrams of disrupted plant material were used for DNA extraction, using the DNeasy PowerPlant Pro kit (Qiagen).

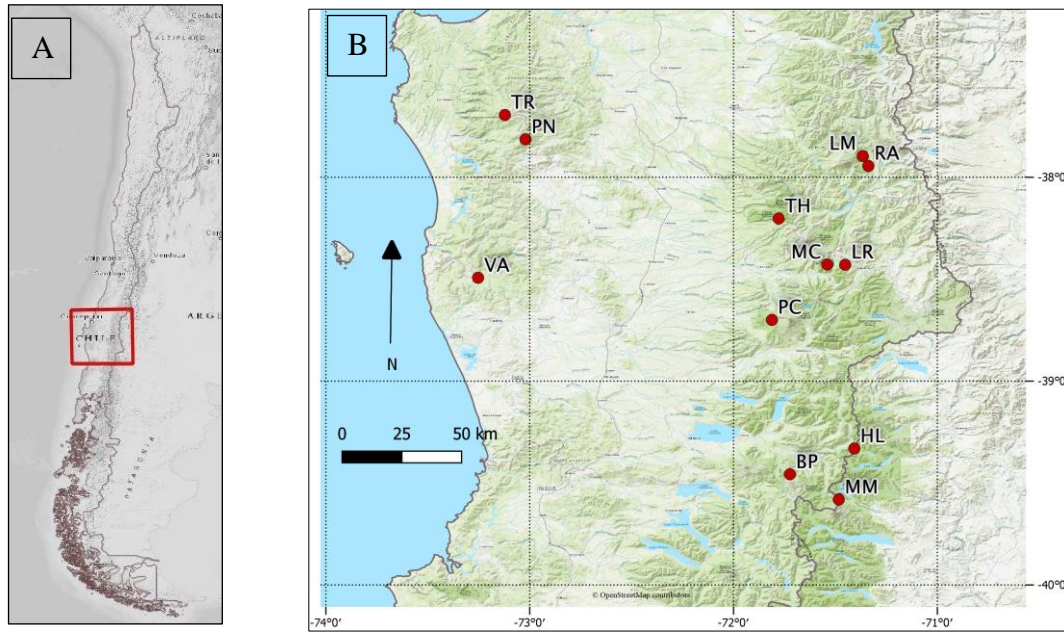


Figure 1. Sampling locations map of *A. araucana*. Panel A): regional map. The red box indicates sampling area. Panel B): detailed map. Red dots show sampling locations and its code.

DNA was quantified in a Qubit 3.0 instrument (Thermo Fisher Scientific) using the Qubit double-stranded DNA (dsDNA) high-sensitivity (HS) assay kit. Genomic DNA from 190 individuals was extracted using a standardized protocol (Doyle, 1991). Libraries for RAD-Seq were prepared by digesting DNA with a *PstI* restriction enzyme. Sequencing was carried out by Floragenex Inc. (Portland, USA), using Illumina HiSeq2500 technology. Reads were filtered and processed with ipyrad v.0.9.58, a pipeline usually employed in phylogenetic and population genetic analyses given its handling of variable indels during the clustering process (Eaton and Overcast, 2020).

Table 1. Sampling locations, geographic coordinates (degrees of south latitude and west longitude), sample size (*N*) and Mountain Range of the 12 *A. araucana* sampled sites.

Map Code (Fig. 1)	Site Name	latitude	longitude	N (Valid)	Mountain Range
BP	Bosque Pehuén	-39.46	-71.72	12	Andes
HL	Hualalafquén	-39.33	-71.41	7	Andes
LM	La Mula	-37.90	-71.37	11	Andes
LR	Las Raíces	-38.43	-71.45	14	Andes
MC	Malalcahuelo	-38.43	-71.54	8	Andes
MM	Mamuil Malal	-39.58	-71.48	10	Andes
PC	Conguillío	-38.70	-71.81	9	Andes
PN	Parque Nahuelbuta	-37.81	-73.02	14	Costa
RA	Ralco	-37.94	-71.34	14	Andes
TH	Tolhuaca	-38.20	-71.78	9	Andes
TR	Trongol	-37.69	-73.12	11	Costa
VA	Villa Araucarias	-38.49	-73.25	15	Costa

The data were assembled *de novo*, following the program default parameters, which considered filters for readings with an average *phred* value lower than 33, barcode removal, and a clustering threshold of 85%. The clusters obtained were selected with a ratio of 0.5 maximum heterozygosity per locus, 0.2 Single Nucleotide Polymorphism (SNP) per locus, two maximum alleles per site, and 8 indels per locus. The final consensus were filtered to retain those present in at least 4 individuals per locus. Subsequently, the data set was additionally filtered with VCFtools software (Danecek *et al.*, 2011). Sites with heterozygosity greater than 50%, minor allele frequency less than 5%, and missing data greater than 50%, were excluded. Individuals with more than 50% missing data were also discarded. After these filtering steps, a genotypic matrix of 134 individuals and 49,122 biallelic SNPs was obtained and later used for population structure analyses. In addition, a second matrix of 18,988 SNPs was obtained by taking only one SNP per fragment, to conduct the detection of loci putatively under selection, since multiple polymorphic fragments can cause problems associated with an unreliable estimation of linkage among markers (Hoban *et al.*, 2016). Sequence data can be found in NCBI's Sequence Read Archive database under BioProject PRJNA634877.

2.2.2 *Environmental predictors*

To detect environmental drivers of adaptive divergence, we considered a comprehensive set of 25 edaphoclimatic predictors, which were later subjected to a variable selection procedure to build a set of few meaningful and uncorrelated variables. This initial dataset considered 19 bioclimatic variables and elevations from the WorldClim version 1.4 database (Hijmans *et al.*, 2005), representative of the average

conditions for the years 1960-1990. In addition, five soil variables obtained from the Soilgrids 2.0 database (de Sousa *et al.*, 2020) were considered. Then, we conducted a two-step variable reduction procedure (Fitzpatrick *et al.*, 2011).

First, we reduced the full set of variables by selecting ten bioclimatic and soil variables that minimized correlation ($r < 0.7$), retaining those variables of correlated pairs that were, in our opinion, the most biologically significant. These included: annual mean temperature ($^{\circ}\text{C}$), isothermality (%), precipitation of the wettest month (mm), precipitation of the driest month (mm), temperature annual range ($^{\circ}\text{C}$), precipitation seasonality (%), elevation (m), cationic exchange capacity (CEC, mmol kg^{-1}), total soil nitrogen (cg kg^{-1}), and soil pH. This set, which includes annual trends, seasonality, limiting environmental factors and soil variables, was employed for the environmental association analysis to detect potentially adaptive loci, which are described below.

Second, to fit the adaptive divergence model, we further reduced this set of ten variables by using a backward-elimination technique (Fitzpatrick *et al.*, 2011), as implemented in the *gdm.varImp* function in the *gdm* R-package (Fitzpatrick *et al.*, 2021). Briefly, this algorithm started removing less important variables, assessing the significance of each through permutation tests. This procedure continued until the difference in deviance between the model with and without the variable became significant and no more variables could be discarded without decreasing the model performance. After this step, temperature annual range ($^{\circ}\text{C}$), precipitation of the wettest month (mm), precipitation of the driest month (mm), elevation (m) and soil pH were retained.

To project the bioclimatic variables used to fit the adaptive divergence model to future climate conditions, we considered four scenarios composed by two Representative Concentration Pathways (RCP4.5 and RCP8.5), reflecting moderate and extreme future conditions, respectively; and two projection periods (2050 and 2070). In order to account for uncertainty in climate projections (Sanderson *et al.*, 2015), we chose two general circulation models (GCM) for each of the four scenarios: ACCESS1-0 (Bi *et al.*, 2013) and HadGEM2-ES (Jones *et al.*, 2011), from the fifth phase of the Coupled Model Intercomparison Project (CMIP5), which were ensemble using average. Future climate data were obtained from the Worldclim 1.4 (Hijmans *et al.*, 2005). Lastly, the genetic offset was mapped using a mask corresponding to previous niche-based species distribution model reported for *A. araucana* (Alarcón and Cavieres, 2015), covering an area of 28.576 km^2 . All raster data had a resolution of 30 arc-second.

2.2.3 Building the adaptive divergence model

The spatially explicit distance-based adaptive divergence model was fitted prior to a population structure analysis and detection of potentially adaptive loci, detailed below.

2.2.3.1 *Population structure.* To evaluate the population structure, the genetic ancestry coefficients and the number of ancestral populations (K), the sNMF algorithm (scattered non-negative matrix factorization algorithm (Frichot *et al.*, 2014)) was used as in the *LEA* R-package (Frichot and François, 2015). This approach provides estimates of ancestry coefficients through least squares and incorporates a cross-entropy criterion to evaluate the model fit. K values from 1 to 5 were tested with 20 repetitions, selecting the optimal K value as the one corresponding to the run showing the lowest cross entropy.

2.2.3.2 *Identification of potentially adaptive loci.* Potentially adaptive loci were identified following a consensus approach between two environmental association analysis methods (EAAs). The first one was a latent factor mixed model (LFMM), which identifies loci showing significant correlations with environmental gradients, while controlling for neutral population structure by integrating K as a random component (Frichot *et al.*, 2013). This model was fitted using the *lfmm_ridge* function in the *lfmm2* R-package (Caye *et al.*, 2019). As the number of latent factors, K , was considered as the optimal number of ancestral populations identified through sNMF. The second EAA was the redundancy analysis (RDA), a multivariate constrained ordination technique, frequently used in landscape genomic analyses (Hecht *et al.*, 2015; Capblancq *et al.*, 2018; Shryock *et al.*, 2020). RDA is an extension of multiple regression modelling to multivariate response data (Legendre and Legendre, 2012), which has been shown to be more effective in detecting potentially adaptive loci than other techniques (Capblancq *et al.*, 2018; Forester *et al.*, 2018). To perform RDA, a biallelic genotypic matrix was used as the response matrix Y , and a set of environmental (climate and soil) variables was used as the predictor matrix X . After this constrained ordination step, we employed the methodology implemented in the *pcadapt* R-package to find outlier loci (Luu *et al.*, 2017), based on the Mahalanobis distance. For both methods (LFMM and RDA), a false discovery rate of 5% was adopted.

2.2.3.3 *Fitting adaptive divergence model.* To model adaptive divergence, a generalized dissimilarity model (GDM) (Ferrier *et al.*, 2007) was fitted at the individual level, using all of 134 individuals

described in section 2.1. GDM is a non-linear distance-based approach that models genotype-environmental associations and adapts to a variable rate of change in allele frequencies along environmental gradients. For accomplished this, we used the *gdm* R-package (Fitzpatrick *et al.*, 2021), adopting 6 knots per spline for all predictors. As response variable, a distance matrix containing pairwise Jaccard's dissimilarity index among individuals was used, which was obtained by applying the *vegdist* function of the *vegan* R-package (Oksanen *et al.*, 2020) on the biallelic matrix containing loci putatively under selection. As predictors, the set of environmental variables retained after the backward elimination process, in addition to the geographic distance, were used. Once the model was fitted, and in order to identify specific zones where genetic offset could be higher, we first created a continuous adaptive divergence map (Fig. S4) obtained according to Fitzpatrick and Keller (2015). Then, we grouped this continuous raster into six adaptive homogeneous zones by using the *clara* function of the *cluster* R-package (Maechler *et al.*, 2021). These zones (Fig. 4) represent areas with similar adaptive allelic compositions and were used as spatial units of genetic offset quantification.

2.2.4 Distance-based redundancy analysis

For better understanding the structure of whole and Adaptive genomic variation, we assessed the degree to which both are influenced by environmental variables, geographic distance and genetic ancestry. To achieve this goal, we perform a distance-based redundancy analysis (db-RDA) (Legendre and Anderson, 1999) at the individual level using both the Adaptive and the full set of SNPs. To do this, we use the function *varpart* function of the *vegan* R-package (Oksanen *et al.*, 2020), which partitioned the variation of genomic dissimilarities in components explained by three explanatory tables (environment, ancestry and geography) and their combinations. To evaluate the significance of each partition, functions *capscale* and *anova* from *vegan* R package were applied running 999 permutations. For assessing the effect of geographic distance, we utilized a principal coordinates of neighborhood matrices (PCNMs), also known as Moran's eigenvector maps (MEM). PCNMs are a set of orthonormal variables calculated through eigenvalue decomposition of a spatial weighting matrix based on geographic coordinates (Legendre and Legendre, 2012). PCNMs were calculated by using the *pcnm* function of the *vegan* R-package. The environmental matrix employed for db-RDA was the same as used for selective loci detection, which was scaled and centred.

2.2.5 Genetic offset predictions

We performed the genetic offset predictions by using the *predict.gdm* function of the *gdm* R-package (Fitzpatrick *et al.*, 2021), which directly calculates the adaptive distance among two environmental raster stacks (present and future), and outputs a single raster layer containing genetic offset values. We assumed soil characteristics were invariable over time, so they had no influence on the genetic offset prediction. Given that two RCPs were evaluated in 2 different periods, the result was a set of four genetic offset maps. Each of these four rasters was overlaid onto the adaptive raster zones described earlier, and a pixel-level data frame containing genetic offset, area in km² and adaptive zone number was assembled for summarize the results.



2.3 RESULTS

2.3.1 *Adaptive divergence model*

2.3.1.1 *Population Structure*

The sNMF algorithm showed a minimum cross entropy value with the parameter $K = 2$, which was used as a surrogate for the number of ancestral populations in the subsequent analyses. The ancestry coefficients revealed a clear differentiation between the populations in the Andes and the Costa Mountain Ranges (Fig. 2). Regarding ancestral admixtures, the Andes populations exhibited homogeneous levels of admixture, close to 20% of cluster G1 and 80% of cluster G2. In contrast, coastal populations showed heterogeneous levels of admixture. While individuals from locations PN and TR accounted for about 60% of cluster G1 and 40% of cluster G2, individuals from VA exhibited admixture levels about to 90% of cluster G1 and 10% of cluster G2.

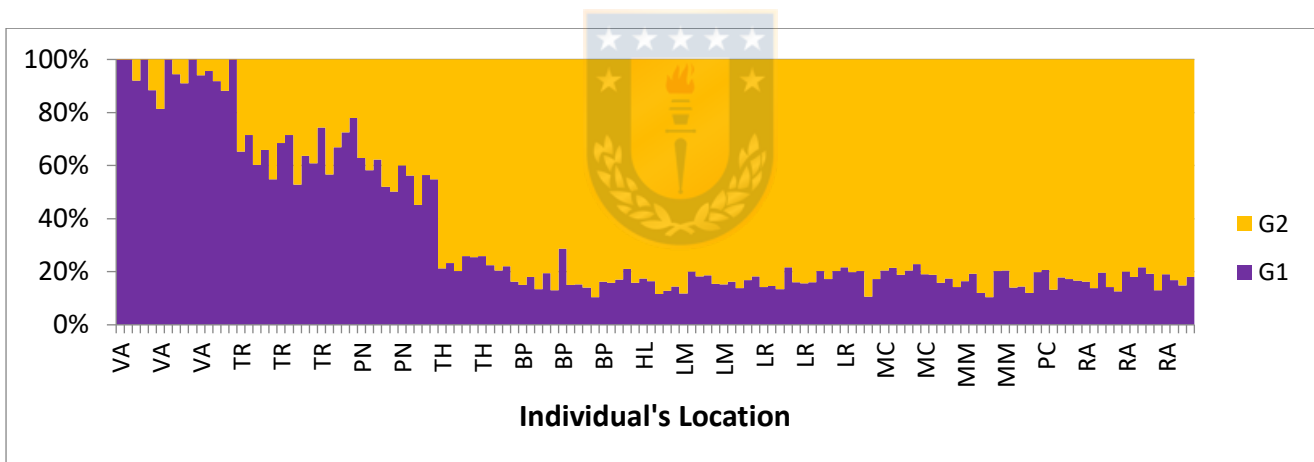


Figure 2. Population memberships ($K = 2$) of the 134 individuals of *A. Araucana*, based on sNMF algorithm. Vertical bars represent the percentage of individual's genome that belonged to each of the two ancestral groups (G1 and G2). Note that individuals from the VA location represent a distinctive class of admixture, showing a high presence of G1. Individuals from TR and PN constituted a second class of admixture, with a relatively balanced presence of G1 and G2, and the remaining individuals from Andes Mountain Range made up a third admixture class, with a low presence of G1.

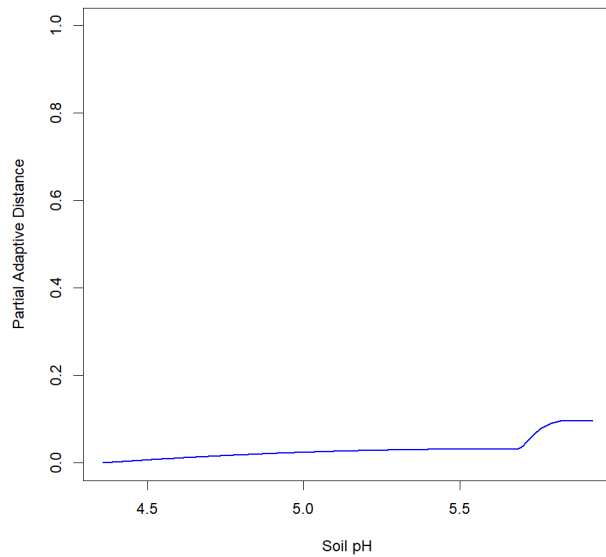
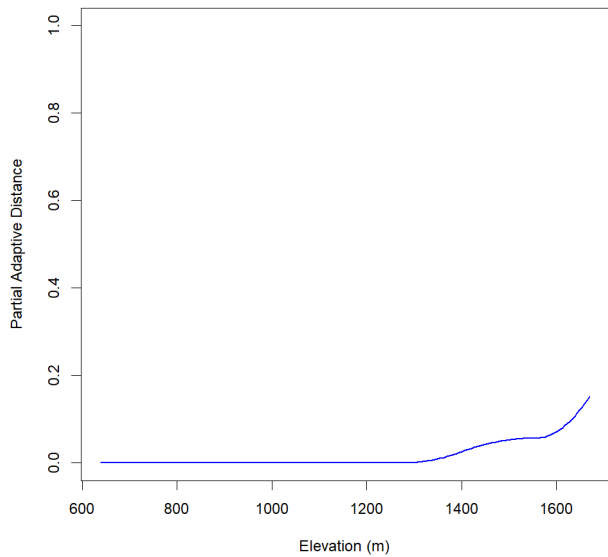
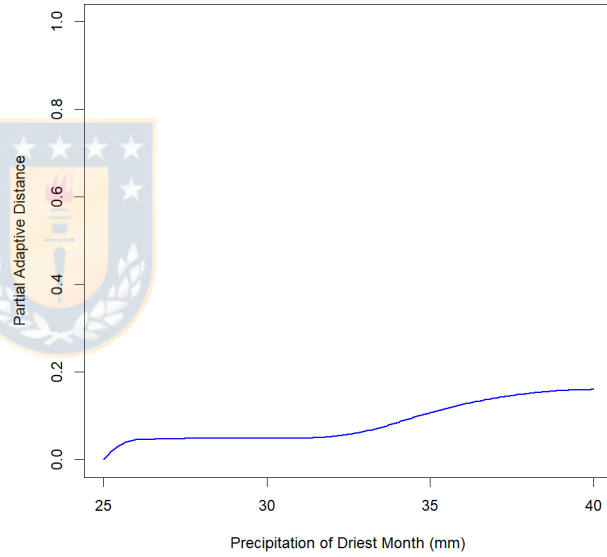
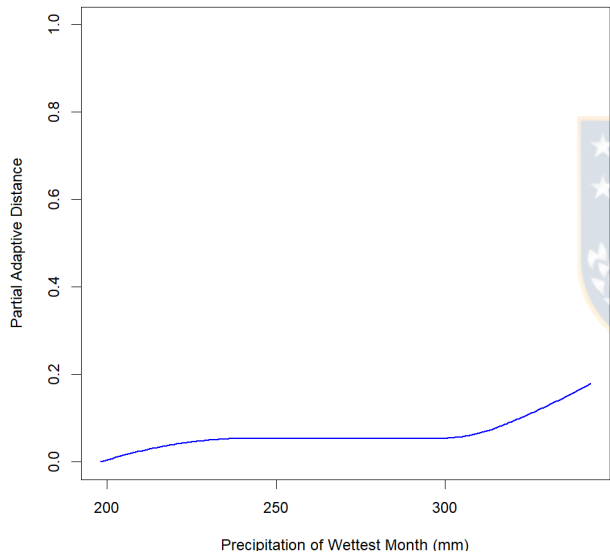
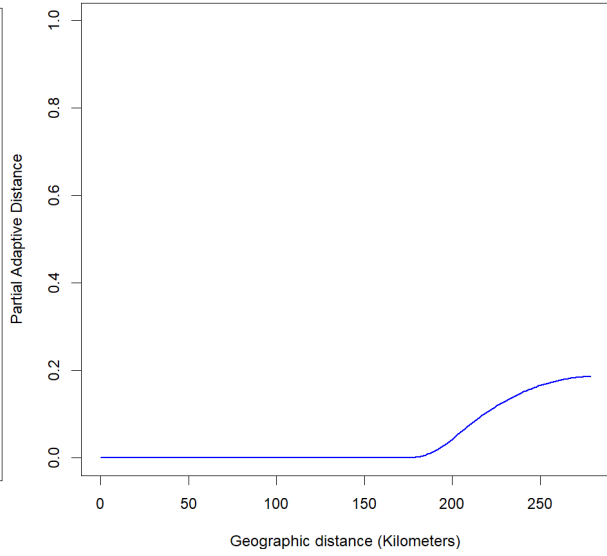
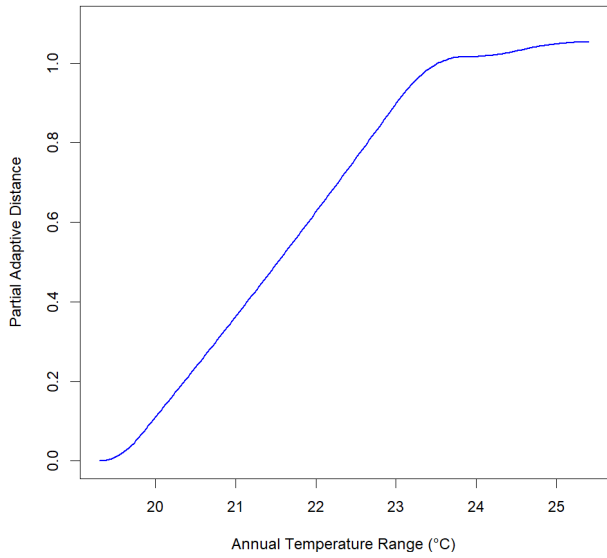


Figure 3. GDM *I*-splines for each environmental predictor and geographic distance. The maximum height of each curve indicates the total amount of change in allele frequencies associated with that predictor (variable importance). The shape of each curve indicates how the rate of change in allele frequencies varies along the predictor gradient.

2.3.1.2 Potentially Adaptive loci

After EAAs modelling, the LFMM detected 1,113 loci as potentially Adaptive. The predictor that returned the highest number of adaptive loci was temperature annual range, followed by soil pH and cationic exchange capacity. On the other hand, RDA returned 2,760 loci as potentially under selection. Observing the projection of each SNP in the RDA space, we were able to detect associations between adaptive SNPs and environmental predictors (Fig. S3). The RDA1 axis, which explained almost 50% of the variation, was strongly correlated with the temperature annual range, cationic exchange capacity and elevation, while the RDA 2 axis, which explained a minor amount of variation (9.7%), was mainly associated with the annual mean temperature and precipitation seasonality. Finally, 896 loci were detected as potentially adaptive by both methods.



2.3.1.3 Adaptive divergence model (GDM) fit

The GDM explained 85.1% of the variation in individual adaptive genetic distance. The maximum height of each *I*-Spline (Fig. 3) indicates the total amount of change in allele frequencies associated with that predictor (variable importance). Temperature annual range proved to be the most important predictor, followed by geographic distance, precipitation of the wettest month, precipitation of the driest month, elevation and soil pH. Regarding the shape of the *I*-Splines (Fig. 3), the temperature annual range exhibited a linear pattern between 19 and 23 °C, beyond which it showed a sharp decrease in slope, suggesting a plateau behaviour after this threshold. Geographic distance showed a different pattern: between 0 and 160 km. its importance was null, after which its significance increased steadily. Adaptive zones map showed a differentiation among Andes and Costa Mountain ranges (Fig. 4). While almost the entirety of zone 3 is located in the Costa Mountain range and central valley, the remaining zones were located in the Andes Mountain Range, following a pattern of diagonal bands oriented from the northeast to the southwest. This suggest that environmental drivers that are adaptively significant have a wider range in the Andes Mountain range than the Costa Mountain range.

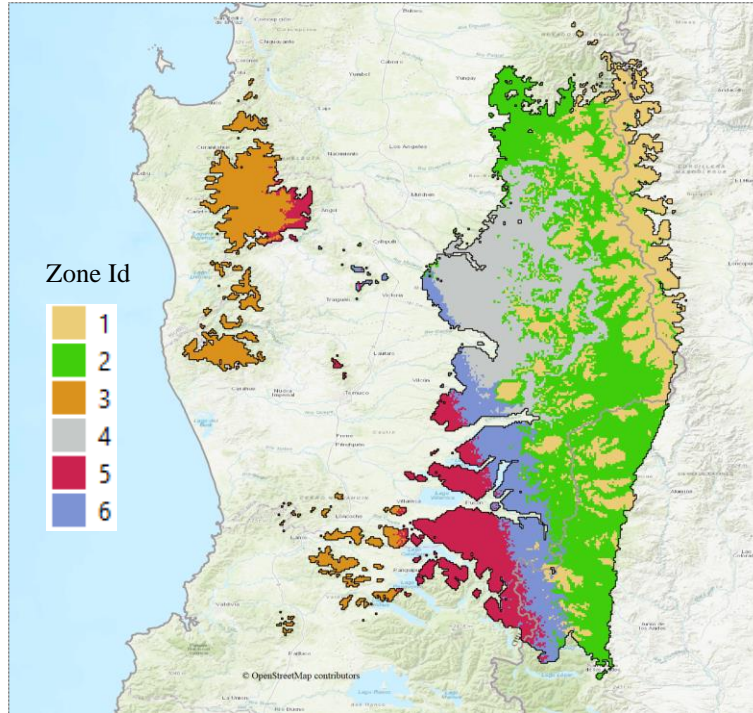


Figure 4. Map showing six zones of similar adaptive allelic compositions for *A. araucana*, built with the K-means clustering algorithm applied to GDM transformed predictor's data frame. These zones were used as genetic offset quantification units.

Combined Fractions	All SNPs			Adaptive SNPs		
	Adj. R ²	P(>F)		Adj. R ²	P(>F)	
F~env.	0.12379	0.001	***	0.50592	0.001	***
F~geo.	0.12396	0.001	***	0.50247	0.001	***
F~anc.	0.09680	0.001	***	0.43595	0.001	***
Individual Fractions						
F~env. (geo. + anc.)	0.00735	0.001	***	0.00838	0.001	***
F~geo. (env. + anc.)	0.00853	0.001	***	0.00953	0.002	**
F~anc. (env. + geo.)	0.00062	0.137	ns	0.00000	0.466	ns
F~env.+geo. anc.	0.02129			0.06163		
F~geo.+anc. env.	0.00103			0.00003		
F~env.+anc. geo.	0.00203			0.00464		
F~env. + anc. + geo	0.09312			0.43128		
Total explained	0.13397			0.51549		
Total unexplained	0.86603			0.48451		
Total	1.00000			1.00000		

Table 2. Distance-based redundancy analysis (db-RDA). The amount of variance explained by each component or its combination is expressed as adjusted R², in addition to its significance level. **Note:** F: Dependent matrix of individual genetic distances; the db-RDA tests are in the form of independent matrices F ~ | covariate matrices.

*env.: environment; geo.: geographic distance; anc.: ancestry. Total Explained: Adjusted R² total of individual fractions. Abbreviation: ns: not significant. * p <.05; ** p <.01; *** p <.001.*

2.3.2 Distance based redundancy analysis

The results of the db-RDA, using both adaptive (896) and full set of SNPs (18,889), revealed a high degree of environmental adaptation (Table 2) in *A. araucana*, showing a great amount of total variance explained by adaptive SNPs (51.5%), regarding all SNPs (13.4%). Concerning the combined fractions in the full set of SNPs, environment and geography explained 12.4% of the total variance, while genetic ancestry explained only 9.7%. For the adaptive SNPs, environment and geography explained a much higher amount (50.6% and 50.2%, respectively), while the lowest value was returned by genetic ancestry (43.6%). Regarding individual fractions (i.e., the variation explained exclusively by a factor, excluding combined effects), the amount of variance explained was low, with the greatest proportion of variance being confounded among the three factors. As in the combined fractions, genetic ancestry explained the lowest amounts of variance in all SNPs, and in adaptive SNPs (0.06 and 0.00% respectively), while geographic distance explained the highest amounts, 0.85 and 0.95%, respectively.

Table 3. Genetic offset values for each adaptive zone, according to RCP and year of projection. Values correspond to area-weighted average of each of 6 adaptive zones overlaid onto 4 genetic offset raster (2 RCPs and 2 Years). Genetic offset values correspond to individual adaptive genomic distance computed as Jaccard index

RCP	ZONE	YEAR	
		2050	2070
4.5	1	0.51	0.51
	2	0.53	0.53
	3	0.61	0.55
	4	0.57	0.57
	5	0.66	0.64
	6	0.61	0.61
	Sub-Total	0.56	0.55
8.5	1	0.51	0.51
	2	0.53	0.54
	3	0.60	0.65
	4	0.59	0.61
	5	0.67	0.70
	6	0.62	0.64
	Sub-Total	0.57	0.58
TOTAL	0.56	0.57	

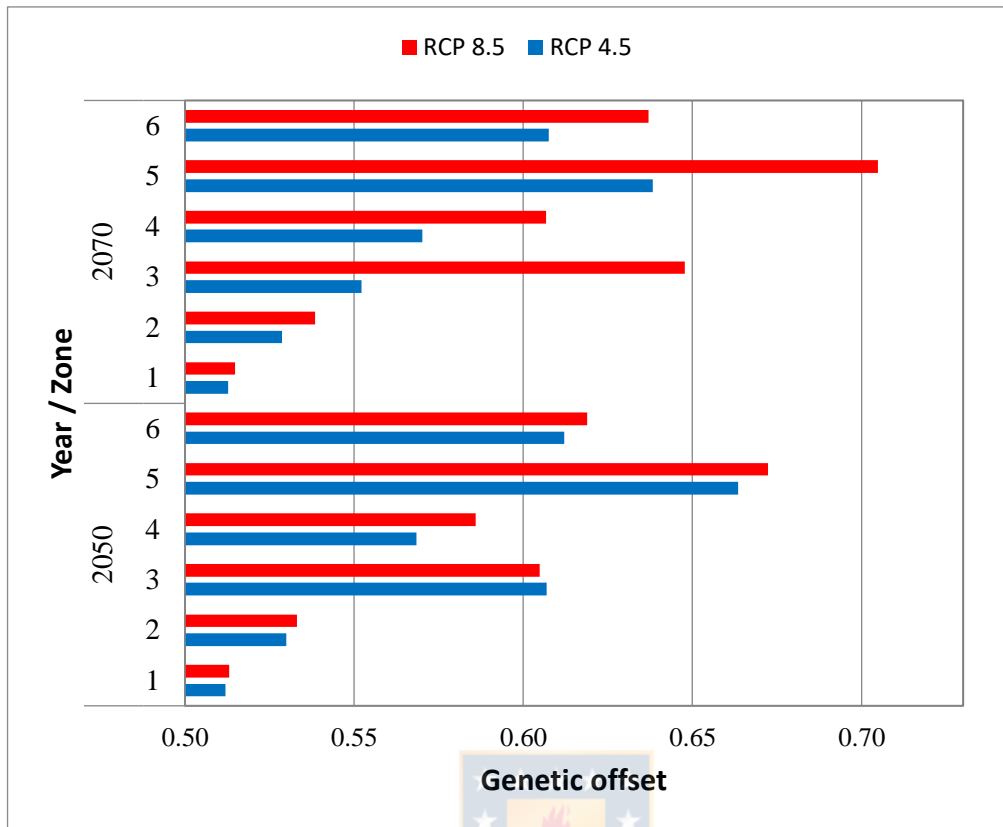


Figure 5. Genetic offset predictions per year and adaptive zone, for each RCP. Bars represent an area-weighted average of each of 6 adaptive zones overlaid onto 4 genetic offset raster (2 RCPs and 2 Years). Genetic offset values (axis x) correspond to individual adaptive genomic distance computed as Jaccard index.

2.3.3 Genetic offset

Predicted mean genetic offset was 0.56 for 2050 and 0.57 for 2070 (Table 3). The RCP8.5-2070 scenario showed the highest genetic offset (0.58), while RCP4.5-2070 exhibited the lowest risk of maladaptation (0.55) (Fig 5). Zone 5, corresponding to the piedmonts of the Andes and the southern border of the range of *A. araucana* (Fig. 6), showed it will be most affected by genetic offset in both 2050 and 2070 in RCP4.5 (0.66 and 0.64, respectively), similar to RCP 8.5 (0.67 and 0.70, respectively). High elevation zones in the Andes Mountain range (Zone 1) appeared to be at the lowest risk (0.51 for both years and RCPs). Zone 3, containing mainly Coastal Mountain range populations, presented an intermediate risk level. It can also be observed that the Bío-Bío river basin is a zone of higher genetic offset than the surrounding areas (Fig. 6).

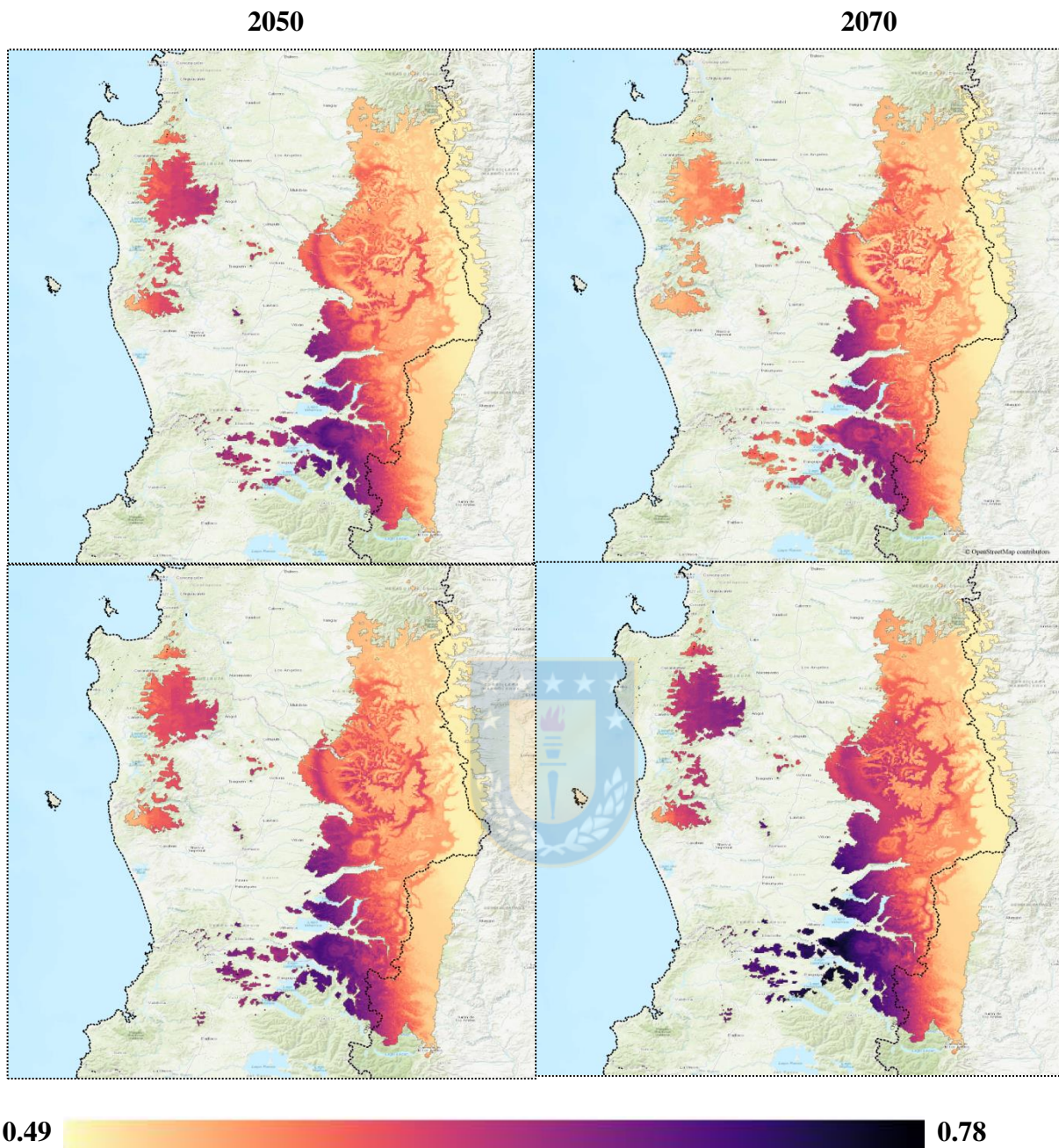


Figure 6. Continuous genetic offset maps of *A. araucana* for the years 2050 and 2070 under the RCP4.5 (up) and RCP8.5 (down) scenarios. Black and yellow represent high and low offset levels, respectively.

2.4 DISCUSSION

This study confirms the presence of adaptive genetic variation in the genome of *A. araucana*, which was highly correlated with environmental variation. Based on our adaptive divergence model (GDM), we predicted that populations living in the piedmonts of southern Andean distribution will be most affected by maladaptation due to future climatic conditions and such populations should be prioritized for conservation. To the best of our knowledge, this is the first study dealing with genetic offset, not only for this emblematic species, but in the southern South American region, which may provide valuable information to ecological restoration, assist gene flow and migration programs.

2.4.1 Population Structure

Our population structure analysis revealed two ancestral clusters in sampled populations ($K=2$), both depicting a disjoint distribution of genetic groups between Costa and Andes Mountains ranges (Fig. 2). This pattern is concurrent with biogeographic studies which propose an important role of glaciation cycles shaping current species distribution and genetic structure of Patagonian temperate forests. Long glacial periods of areal expansion alternated with shorter periods of retraction and isolation during the interglacial warmer periods, would have produced the present disjoint distribution of *A. araucana* populations (Villagrán, 2018). According to pollen and charcoal records, colonization route could have been from glacial refuges located in the Costa Mountain range toward the Andes Mountain range (Sersic *et al.*, 2011; Nanavati *et al.*, 2020).

2.4.2 Local Adaptation

We found a high degree of local adaptation in *A. Araucana* whose main environmental driver was temperature annual range. Coincidentally, this environmental driver has been also detected as most important in a common garden trial study (McIntosh *et al.*, Unpublished results) reinforcing the power of landscape genomic approach to detect and to model adaptive divergence. In addition, we detected limiting factors influencing adaptive divergence (i.e., precipitation of the wettest month and precipitation of the driest month) probably reflecting physiological trade-offs associated with an aridity gradient. This is not surprising, since water stress has formerly been shown to be an active driver of adaptation and possibly the most globally selective force acting on trees, causing mortality in vegetation from arid zones to the understory of tropical rainforests (Engelbrecht *et al.*, 2007; Allen *et al.*, 2010). More importantly, a global

population decline has been documented in Araucariaceae species during the Cretaceous, which has been attributed to climatic drying and cooling (Kershaw and Wagstaff, 2001).

One key functional trait of this species is its remarkable isohydry, a physiological feature present principally in the basal clades of Gymnosperms (Pinaceae, Araucariaceae). Isohydric species reduces stomatal conductance rates as soil water potential decreases and atmospheric conditions dry, maintaining a strict control over the leaf water potential and closing stomata by steadily increasing abscisic acid (ABA) concentrations, showing a high sensibility to drought (Zimmer *et al.*, 2016). Furthermore, according to Brodribb *et al.* (2014), tolerance to xylem failure due to water shortage in Araucariaceae was the least among conifer families. This costly and inefficient strategy to cope with water shortage stress could enhance intra-specific natural selection in abiotic gradients associated with drought in *A. araucana*. In a work that directly supports this hypothesis, Rafii and Dodd (1998), using foliar epicuticular wax alkanes as genetic markers, found longer chains of this wax in populations living in dryer habitats, consistent with a genetic adaptation to resist water loss in more arid conditions (Papú *et al.*, 2021). In addition, Puchi *et al.* (2021) conducted a retrospective analysis for the 1800-2017 period, measuring tree ring growth, wood anatomical traits and $\delta^{18}\text{O}$ stable isotopes to assess dieback causes. They found that drought severity, expressed as SPEI (Standardized Precipitation-Evapotranspiration Index), was the main climatic driver of tree growth.

Regarding of local adaptation structure given by db-RDA analyses, we found that a large amount of the explained among-individual genetic variation co-varied with the effects of environment, geography and ancestry (Table 2, Figure S5) and a small, yet significant, proportion of genetic divergence was explained exclusively by environmental conditions in both full and Adaptive sets of SNPs. Such confounding effects of local adaptation with neutral processes are frequent and not mutually excluding in natural landscapes (Engelbrecht *et al.*, 2007). Two main drivers could explain such confounding effects: a) the presence of selective gradients spatially correlated with demographic history (e.g., postglacial colonization along climatic gradients); and b) the simultaneous action of natural adaptive and neutral processes shaping genetic variation and gene flow among populations. The presence of both drivers is supported by our genetic structure analysis, which showed different levels of ancestral admixtures in three ecologically divergent and distant areas.

2.4.3 Genetic Offset and Conservation Strategy

Our results predict a high risk of maladaptation in overall range of *A. araucana*, particularly in southern Andean piedmonts, which means that the amount of change in adaptive allele frequencies for preserve observed genetic–environmental association will be high. Therefore, we propose that these populations be prioritized for conservation. In contrast, high elevations zones in Andes Mountain range will be the least affected by climate change. This could be seen as the continuation of demographic processes described in point 4.1: while the environment warms up, the cooler highlands become refuges for this species. However, in this anthropogenic era this process is happening at a much faster rate.

Unfortunately, longevity of this species slows down the rate of emergence and spread of new adaptive alleles in populations through *de novo* mutations, making this species particularly vulnerable to rapid climate change (Fajardo *et al.*, 2019). To cope this, we propose an assisted gene flow conservation strategy to help *A. araucana* adapt to novel environments. Assisted gene flow is a controlled migration of individuals or gametes among populations within the species range, which can be effective in accelerating adaptation to future climate conditions (Aitken and Whitlock, 2013). We suggest adopting a predictive seed sourcing strategy, mixing local with pre-adapted non-local seed lots to increase diversity and resilience. Candidate donor populations could be prospected by using an adaptive divergence model like the one proposed in this study, finding areas where current environmental conditions are similar to the future conditions in the recipient population. In addition, common garden experiments must be simultaneously conducted to test donor genotypes before to implement operational conservation programs, in order to evaluate key traits and phenotypic plasticity.

Assisted gene flow have unintended risks. One of these is the outbreeding depression, which consists of a reduction in the reproductive fitness in the first or subsequent generations following attempted crossings of genetically distant populations (Frankham *et al.*, 2011). This risk can, however, be empirically estimated and could therefore be manageable.

2.4.4 Methodological Challenges

There is a consensus in the fact that loci under selection and maladaptation predictions by landscape genomic approach must be functionally validated in common garden and reciprocal transplant experiments (Li *et al.*, 2017). According to this, we adopted conservative controls for putatively adaptive

loci identification in all stages of this analysis. First, we reduced the initial dataset to a set containing only one SNP per fragment, minimizing the risk of linkage among markers. Then, we applied a consensus approach for two GEA analyses (LFMM and RDA), which required the loci to be identified by both genome scan methods, minimizing type I errors. This led us to exclude more than 1800 loci identified by the RDA. As a result, we retained 4.7% of the full set of SNPs as potentially adaptive; less than the 8.9 % reported by Strasburg et al. (Strasburg *et al.*, 2012). Therefore, our findings were based on a high confidence adaptive genotypic matrix. In the second stage, to unravel the structure of local adaptation, we adopted two approaches. First, a nonlinear distance-based model was fitted (GDM), and second, a distance-based redundancy analysis (db-RDA) was performed on both full and adaptive sets of SNPs.

The variation explained by the db-RDA was lower than that explained by the GDM, possibly due to the linear nature of db-RDA, which weakens its power to reproduce non-linear patterns of allelic turnover along environmental gradients. Despite this, both approaches were equally able to detect the most important predictor (Temperature annual range). We believe that the development of novel statistical tools that combine non-linear approaches with ordination techniques could aid in disentangling the evolutionary forces driving adaptive divergence. Finally, a major challenge is to build a reference genome for *A. Araucana* to map adaptive loci for functional trait studies.

LITERATURE CITED

- Ahrens, C.W., Rymer, P.D., Stow, A., Bragg, J., Dillon, S., Umbers, K.D., Dudaniec, R.Y., 2018. The search for loci under selection: trends, biases and progress. *Molecular Ecology* 27, 1342-1356.
- Aitken, S.N., Bemmels, J.B., 2016. Time to get moving: assisted gene flow of forest trees. *Evolutionary Applications* 9, 271-290.
- Aitken, S.N., Whitlock, M.C., 2013. Assisted gene flow to facilitate local adaptation to climate change. *Annual Review of Ecology, Evolution, and Systematics* 44, 367-388.
- Aitken, S.N., Yeaman, S., Holliday, J.A., Wang, T., Curtis-McLane, S., 2008. Adaptation, migration or extirpation: climate change outcomes for tree populations. *Evolutionary Applications* 1, 95-111.

- Alarcón, D., Cavieres, L.A., 2015. In the right place at the right time: habitat representation in protected areas of South American *Nothofagus*-dominated plants after a dispersal constrained climate change scenario. *PLoS One* 10, e0119952.
- Allen, C.D., Macalady, A.K., Chenchouni, H., Bachelet, D., McDowell, N., Vennetier, M., Kitzberger, T., Rigling, A., Breshears, D.D., Hogg, E.T., 2010. A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. *Forest Ecology and Management* 259, 660-684.
- Bay, R.A., Harrigan, R.J., Underwood, V.L., Gibbs, H.L., Smith, T.B., Rugg, K., 2018. Genomic signals of selection predict climate-driven population declines in a migratory bird. *Science* 359, 83-86.
- Bi, D., Dix, M., Marsland, S.J., O'Farrell, S., Rashid, H., Uotila, P., Hirst, A.C., Kowalczyk, E., Golebiewski, M., Sullivan, A., 2013. The ACCESS coupled model: description, control climate and evaluation. *Australian Meteorological and Oceanographic Journal (AMOJ)* 63, 41-64.
- Brodribb, T.J., McAdam, S.A., Jordan, G.J., Martins, S.C., 2014. Conifer species adapt to low-rainfall climates by following one of two divergent pathways. *Proceedings of the National Academy of Sciences* 111, 14489-14493.
- Capblancq, T., Forester, B.R., 2021. Redundancy Analysis (RDA): a Swiss Army knife for landscape genomics. *Methods in Ecology and Evolution*, Doi:10.1111/2041-210X.13722.
- Capblancq, T., Luu, K., Blum, M.G., Bazin, E., 2018. How to make use of ordination methods to identify local adaptation: a comparison of genome scans based on PCA and RDA. *bioRxiv*, 258988.
- Caye, K., Jumentier, B., Lepeule, J., François, O., 2019. LFMM 2: Fast and Accurate Inference of Gene-Environment Associations in Genome-Wide Studies. *Molecular Biology and Evolution* 36, 852-860.
- Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., 2011. The variant call format and VCFtools. *Bioinformatics* 27, 2156-2158.
- de Sousa, L.M., Poggio, L., Batjes, N.H., Heuvelink, G., Kempen, B., Riberio, E., Rossiter, D., 2020. SoilGrids 2.0: producing quality-assessed soil information for the globe. *SOIL Discussions*, 1-37.

- Derry, A.M., Fraser, D.J., Brady, S.P., Astorg, L., Lawrence, E.R., Martin, G.K., Matte, J.M., Negrín Dastis, J.O., Paccard, A., Barrett, R.D., 2019. Conservation through the lens of (mal) adaptation: Concepts and meta-analysis. *Evolutionary Applications* 12, 1287-1304.
- Doyle, J., 1991. DNA Protocols for Plants. In: Hewitt, G.M., Johnston, A.W.B., W., Y.J.P. (Eds.), *Molecular Techniques in Taxonomy*. Springer-Verlag, Berlin, Germany, pp. 283-293.
- Eaton, D.A., Overcast, I., 2020. ipyrad: Interactive assembly and analysis of RADseq datasets. *Bioinformatics* 36, 2592-2594.
- Engelbrecht, B.M., Comita, L.S., Condit, R., Kursar, T.A., Tyree, M.T., Turner, B.L., Hubbell, S.P., 2007. Drought sensitivity shapes species distribution patterns in tropical forests. *Nature* 447, 80-82.
- Fajardo, A., McIntire, E.J., Olson, M.E., 2019. When short stature is an asset in trees. *Trends in Ecology & Evolution* 34, 193-199.
- Ferrier, S., Manion, G., Elith, J., Richardson, K., 2007. Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. *Diversity and Distributions* 13, 252-264.
- Fitzpatrick, M.C., Keller, S.R., 2015. Ecological genomics meets community-level modelling of biodiversity: Mapping the genomic landscape of current and future environmental adaptation. *Ecology Letters* 18, 1-16.
- Fitzpatrick, M.C., Mokany, K., Manion, G., Lisk, M., Ferrier, S., Nieto-Lugilde, D., Fitzpatrick, M.M.C., Rcpp, L., Rcpp, I., 2021. Package 'gdm'. A toolkit with functions to fit, plot, and summarize Generalized Dissimilarity Models: CRAN Repository, R. Version 1.4.2.2.
- Fitzpatrick, M.C., Sanders, N.J., Ferrier, S., Longino, J.T., Weiser, M.D., Dunn, R., 2011. Forecasting the future of biodiversity: a test of single-and multi-species models for ants in North America. *Ecography* 34, 836-847.
- Forester, B.R., Lasky, J.R., Wagner, H.H., Urban, D.L., 2018. Comparing methods for detecting multilocus adaptation with multivariate genotype–environment associations. *Molecular Ecology* 27, 2215-2233.
- Frankham, R., Ballou, J.D., Eldridge, M.D., Lacy, R.C., Ralls, K., Dudash, M.R., Fenster, C.B., 2011. Predicting the probability of outbreeding depression. *Conservation Biology* 25, 465-475.
- Frichot, E., François, O., 2015. LEA: An R package for landscape and ecological association studies. *Methods in Ecology and Evolution* 6, 925-929.

- Frichot, E., Mathieu, F., Trouillon, T., Bouchard, G., François, O., 2014. Fast and efficient estimation of individual ancestry coefficients. *Genetics* 196, 973-983.
- Frichot, E., Schoville, S.D., Bouchard, G., François, O., 2013. Testing for associations between *loci* and environmental gradients using latent factor mixed models. *Molecular Biology and Evolution* 30, 1687-1699.
- Gougherty, A.V., Keller, S.R., Fitzpatrick, M.C., 2021. Maladaptation, migration and extirpation fuel climate change risk in a forest tree species. *Nature Climate Change* 11, 166-171.
- Gugger, P.F., Liang, C.T., Sork, V.L., Hodgskiss, P., Wright, J.W., 2018. Applying landscape genomic tools to forest management and restoration of Hawaiian koa (*Acacia koa*) in a changing environment. *Evolutionary Applications* 11, 231-242.
- Hasbún, R., González, J., Iturra, C., Fuentes, G., Alarcón, D., Ruiz, E., 2016. Using genome-wide SNP discovery and genotyping to reveal the main source of population differentiation in *Nothofagus dombeyi* (Mirb.) Oerst. in Chile. *International Journal of Genomics* 2016.
- Hecht, B.C., Matala, A.P., Hess, J.E., Narum, S.R., 2015. Environmental adaptation in Chinook salmon (*Oncorhynchus tshawytscha*) throughout their North American range. *Molecular Ecology* 24, 5573-5595.
- Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology: A Journal of the Royal Meteorological Society* 25, 1965-1978.
- Hoban, S., Kelley, J.L., Lotterhos, K.E., Antolin, M.F., Bradburd, G., Lowry, D.B., Poss, M.L., Reed, L.K., Storfer, A., Whitlock, M.C., 2016. Finding the genomic basis of local adaptation: pitfalls, practical solutions, and future directions. *The American Naturalist* 188, 379-397.
- Hoffmann, A.A., Sgrò, C.M., 2011. Climate change and evolutionary adaptation. *Nature* 470, 479-485.
- Holderegger, R., Wagner, H.H., 2008. Landscape genetics. *Bioscience* 58, 199-207.
- Holliday, J.A., Aitken, S.N., Cooke, J.E., Fady, B., González-Martínez, S.C., Heuertz, M., Jaramillo-Correa, J.P., Lexer, C., Staton, M., Whetten, R.W., 2017. Advances in ecological genomics in forest trees and applications to genetic resources conservation and breeding. In. *Wiley Online Library*.
- Jia, K.H., Zhao, W., Maier, P.A., Hu, X.G., Jin, Y., Zhou, S.S., Jiao, S.Q., El-Kassaby, Y.A., Wang, T., Wang, X.R., 2020. Landscape genomics predicts climate change-related genetic offset for the widespread *Platyclusus orientalis* (Cupressaceae). *Evolutionary Applications* 13, 665-676.

- Jones, C., Hughes, J., Bellouin, N., Hardiman, S., Jones, G., Knight, J., Liddicoat, S., O'connor, F., Andres, R.J., Bell, C., 2011. The HadGEM2-ES implementation of CMIP5 centennial simulations. *Geoscientific Model Development* 4, 543-570.
- Jump, A.S., Penuelas, J., 2005. Running to stand still: adaptation and the response of plants to rapid climate change. *Ecology Letters* 8, 1010-1020.
- Kershaw, P., Wagstaff, B., 2001. The southern conifer family Araucariaceae: history, status, and value for paleoenvironmental reconstruction. *Annual Review of Ecology and Systematics* 32, 397-414.
- Legendre, P., Anderson, M.J., 1999. Distance-based redundancy analysis: testing multispecies responses in multifactorial ecological experiments. *Ecological Monographs* 69, 1-24.
- Legendre, P., Legendre, L., 2012. *Numerical ecology*. Elsevier.
- Li, Y., Zhang, X.-X., Mao, R.-L., Yang, J., Miao, C.-Y., Li, Z., Qiu, Y.-X., 2017. Ten years of landscape genomics: challenges and opportunities. *Frontiers in Plant Science* 8, 2136.
- Luu, K., Bazin, E., Blum, M.G., 2017. pcadapt: an R package to perform genome scans for selection based on principal component analysis. *Molecular Ecology Resources* 17, 67-77.
- Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., Hornik, K., 2021. *cluster: Cluster Analysis Basics and Extensions*. <https://CRAN.R-project.org/package=cluster>.
- Manel, S., Schwartz, M.K., Luikart, G., Taberlet, P., 2003. Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology & Evolution* 18, 189-197.
- Marchelli, P., Baier, C., Mengel, C., Ziegenhagen, B., Gallo, L., 2010. Biogeographic history of the threatened species *Araucaria araucana* (Molina) K. Koch and implications for conservation: a case study with organelle DNA markers. *Conservation Genetics* 11, 951-963.
- Martín, M.A., Mattioni, C., Lusini, I., Molina, J.R., Cherubini, M., Drake, F., Herrera, M.A., Villani, F., Martín, L.M., 2014. New insights into the genetic structure of *Araucaria araucana* forests based on molecular and historic evidences. *Tree Genetics & Genomes* 10, 839-851.
- Mattera, M.G., Pastorino, M.J., Lantschner, M.V., Marchelli, P., Soliani, C., 2020. Genetic diversity and population structure in *Nothofagus pumilio*, a foundation species of Patagonian forests: defining priority conservation areas and management. *Scientific Reports* 10, 1-13.
- McIntosh, M., González-Campos, J., Demaree, P., Toro, O., Ipinza, R., Bustamante-Sánchez, M.A., Hasbún, R., Nelson, C.R., Unpublished results. Trait variation between and within Andes and coastal mountain ranges in the iconic South American tree *Araucaria araucana* in Chile.

- Nanavati, W., Whitlock, C., Outes, V., Villarosa, G., 2020. A Holocene history of monkey puzzle tree (pehuén) in northernmost Patagonia. *Journal of Biogeography*.
- Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P.R., O'hara, R., Simpson, G.L., Solymos, P., Stevens, M.H.H., Wagner, H., 2020. Package 'vegan'. Community ecology package. R Foundation for Statistical Computing, Vienna, Austria. Version 2.5-7, 1-298.
- Papú, S., Berli, F., Piccoli, P., Patón, D., Rodriguez, D.O., Roig, F., 2021. Physiological, biochemical, and anatomical responses of *Araucaria araucana* seedlings to controlled water restriction. *Plant Physiology and Biochemistry* 165, 47-56.
- Premoli, A., Quiroga, P., Gardner, M., 2013. Premoli, A., Quiroga, P. & Gardner, M. 2013. *Araucaria araucana*. The IUCN Red List of Threatened Species 2013: e.T31355A2805113. <https://dx.doi.org/10.2305/IUCN.UK.2013-1.RLTS.T31355A2805113.en>. Downloaded on 11 August 2021. In.
- Premoli, A.C., Kitzberger, T., Veblen, T.T., 2000. Conservation genetics of the endangered conifer *Fitzroya cupressoides* in Chile and Argentina. *Conservation Genetics* 1, 57-66.
- Puchi, P.F., Camarero, J.J., Battipaglia, G., Carrer, M., 2021. Retrospective analysis of wood anatomical traits and tree-ring isotopes suggests site-specific mechanisms triggering *Araucaria araucana* drought-induced dieback. *Global Change Biology*.
- Rafii, Z.A., Dodd, R.S., 1998. Genetic diversity among coastal and Andean natural populations of *Araucaria araucana* (Molina) K. Koch. *Biochemical Systematics and Ecology* 26, 441-451.
- Rellstab, C., Gugerli, F., Eckert, A.J., Hancock, A.M., Holderegger, R., 2015. A practical guide to environmental association analysis in landscape genomics. *Molecular Ecology* 24, 4348-4370.
- Sanderson, B.M., Knutti, R., Caldwell, P., 2015. A representative democracy to reduce interdependency in a multimodel ensemble. *Journal of Climate* 28, 5171-5194.
- Schoville, S.D., Bonin, A., François, O., Lobreaux, S., Melodelima, C., Manel, S., 2012. Adaptive genetic variation on the landscape: methods and cases. *Annual Review of Ecology, Evolution, and Systematics* 43, 23-43.
- Sersic, A.N., Cosacov, A., Cocucci, A.A., Johnson, L.A., Pozner, R., Avila, L.J., Sites Jr, J.W., Morando, M., 2011. Emerging phylogeographical patterns of plants and terrestrial vertebrates from Patagonia. *Biological Journal of the Linnean Society* 103, 475-494.

- Shryock, D.F., Washburn, L.K., DeFalco, L.A., Esque, T.C., 2020. Harnessing landscape genomics to identify future-climate resilient genotypes in a desert annual. *Molecular Ecology*.
- Souto, C.P., Mathiasen, P., Acosta, M.C., Quiroga, M.P., Vidal-Russell, R., Echeverría, C., Premoli, A.C., 2015. Identifying genetic hotspots by mapping molecular diversity of widespread trees: when commonness matters. *Journal of Heredity* 106, 537-545.
- Strasburg, J.L., Sherman, N.A., Wright, K.M., Moyle, L.C., Willis, J.H., Rieseberg, L.H., 2012. What can patterns of differentiation across plant genomes tell us about adaptation and speciation? *Philosophical Transactions of the Royal Society B: Biological Sciences* 367, 364-373.
- Sultan, S.E., 2000. Phenotypic plasticity for plant development, function and life history. *Trends in Plant Science* 5, 537-542.
- Tella, J.L., Lambertucci, S.A., Speziale, K.L., Hiraldo, F., 2016. Large-scale impacts of multiple co-occurring invaders on monkey puzzle forest regeneration, native seed predators and their ecological interactions. *Global Ecology and Conservation* 6, 1-15.
- Urban, M.C., 2015. Accelerating extinction risk from climate change. *Science* 348, 571-573.
- Velez, M., Salgado Salomón, M., Marfetan, A., Tirante, S., Mattes Fernández, H., Avila, M., Szychowski, A., Viviani, D., Sanguinetti, J., 2018. Caracterización de la desecación del dosel y sanidad de *Araucaria araucana* en Argentina.
- Villagrán, C., 2018. Biogeografía de los bosques subtropical-templados del sur de Sudamérica. Hipótesis históricas. *Magallania (Punta Arenas)* 46, 27-48.
- Wiens, J.J., 2016. Climate-Related Local Extinctions Are Already Widespread among Plant and Animal Species. *PLoS Biology* 14.
- Zimmer, H.C., Brodribb, T.J., Delzon, S., Baker, P.J., 2016. Drought avoidance and vulnerability in the Australian Araucariaceae. *Tree Physiology* 36, 218-228.

SUPPORTING INFORMATION

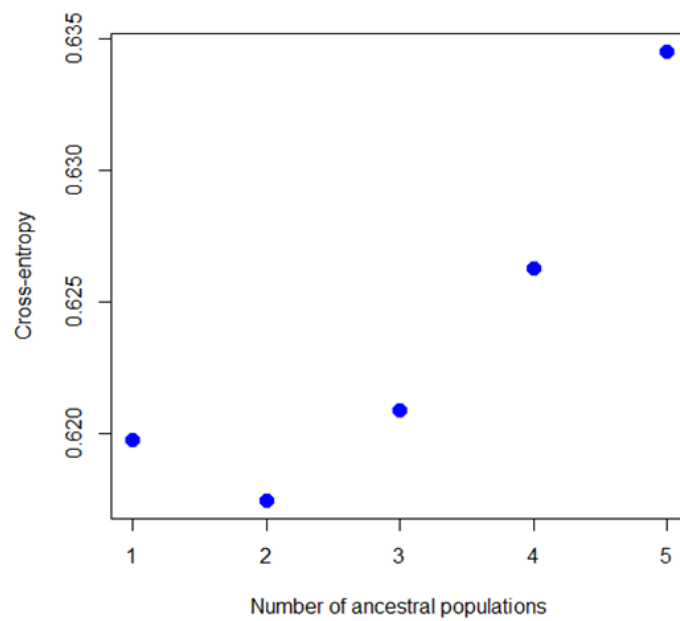
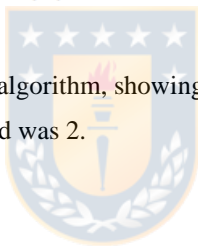


Figure S.1. Cross Entropy values returned by sNMF algorithm, showing a minimal value for $K=2$. Minimal cross entropy indicates that number of ancestral populations adopted was 2.



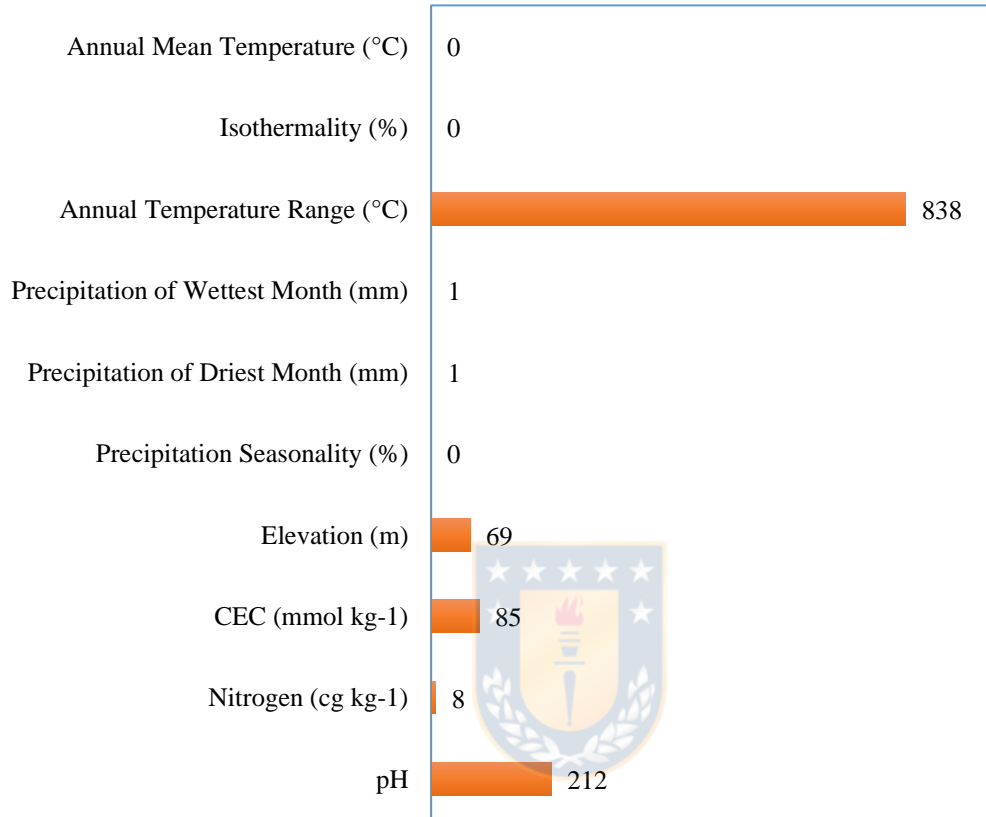


Figure S.2. Number of loci identified as potentially adaptive per predictor in LFMM, showing the temperature annual range as the variable that yielded the most adaptive loci, followed by pH. Note that final adaptive loci matrix was built after RDA analysis, containing only loci detected as putatively Adaptive by both methods.

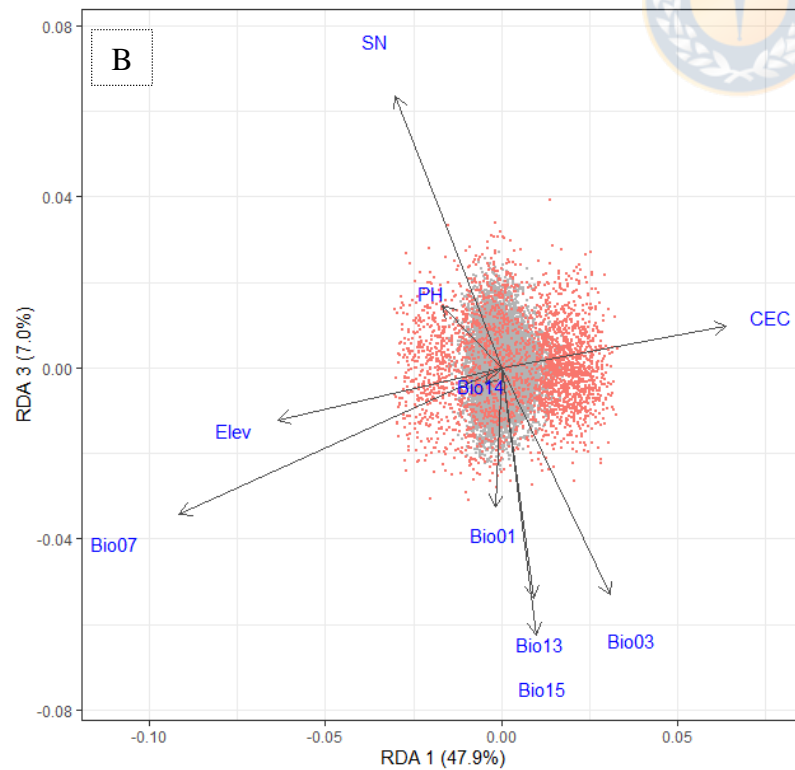
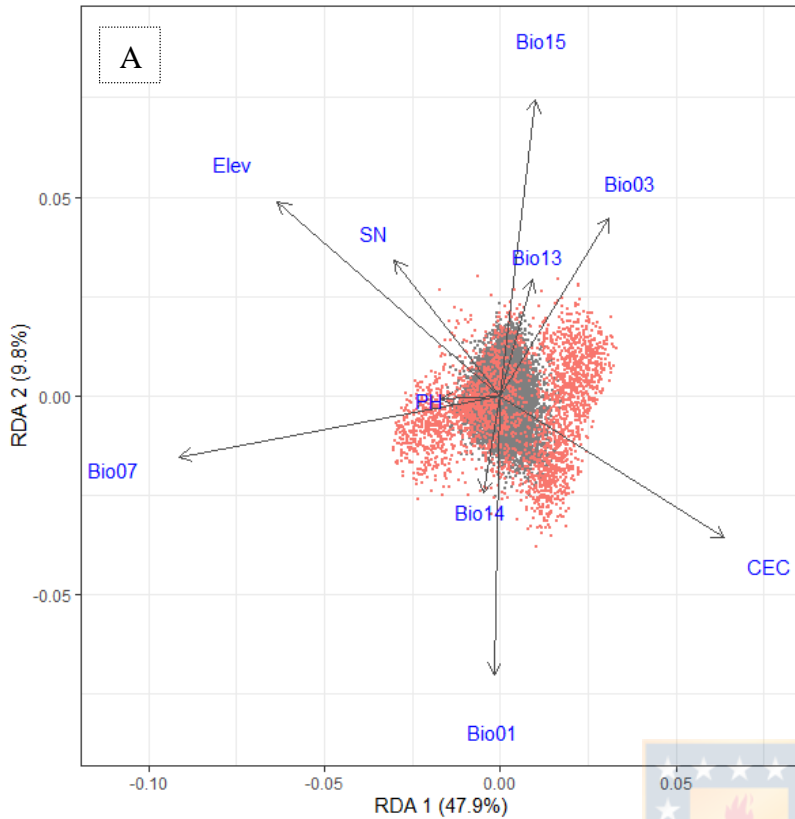


Figure S.3. SNPs (dots) and environmental predictors (arrows) in RDA space performed for adaptive loci detection. Panel A) axes 1 and 2. Panel B) axes 1 and 3. In parentheses the variance explained by each canonical axis. Red dots are adaptive SNPs; grey dots are neutral SNPs. Legend: Bio01: Annual Mean Temperature; Bio03: Isothermality; Bio07: Annual Temperature Range; Bio13: Precipitation of Wettest Month; Bio14: Precipitation of Driest Month; Bio15: Precipitation Seasonality; Elev: Elevation; CEC: Cationic Exchange Capacity; SN: Total soil Nitrogen; PH: Soil pH.

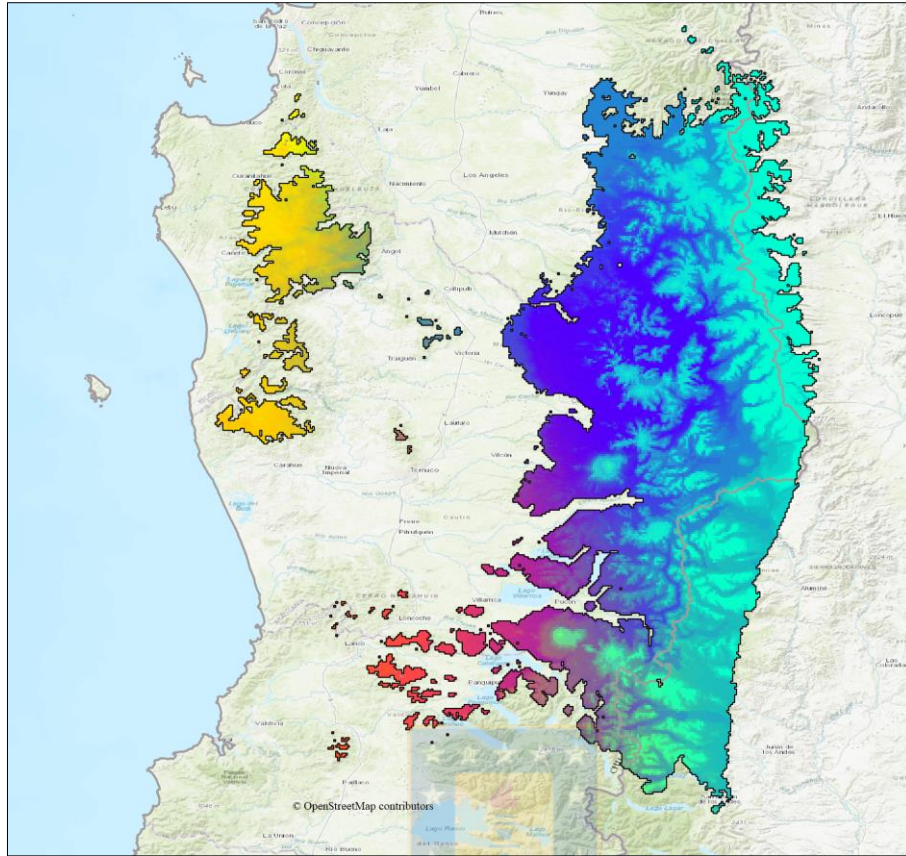


Figure S4. Map of spatial interpolations of adaptive divergence for *Araucaria araucana*. RGB color combinations results from principal components analysis of GDM-transformed environmental variables. Similar colors represent similar patterns of local adaptation.

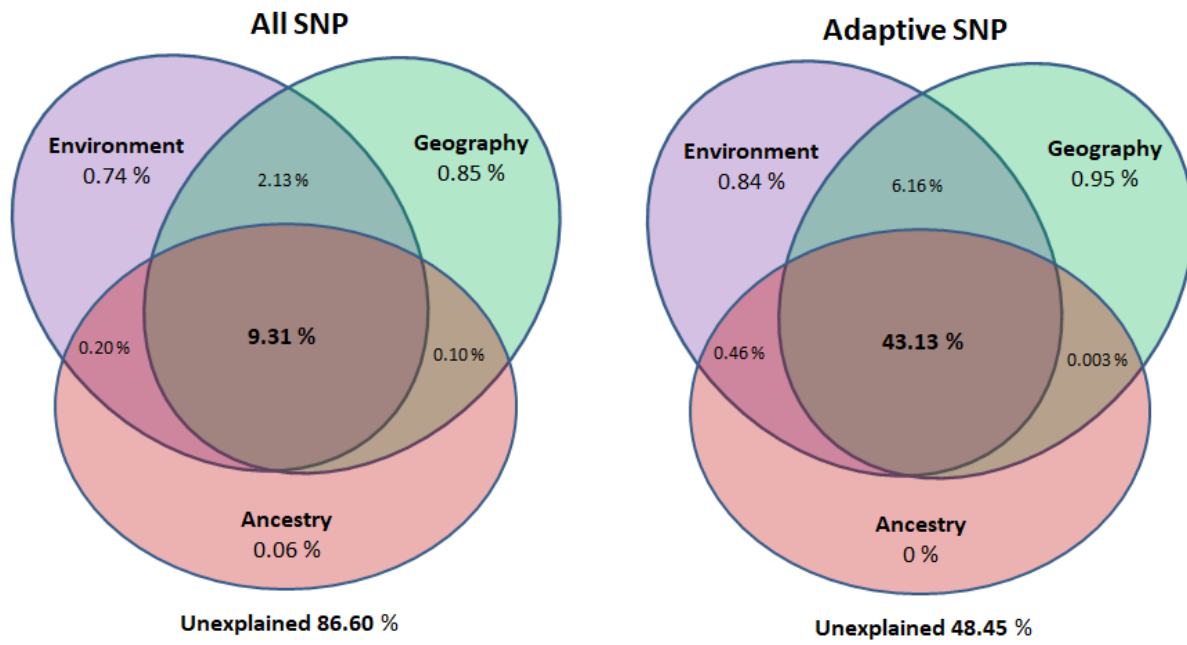


Figure S5. Venn diagrams showing the proportion of among-individuals genetic variation explained by environment, geography, and ancestry in (db-RDA) using all SNPs (left) and adaptive SNPs (right) in *A. araucana*.



CAPÍTULO III

Searching for Pre-Adapted Genotypes Using an Adaptive Divergence Model for the iconic South American conifer *Araucaria araucana*

Running head: Climate adjusted seed sourcing in a South American tree species

Antonio Varas-Myrik¹, Francisco Sepúlveda-Espinoza¹, Alex Fajardo², Diego Alarcón³, Óscar Toro-Núñez⁴ and Rodrigo Hasbún¹.

¹Laboratorio de Epigenética Vegetal, Departamento de Silvicultura, Facultad de Ciencias Forestales, Universidad de Concepción, Victoria 631, Concepción, Chile.

²Instituto de Investigación Interdisciplinario (I³), Universidad de Talca, Campus Lircay, Talca 3460000, Chile.

³Departamento de Ciencias Ecológicas, Universidad de Chile. Instituto de Ecología y Biodiversidad. Las Palmeras 3425, Ñuñoa, Santiago, Chile.

⁴Departamento de Botánica, Facultad de Ciencias Naturales y Oceanográficas, Universidad de Concepcion, Victoria 631, Concepción, Chile.



Type of contribution: Article

*Corresponding author: Rodrigo Hasbún; email: rhasbun@udec.cl

ABSTRACT

Climate-matching seed sourcing strategies have raised as a powerful tool to cope maladaptation expected in forests due to human-induced climate change. These strategies consider assisted migration of pre-adapted genotypes from populations living under current climatic conditions that recipient populations will experiment in the future, thus reducing climate change-related genetic offset and increasing mean fitness. Here, we used a previously built genomic adaptive divergence model to answer whether populations containing pre-adapted genotypes or populations are possible to find in natural populations of the iconic South American conifer *Araucaria araucana*. We concluded that climate-matching seed sourcing strategies will be difficult to implement in this species because of a lack of populations containing pre-adapted genotypes; i.e., *in situ* genetic offset will be lower than expected maladaptation of any immigrant genotypes. This study is the first one of this kind in a South American tree species and its novel methodology might constitute an important contribution to the conservation of *A. araucana*.

Key words: local adaptation, climate change, landscape genomics, conservation genetics, RAD-Seq.

3.1 INTRODUCTION

Climate change is recognized as one of the main threats that ecosystems are fronting and will have to face with greater emphasis in the near future. Many woody species are reaching their physiological limits (Adams *et al.*, 2017), resulting in massive forest dieback (Evans *et al.*, 2017). Such mortality is mainly explained by two physiological phenomena: carbon starvation and hydraulic failure (Menezes-Silva *et al.*, 2019). These organisms have historically handled environmental changes through migration, *in-situ* adaptation, or phenotypic plasticity. Nonetheless, the speed of anthropogenic environmental changes outpaces the adaptive response of many species (Radchuk *et al.*, 2019). This entails an increasing risk of maladaptation, mortality, and finally the extinction of local populations, bringing catastrophic ecological, social and even economic consequences for humanity. These pressures demand land managers to reassess forest management practices (Dumroese *et al.*, 2015). Some alternatives to face climate change are ecological restoration, assisted migration (AM), and bioengineering; all of them with the potential to be applied independently or concurrently at different scales (Dumroese *et al.*, 2015).

Assisted migration is a conservation strategy that arises as an anticipated response to climate change that is expected to occur in the future, and involves anthropogenic movement of pre-adapted individuals within or outside the range of the species (Havens and Vitt, 2011; Aitken and Whitlock, 2013). For the purposes of this paper, we will use assisted migration concept to refer the incorporation of adaptive genetic variation within the range of a species. This approach can be useful for management or conservation option of species and/or populations, which currently are or will be under the negative effects of climate change, especially for those isolated or less adaptable (e.g. low frequency of climate-adapted alleles) (Chen *et al.*, 2021).

A second concept that is key for understand what we call “pre-adapted” is named genetic offset. This concept was coined by Fitzpatrick and Keller (2015), and is defined as the measure of the amount of adaptive genetic change between present and future climate conditions required for to maintain species current genetic–environmental association. This idea is based on the assumption that in the past, a sufficiently long period of relative environmental stability plus natural selection has produced a association between adaptive divergence and the environmental gradients. Such relationship can be

detected in the genome of individuals and spatially modelled (Fitzpatrick and Keller, 2015; Gugger *et al.*, 2018; Jia *et al.*, 2020; Varas-Myrik *et al.*, 2022).

Integrating both concepts, genetic offset brings a measure of the distance between an adaptive allelic configuration and an environmental scenario, bringing an objective criterion for how “pre-adapted” an individual or population is to given future environmental scenario. Thus, we can assume that *in-situ* genetic offset is maladaptation that will suffer a population growing in its site as a result of climate change, and *ex-situ* genetic offset as maladaptation of an individual (or propagules) in a different and distant place that is growing. Hence, we are capable to define as positive or benefit genetic interchange, or pre-adapted individual or population, as one whose *ex-situ* genetic offset (donor genetic material) is less than the *in-situ* genetic offset (recipient population). In this way, individuals or propagules incorporated in a recipient population will contribute to decrease expected maladaptation and increasing evolutionary fitness.

For to detect adaptive divergence in genome of individuals and to build spatially explicit adaptive divergence models, a new approach called Landscape Genomics (LG) has arisen as the most powerful framework. This discipline, which combines elements of population genetics, landscape ecology and spatial statistics, assumes that landscape elements such as geography, climate and soil have a causal influence on adaptive divergence (Li *et al.*, 2017). Through the use of next generation DNA sequencing techniques (NGS), which provides a broad sampling of the genome and large number of single nucleotide polymorphisms (SNP), LG provide mathematical tools for to detect loci subject to natural selection and then to model adaptive divergence using as predictors environmental variables, informing scientifically genetic offset and assisted migration (Fitzpatrick and Keller, 2015). Some studies have already used this framework to genetic offset predictions and have proposed spatially explicit tools to search for populations pre-adapted to future conditions (Supple *et al.*, 2018; Rossetto *et al.*, 2019; Shryock *et al.*, 2020). However, these studies had point-based outcomes and did not evaluate the possibility of AM in the entire range of the species. In this work, we present a methodology and study case to determine if it is feasible to implement an assisted migration strategy, evaluating the existence of populations containing pre-adapted genotypes in the complete range of the study species.

One of the iconic conifers of temperate forests, regarded with an important social and ancestral value, narrow geographic distribution in latitude and wide in longitude, is *Araucaria araucana*. Currently, local populations are considered under a critical threat of extinction, given

the observed levels of degradation of their habitats due to anthropogenic pressure (mostly fires, illegal logging and overgrazing; Premoli *et al.* (2013)). These elements suggest that AM is an appropriate conservation tool to i) to cope maladaptation that will arise due to climate change, and ii) to maintain high levels of adaptive genetic diversity that prompts an adaptive response for climate change in the near future in. This task could be accomplished by using newly genomic data and genetic offset analyses produced in *A. araucana*, which has revealed high strength of local adaptation in this species and predicted high levels of genetic offset because of climate change (Varas-Myrik *et al.*, 2022).

Here, we propose a new approach and a study case to assess the viability of implement an AM program, mainly determined by the existence of pre-adapted genotypes or populations which contribute to reducing the *in-situ* genetic offset of recipient and most endangered populations. Our work constitutes an important contribution to the conservation of the species, informing conservation programs such as assisted gene flow, and seed supply strategies for ecological restoration.

3.2 MATERIALS AND METHODS

To find candidate seed sources potentially containing pre-adapted genotypes from another distant population, we followed a previously built adaptive divergence model for this purpose (Varas-Myrik *et al.*, 2022). The methodology to fit the model are clearly detailed in this work, which we will briefly describe as follows:

One hundred fifty *A. araucana* trees were sampled covering their entire distribution. Nuclear DNA from needle tissue was extracted and a genomic library was constructed using the RAD-seq. After bioinformatics processing (*de novo* assembling and datasets filtering), a matrix of more than 49 thousand SNPs and 135 individuals was assembled, which was used to conduct a preliminary population structure analysis. This genotypic matrix was once again filtered to retain only one SNP per fragment, in order to avoid issues of Linkage Disequilibrium, resulting in a matrix of 18 thousand SNPs. With this dataset, environmental association analyses were performed, detecting near 900 loci putatively under selection.

To provide the explanatory dataset for adaptive divergence model, an environmental explanatory matrix was obtained from the worldclim 1.4 (Hijmans *et al.*, 2005) and soilgrid 2.0 (de Sousa *et al.*, 2020) databases, which was subjected to a procedure of elimination of correlated variables. Then, the adaptive

divergence model was fitted using a GDM (Generalized Dissimilarity Model, Ferrier *et al.* (2007)), with individual adaptive genomic distances estimated using environmental distances as predictors. Fitting results showed an 85% of variance explained by the model, being temperature annual range the most important predictor, followed by geographic distance and precipitation of wettest month. This model was used for estimate adaptive divergence using as predictor the distance between 2 environmental scenarios.

To harness the proposed adaptive divergence model and search for candidate donor populations, we estimated two values: a) *in-situ* genetic offset of the recipient population and b) *ex-situ* genetic offset of a given donor population respect to recipient population. If ($b < a$), we assumed there are a potential beneficial adaptive genetic interchange from b) to a), since the incorporation of such genotypes can reduce the climate change-related genetic offset of the recipient population, increasing mean fitness and genetic diversity. Since pairwise pixel to pixel or point to point analysis is not feasible, we homologated the concept of "population" with that of "adaptive zone", hence, we grouped the pixels of niche-based raster mask into 4 and 8 homogeneous adaptive zones by clustering the matrix of GDM transformed predictors (Varas-Myrik *et al.*, 2022). This was conducted with the "*clara*" function from the *Cluster* R package (Maechler *et al.*, 2021). Then, we obtained the mean values (present climate) of the GDM untransformed predictors for each adaptive zone, and then created a raster for each predictor with that average value, which we later packed into a raster stack. This stack was used as a present raster in the *gdm.predict* function from the GDM R package (Fitzpatrick *et al.*, 2021). As future raster, the stack of GDM untransformed climatic variables

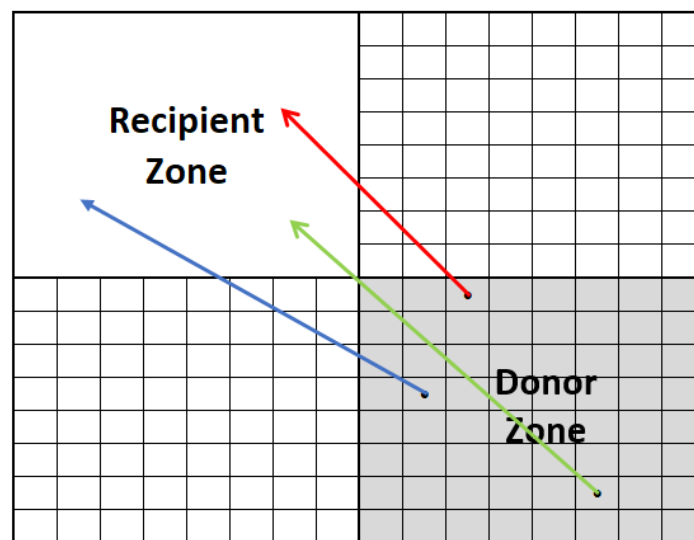


Figure 7. Illustration of *ex-situ* genetic offset calculation of a pixel regarding the mean future environmental conditions of a recipient adaptive zone. Offset is then evaluated between the present environmental data of a given pixel and the mean future conditions of an adaptive recipient zone. In this case, the recipient zone may be assumed as a single pixel, since its future environmental conditions (Worldclim 1.4, CMIP5, year 2070, RCP4.5) were averaged. On the other hand, in the candidate donor zone, each pixel is considered in the evaluation of genetic offset.

for the year 2070 and the moderate scenario RCP4.5 (worldclim 1.4) was adopted. Thus, the output raster contains the *ex-situ* genetic offset values for pixels not belonging to the recipient adaptive zone, and *in-situ* genetic offset for pixel that belong to the zone (Figure 7). This can be interpreted as the places where potentially pre-adapted genotypes can be found for that recipient adaptive zone. The analysis was carried out considering 4 and 8 adaptive zones, in order to reduce the climatic variation within each zone and increase the resolution of the analysis.



3.3 RESULTS

3.3.1 Four Zones Resolution

Regarding adaptive zones, clear differentiation can be observed between Coastal and Andes Mountain ranges in four adaptive zones map (Figure 8), showing a pattern of diagonal bands. Zone 1 represents high lands in Andes Mountain range. Zone 3 was almost exclusively from Coastal Mountain range. Zone 2 is an intermediate between zone 4 and zone 1, and zone 4 represents Andean piedmont and will suffer mostly climate change effects (Varas-Myrik *et al.*, 2022). Investigating a possible beneficial genetic interchange, we observe that average genetic offset between mean future environmental conditions of a given receiving zone and present environmental conditions of any of the remaining zones (candidate donor zone) was always higher than genetic offset of receiving zone with itself (Figure 9, Table 4). Thus, on average, given an adaptive zone, none of the three remaining could present pre-adapted genotypes to ameliorate the impact of climate change.

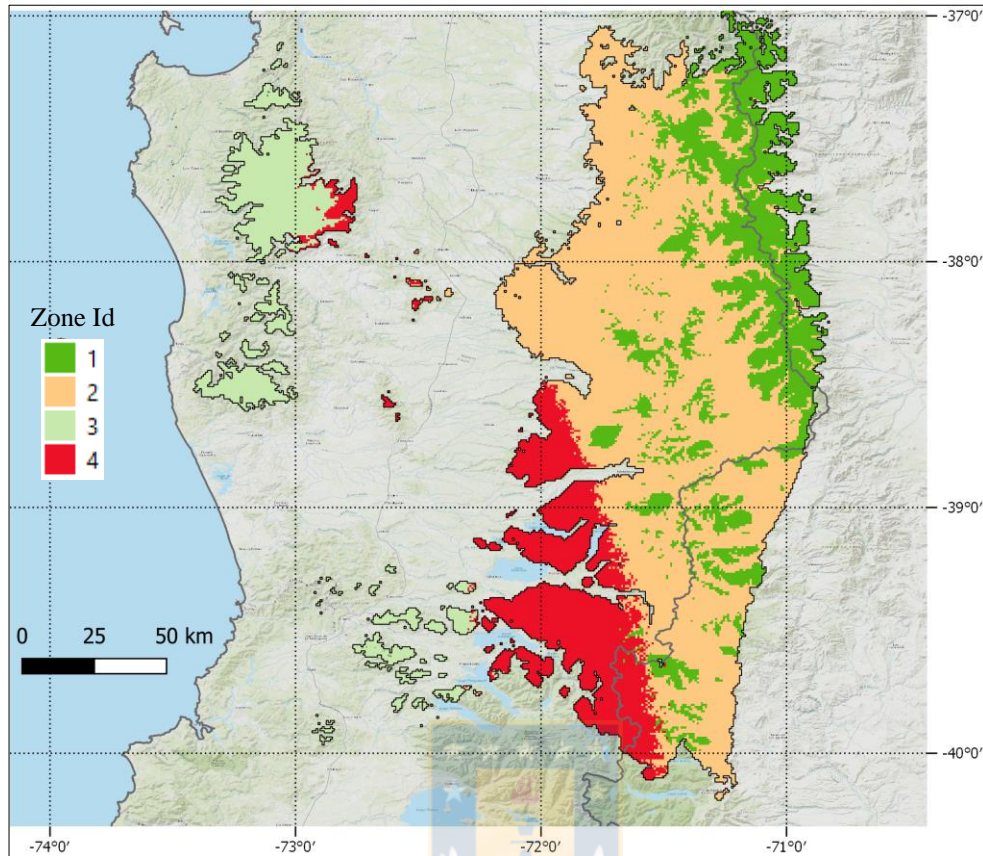
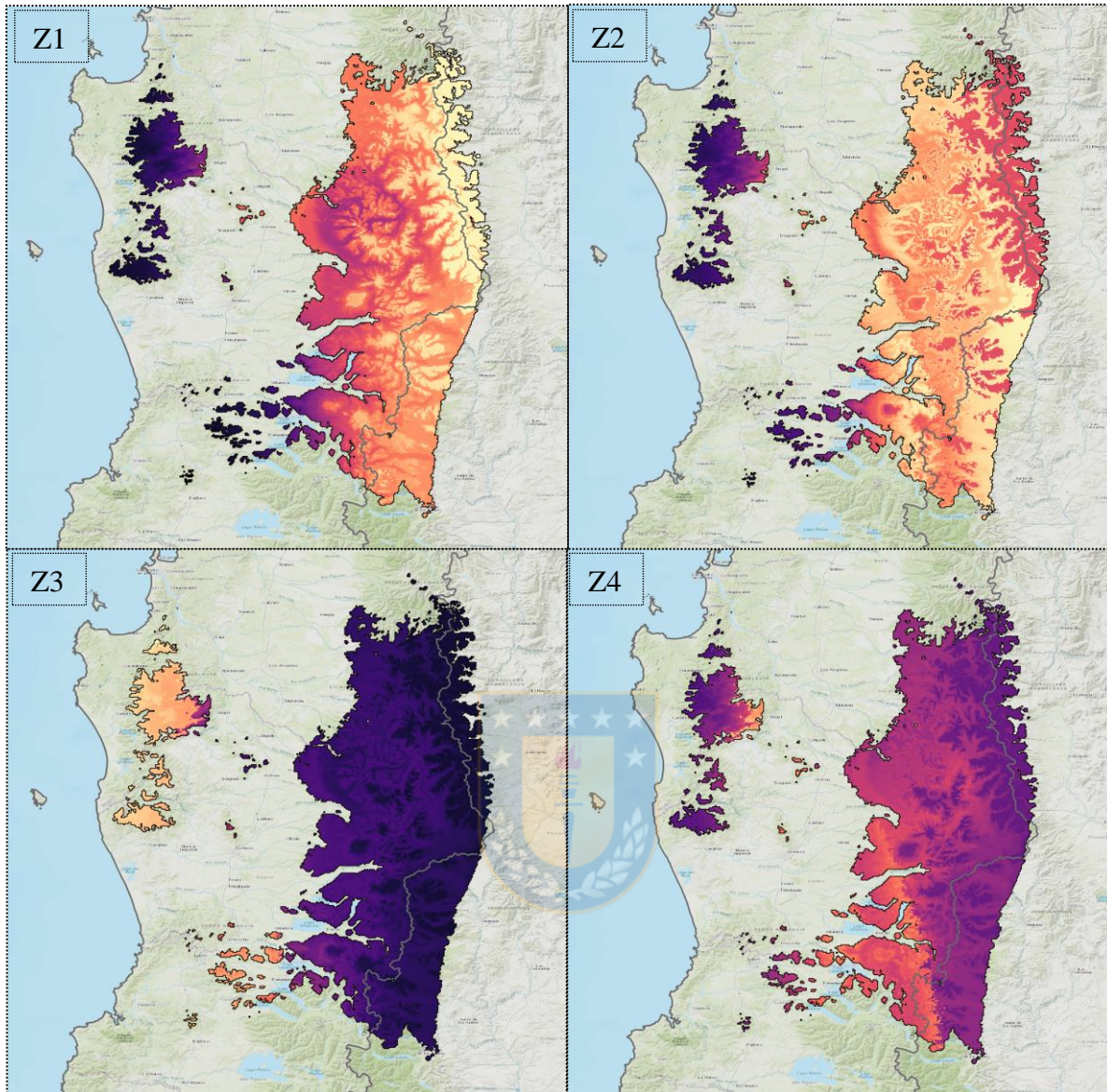


Figure 8. Map showing four zones of similar adaptive allelic compositions for *A. araucana*, built using the K-means clustering algorithm applied to GDM transformed predictor's data frame.

Table 4. Pairwise genetic offset values of four adaptive zones. Diagonal values in bold are in-situ genetic offset. If in a row there are a lower value than bold number, the zone matching to the column of this cell was assumed as a candidate donor zone.

Adaptive zone	Candidate donor zone			
	Z1	Z2	Z3	Z4
Z1	0.51	0.61	0.84	0.67
Z2	0.64	0.54	0.80	0.58
Z3	0.86	0.83	0.55	0.79
Z4	0.77	0.71	0.74	0.63



0.49  0.90

Figure 9. Genetic offset continuous map for each four adaptive zone. Each pixel represents genetic offset among current environmental conditions of that pixel and average future environmental conditions of the adaptive zone.

3.3.2 Eight Zones Resolution

To catch possible missed variation due to adopting environmental average per adaptive zone, we increased resolution and repeated the analysis using eight adaptive zones (Figure 10). The result was nearly to 4 zones resolution, except by interchange from zone 5 to zone 7 which appear be benefit (Table

5). Here, average genetic offset of zone 7 was 0.64, and mean genetic offset from candidate donor zone 5 to zone 7 was 0.63, a bit lower. However, zones 5 and 7 correspond to zone 4 in 4 zones resolution map, and we can interpret this as an intra-zone assisted gene flow.

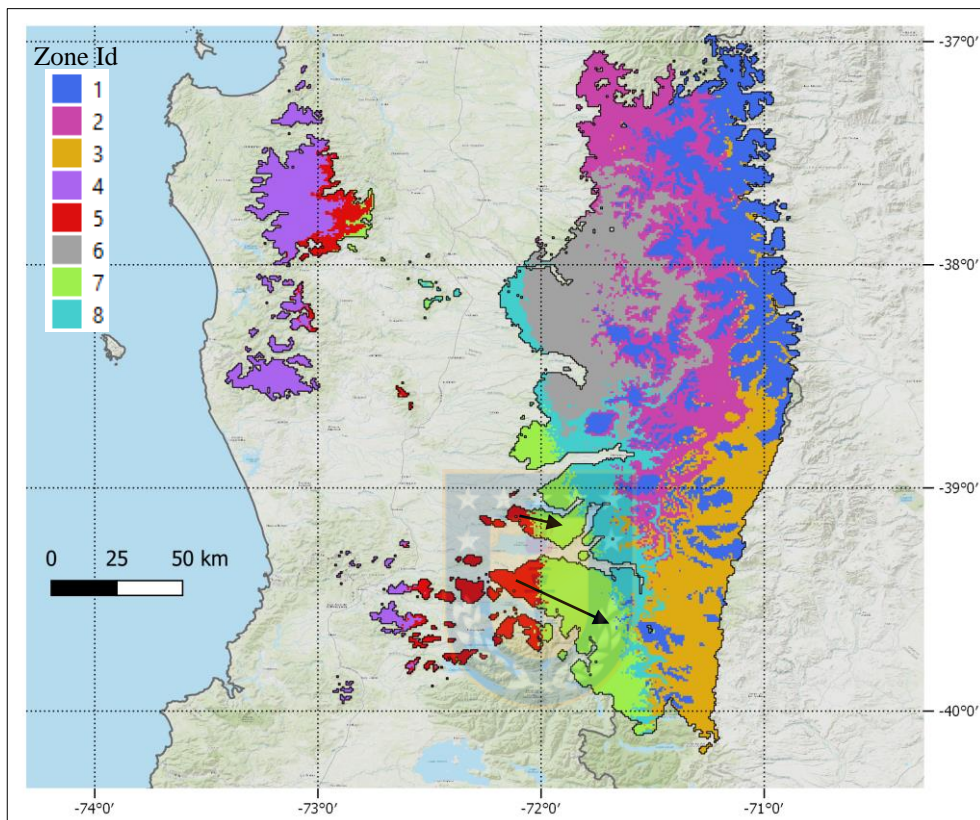


Figure 10. Map showing eight zones of similar adaptive allelic compositions for *A. araucana*, built with the *K*-means clustering algorithm applied to GDM transformed predictor's data frame. Black arrows represent AM from zone 5 to zone 7.

Table 5. Pairwise genetic offset values of eight adaptive zones. Diagonal values in bold are mean in-situ genetic offset. If in a row there are a lower value than bold number (*), the zone matching to the column of this cell was assumed as a candidate donor zone.

Adaptive Zone	Candidate donor Zone							
	Z1	Z2	Z3	Z4	Z5	Z6	Z7	Z8
Z1	0.51	0.60	0.58	0.84	0.79	0.66	0.65	0.63
Z2	0.63	0.53	0.58	0.81	0.73	0.55	0.58	0.54
Z3	0.61	0.57	0.52	0.81	0.74	0.60	0.59	0.56
Z4	0.86	0.83	0.84	0.55	0.73	0.83	0.81	0.82
Z5	0.85	0.82	0.83	0.70	0.59	0.78	0.78	0.79
Z6	0.72	0.66	0.68	0.84	0.72	0.57	0.64	0.62
Z7	0.77	0.72	0.74	0.76	0.63*	0.69	0.64	0.68
Z8	0.71	0.64	0.66	0.81	0.69	0.60	0.59	0.59

3.4 DISUSSION

Within the strategies for adaptive conservation, assisted migration (AM), in the form of assisted gene flow, involve a lower risk of maladaptation than range expansion and long-distance colonization, so it should be considered as a first option (Aitken and Bemmels, 2016). In this work, we have demonstrated that AM in *A. araucana* will be hard to implement due to the lack of donor populations that contain pre-adapted genotypes to future environmental conditions, or less maladaptive than the in-situ populations will be. Only 1 of 56 possible genetic interchanges could be beneficial, which represents only 1.7%; in addition, with a marginal reduction of the genetic offset in-situ. This makes us conclude that there is no scientific evidence to support an AM strategy that considers the movement of genotypes pre-adapted to future conditions in a moderate climate change scenario. To the best of our knowledge, this is the first study to apply the genomic approach to evaluate the feasibility of implementing an assisted gene flow strategy in a south American tree species.

Traditionally, the feasibility of assisted migration has been assessed via trait data measured in common garden or reciprocal transplant trials, mainly in conifers of the northern hemisphere. Outcomes of these studies has reported a common trend: the habitat of these species will occur at more polar latitudes and at higher altitudes, and the most negative impacts will be on the rear-edge in its range, (lower latitude, altitude, or xeric limit). Sáenz-Romero *et al.* (2021) suggests that the impact of climate change can be mitigated by using seed sources growing in sites 3 °C colder than the recipient population, or by raising its upper limit 400 meters upward. This is partially supported by the genetic offset predictions obtained for *A. araucana* (Varas-Myrik *et al.*, 2022), where the risk of maladaptation due to climate change is

greater in the low elevation foothills of the Andes Mountain range. In a similar study, Gray et al. (2011) recommended that seed sources be moved northward between 1 and 2 degrees of latitude for *Populus tremuloides* in Canada. In another conifer, (Nagamitsu *et al.*, 2015), working with *Abies sachalinensis* in Japan, recommended to move the plantations of this species to higher altitudes to mitigate the effects of climate change.

Interestingly, the MA rule about moving populations to higher altitudes and lower temperatures has not always been observed, and it is possible that adaptation strategies that consider climate change are not always applicable through a change of seed sources boundaries along environmental gradients (Malavasi *et al.*, 2018). In a recent work, (Notivol *et al.*, 2020), using multi-site provenance trials in *Pinus sylvestris* in Spain, concluded that, despite being theoretically possible, seed sourcing based on future climatic coincidence will not be feasible to implement due to the lack of donor populations containing pre-adapted genotypes. Both species (*A. araucana* and *P. sylvestris*) are long-lived conifers with similar demographic histories associated with the expansion and contraction of their distribution according to the dynamics of glaciation cycles (Tóth *et al.*, 2017; Villagrán, 2018). Despite the fact that *P. sylvestris* has a very wide distribution in Eurasia, the conclusions of (Notivol *et al.*, 2020) are based on populations located at the rear-edge of its distribution (Spain), similar to the reduced and fragmented distribution that *A. araucana* presents today. Another important factor when genotypes are translocated to different latitudes is photoperiod. Differences in flowering time can severely restrict crossing opportunities between individuals from different latitudes (Wadgymar *et al.*, 2015). It has already been shown that the molecular system that rules plants phenology called circadian clock has strong genetic control (de Dios *et al.*, 2017).

The findings presented in the previous paragraph are explained because forest species can be adapted to seasonal variation in temperature or precipitation, rather than annual averages. This is the case of *A. araucana* (Varas-Myrik *et al.*, 2022), where the environmental variable that explains by far the adaptive genomic divergence is the annual temperature range, that is, the maximum temperature of the warmest month minus the minimum temperature of the coldest month. Under this premise, it is possible that two distant populations of *A. araucana* are growing in environments with the same mean annual temperature, but whose annual temperature ranges are very different. Therefore, making AM recommendations based only on the mean annual temperature, geographical distance or elevation is not advisable in this species.

Local adaptation in trees can be more complex than what has been observed in field trials to date, encompassing other different climatic drivers, with different levels of influence and in nonlinear gradients, which was detected by our adaptive divergence model. Landscape genomics can analyze changes in adaptive allele frequencies along environmental gradients that span different types of variables (annual average, seasonality and limiting factors). These genotype-environment relationships are the result of natural selection over many reproductive generations (evolutionary fitness). Furthermore, the effects of neutral evolutionary features that confound the signals of natural selection are considered both in the models for the detection of putatively adaptive loci and in the model of adaptive divergence. A disadvantage of field trials to assess assisted migration in the face of climate change is the difficulty of subjecting the species under study to future climatic scenarios, since these scenarios may not be found actually in nature. Furthermore, evolutionary fitness is not necessarily reflected by growth measurements in common garden tests in a relatively short time (Gray *et al.*, 2011), which means that these trials probably do not reflect long-term evolutionary fitness, since some environments require survival adaptations that result in a trade-off with growth-maximizing adaptations, which can confound the local adaptation signals observed in growth-associated traits.



3.4.1 *In-situ* conservation

As we saw in previous paragraphs, the genetic exchange between adaptive zones does not report any benefit in the face of climate change. However, one of the goals of assisted migration is also the increase of total genetic diversity. For this, total genetic variation in the donor populations must be carefully evaluated to guarantee the greater adaptability in the populations established by MA (Hoffmann *et al.*, 2021). This greater total genetic variation can support AM established populations to cope with new stressors and increase adaptive potential (Burton and Burton, 2002). Considering that the selection of the founder individuals only to maximize overall genetic diversity could reduce the adaptive traits and increase frequency of phenotypes poorly adapted to the new environments, we propose a genetic exchange within the adaptive zones. This guarantees an increase in genetic diversity and adaptive potential, while reducing the risks of maladaptation. We also propose that four adaptive zones resolution zones could be most suitable, since in this there is only one zone for the Coast and 3 for the Andes. This would avoid exchanging materials between both mountain ranges, which has been recommended in a recent work by (McIntosh *et al.*, 2022).

LITERATURE CITED

- Adams, H.D., Zeppel, M.J., Anderegg, W.R., Hartmann, H., Landhäusser, S.M., Tissue, D.T., Huxman, T.E., Hudson, P.J., Franz, T.E., Allen, C.D., 2017. A multi-species synthesis of physiological mechanisms in drought-induced tree mortality. *Nature ecology & evolution* 1, 1285-1291.
- Aitken, S.N., Bemmels, J.B., 2016. Time to get moving: assisted gene flow of forest trees. *Evolutionary Applications* 9, 271-290.
- Aitken, S.N., Whitlock, M.C., 2013. Assisted gene flow to facilitate local adaptation to climate change. *Annual Review of Ecology, Evolution, and Systematics* 44, 367-388.
- Burton, P.J., Burton, C.M., 2002. Promoting genetic diversity in the production of large quantities of native plant seed. *Ecological Restoration* 20, 117-123.
- Chen, Z., Grossfurthner, L., Loxterman, J.L., Masingale, J., Richardson, B.A., Seaborn, T., Smith, B., Waits, L.P., Narum, S.R., 2021. Applying genomics in assisted migration under climate change: framework, empirical applications, and case studies. *Evolutionary Applications* n/a.
- de Dios, V.R., Gessler, A., Ferrio, J.P., Alday, J.G., Bahn, M., Del Castillo, J., Devidal, S., García-Muñoz, S., Kayler, Z., Landais, D., 2017. Circadian rhythms regulate the environmental responses of net CO₂ exchange in bean and cotton canopies. *Agricultural and Forest Meteorology* 239, 185-191.
- Dumroese, R.K., Williams, M.I., Stanturf, J.A., Clair, J.B.S., 2015. Considerations for restoring temperate forests of tomorrow: forest restoration, assisted migration, and bioengineering. *New Forests* 46, 947-964.
- Evans, P.M., Newton, A.C., Cantarello, E., Martin, P., Sanderson, N., Jones, D.L., Barsoum, N., Cottrell, J.E., A'Hara, S.W., Fuller, L., 2017. Thresholds of biodiversity and ecosystem function in a forest ecosystem undergoing dieback. *Scientific reports* 7, 1-9.
- Fitzpatrick, M.C., Keller, S.R., 2015. Ecological genomics meets community-level modelling of biodiversity: Mapping the genomic landscape of current and future environmental adaptation. *Ecology Letters* 18, 1-16.
- Fitzpatrick, M.C., Mokany, K., Manion, G., Lisk, M., Ferrier, S., Nieto-Lugilde, D., Fitzpatrick, M.M.C., Rcpp, L., Rcpp, I., 2021. Package 'gdm'. A toolkit with functions to fit, plot, and summarize Generalized Dissimilarity Models: CRAN Repository, R. Version 1.4.2.2.

- Gray, L.K., Gylander, T., Mbogga, M.S., Chen, P.-y., Hamann, A., 2011. Assisted migration to address climate change: recommendations for aspen reforestation in western Canada. *Ecological applications* 21, 1591-1603.
- Gugger, P.F., Liang, C.T., Sork, V.L., Hodgskiss, P., Wright, J.W., 2018. Applying landscape genomic tools to forest management and restoration of Hawaiian koa (*Acacia koa*) in a changing environment. *Evolutionary Applications* 11, 231-242.
- Havens, K., Vitt, P., 2011. Assisted Migration of Plants: Changes in Latitudes, Changes in Attitudes. In, GPPC 2011.
- Hoffmann, A.A., Miller, A.D., Weeks, A.R., 2021. Genetic mixing for population management: From genetic rescue to provenancing. *Evolutionary applications* 14, 634-652.
- Jia, K.H., Zhao, W., Maier, P.A., Hu, X.G., Jin, Y., Zhou, S.S., Jiao, S.Q., El-Kassaby, Y.A., Wang, T., Wang, X.R., 2020. Landscape genomics predicts climate change-related genetic offset for the widespread *Platycladus orientalis* (Cupressaceae). *Evolutionary Applications* 13, 665-676.
- Li, Y., Zhang, X.-X., Mao, R.-L., Yang, J., Miao, C.-Y., Li, Z., Qiu, Y.-X., 2017. Ten years of landscape genomics: challenges and opportunities. *Frontiers in Plant Science* 8, 2136.
- Malavasi, M.d.M., Davis, A.S., Malavasi, U.C., 2018. TREE SEED SOURCING FOR LANDSCAPE RESTORATION UNDER CLIMATE CHANGES. *Ciência Florestal* 28, 446-455.
- McIntosh, M., González-Campos, J., Demaree, P., Toro-Salamanca, O., Ipinza, R., Bustamante-Sánchez, M.A., Hasbún, R., Nelson, C.R., 2022. Trait variation between and within Andes and coastal mountain ranges in the iconic South American tree *Araucaria araucana* in Chile. *bioRxiv*, 2022.2001.2004.474828.
- Menezes-Silva, P.E., Loram-Lourenço, L., Alves, R.D.F.B., Sousa, L.F., Almeida, S.E.d.S., Farnese, F.S., 2019. Different ways to die in a changing world: Consequences of climate change for tree species performance and survival through an ecophysiological perspective. *Ecology and evolution* 9, 11979-11999.
- Nagamitsu, T., Shimada, K.-i., Kanazashi, A., 2015. A reciprocal transplant trial suggests a disadvantage of northward seed transfer in survival and growth of Japanese red pine (*Pinus densiflora*) trees. *Tree genetics & genomes* 11, 813.
- Notivol, E., Santos-del-Blanco, L., Chambel, R., Climent, J., Alía, R., 2020. Seed Sourcing Strategies Considering Climate Change Forecasts: A Practical Test in Scots Pine. *Forests* 11, 1222.

- Premoli, A., Quiroga, P., Gardner, M., 2013. Premoli, A., Quiroga, P. & Gardner, M. 2013. *Araucaria araucana*. The IUCN Red List of Threatened Species 2013: e.T31355A2805113.
<https://dx.doi.org/10.2305/IUCN.UK.2013-1.RLTS.T31355A2805113.en>. Downloaded on 11 August 2021. In.
- Radchuk, V., Reed, T., Teplitsky, C., Van De Pol, M., Charmantier, A., Hassall, C., Adamík, P., Adriaensen, F., Ahola, M.P., Arcese, P., 2019. Adaptive responses of animals to climate change are most likely insufficient. *Nature communications* 10, 1-14.
- Rossetto, M., Bragg, J., Kilian, A., McPherson, H., van der Merwe, M., Wilson, P.D., 2019. Restore and renew: A genomics-era framework for species provenance delimitation. *Restoration Ecology* 27, 538-548.
- Sáenz-Romero, C., O'Neill, G., Aitken, S.N., Lindig-Cisneros, R., 2021. Assisted Migration Field Tests in Canada and Mexico: Lessons, Limitations, and Challenges. *Forests* 12, 9.
- Shryock, D.F., Washburn, L.K., DeFalco, L.A., Esque, T.C., 2020. Harnessing landscape genomics to identify future-climate resilient genotypes in a desert annual. *Molecular Ecology*.
- Supple, M.A., Bragg, J.G., Broadhurst, L.M., Nicotra, A.B., Byrne, M., Andrew, R.L., Widdup, A., Aitken, N.C., Borevitz, J.O., 2018. Landscape genomic prediction for restoration of a *Eucalyptus* foundation species under climate change. *Elife* 7, e31835.
- Tóth, E.G., Köbölkuti, Z.A., Pedryc, A., Höhn, M., 2017. Evolutionary history and phylogeography of Scots pine (*Pinus sylvestris* L.) in Europe based on molecular markers. *Journal of Forestry Research* 28, 637-651.
- Varas-Myrik, A., Sepúlveda-Espinoza, F., Fajardo, A., Alarcón, D., Toro-Núñez, Ó., Castro-Nallar, E., Hasbún, R., 2022. Predicting climate change-related genetic offset for the endangered southern South American conifer *Araucaria araucana*. *Forest Ecology and Management* 504, 119856.
- Villagrán, C., 2018. Biogeografía de los bosques subtropical-templados del sur de Sudamérica. Hipótesis históricas. *Magallania (Punta Arenas)* 46, 27-48.
- Wadgyamar, S.M., Cumming, M.N., Weis, A.E., 2015. The success of assisted colonization and assisted gene flow depends on phenology. *Global change biology* 21, 3786-3799.

CAPÍTULO IV

Aplicando el modelo de divergencia adaptativa para investigar zonas de acogida potenciales fuera de la distribución natural de *Araucaria araucana*

En los capítulos anteriores hemos ajustado un modelo de divergencia adaptativa con la finalidad de responder 2 preguntas de investigación: a) ¿Cuáles son las zonas del nicho de *A. araucana* que se verán más afectadas por el offset genético? y b) ¿Es posible implementar una estrategia de flujo genético asistido en esta especie incorporando genotipos pre-adaptados a las zonas más vulnerables al cambio climático? En este capítulo responderemos la siguiente pregunta: c) ¿en qué zonas del territorio chileno o argentino se encuentran las mejores condiciones para la colonización asistida de *A. araucana*?

4.5 Metodología

Para responder la pregunta antes citada, aplicamos el modelo de divergencia adaptativa (GDM) fuera del rango natural de la especie, utilizando el agrupamiento de 4 zonas adaptativas construido en el capítulo III, y así respondiendo la pregunta separadamente para cada una de dichas zonas. Para hacer predicciones más realistas y científicamente basadas, primero creamos un rango geográfico (máscara) dentro del cual proyectar los niveles de offset por zona, es decir, el potencial de un área para acoger poblaciones de *Araucaria araucana* dado el cambio climático (Área de acogida potencial). Todas las proyecciones climáticas para este capítulo están basadas en worldclim 1.4 (CMIP5), RCP4.5 (escenario moderado) y para el año 2070. Esta máscara de proyección se construyó considerando los máximos y mínimos climáticos del nicho ecológico de la especie (Alarcón and Cavieres, 2015), para las variables contenidas en el modelo de divergencia adaptativa (Bio07, Bio13, Bio14, Elevación y pH), pero ampliando el rango un 10% hacia arriba y un 10% hacia abajo (Tabla 6), de forma de considerar la plasticidad fenotípica. Con estos límites climáticos, nos aseguramos de no extrapolar el modelo de forma excesiva, ya que esto implica una mayor probabilidad de estimaciones erróneas.

Luego, para estimar el offset dentro de la máscara, para cada una de las 4 zonas adaptativas, se siguió la metodología descrita en el capítulo anterior, es decir, se calculó el offset entre 2 pilas de ráster (dos stacks), una conteniendo los promedios de cada variable ambiental (clima actual), y otra conteniendo los a nivel de pixel (sin promediar) de las variables

Variable	Min	Max	rango	10%	Min-B	Max-B
Bio07 (°C*10)	187	263	76	7.6	179	271
Bio13 (mm)	136	511	375	37.5	99	549
Bio14 (mm)	14	64	50	5	9	69
Elev (msnm)	107	3323	3216	321.6	-215	3645

Table 6. Rangos climáticos utilizados para construir la máscara de proyección de zonas de acogida potenciales para migración asistida de *A. araucana*. Los rangos utilizados fueron Min-B y Max-B, que corresponden al rango actual del nicho más y menos 10%.

ambientales futuras (CMIP5, RCP4.5, 2070). El resultado fue, por lo tanto, el valor del offset genético entre las condiciones promedio actuales de esa zona adaptativa (Figura 8), con las condiciones futuras de cada pixel de la máscara de potencial acogida

4.6 Resultados

Podemos apreciar que la forma de la máscara resultante tiene una forma de triángulo invertido que incluye a territorio argentino (Figura 11), y también la zona de Coyhaique. Como primer resultado vemos que los valores de offset van de 0.5 a 0.9, lo que constituye una observación desalentadora, ya que el valor mínimo es mayor que el offset promedio para ese escenario dentro del nicho de la especie. Al observar los patrones espaciales de offset genético se puede deducir que, por ejemplo, la zona de Coyhaique no representa la mejor zona de acogida para la especie, exceptuando la zona 3, la cual corresponde a la zona de la costa. La zona adaptativa altoandina (zona 1), que observó el menor offset *in-situ*, presenta una zona de acogida mejor en los sectores altos cercanos a Bariloche, con una presencia preponderante en territorio argentino. La zona andina de altitud media (zona 2) presenta una zona de acogida similar a la zona 1. Para la zona de andina piedmont, (zona 4) que presenta mayor offset *in-situ*, se observa un patrón que recorre el sector medio de la zona de acogida potencial, comprendiendo las altas cumbres andinas del límite Chile-Argentina.

Exploramos los sectores donde el offset resultó menor que el offset *in-situ* promedio de cada zona adaptativa. Sorprendentemente, sólo la zona 4 presentó superficie disponible. Para el resto de las zonas no se observaron lugares con offset menor al offset *in-situ* de la zona (Figura 12).

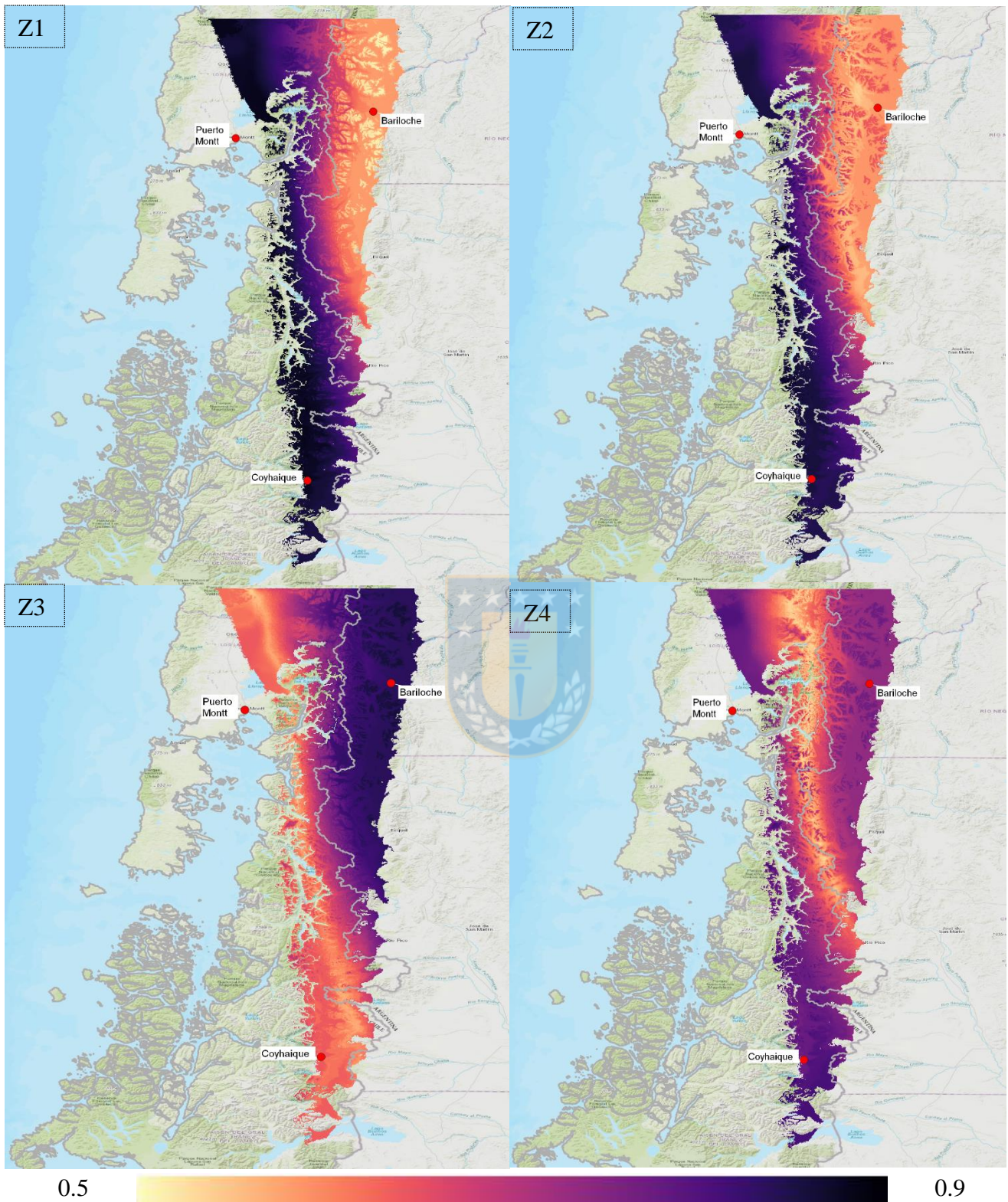
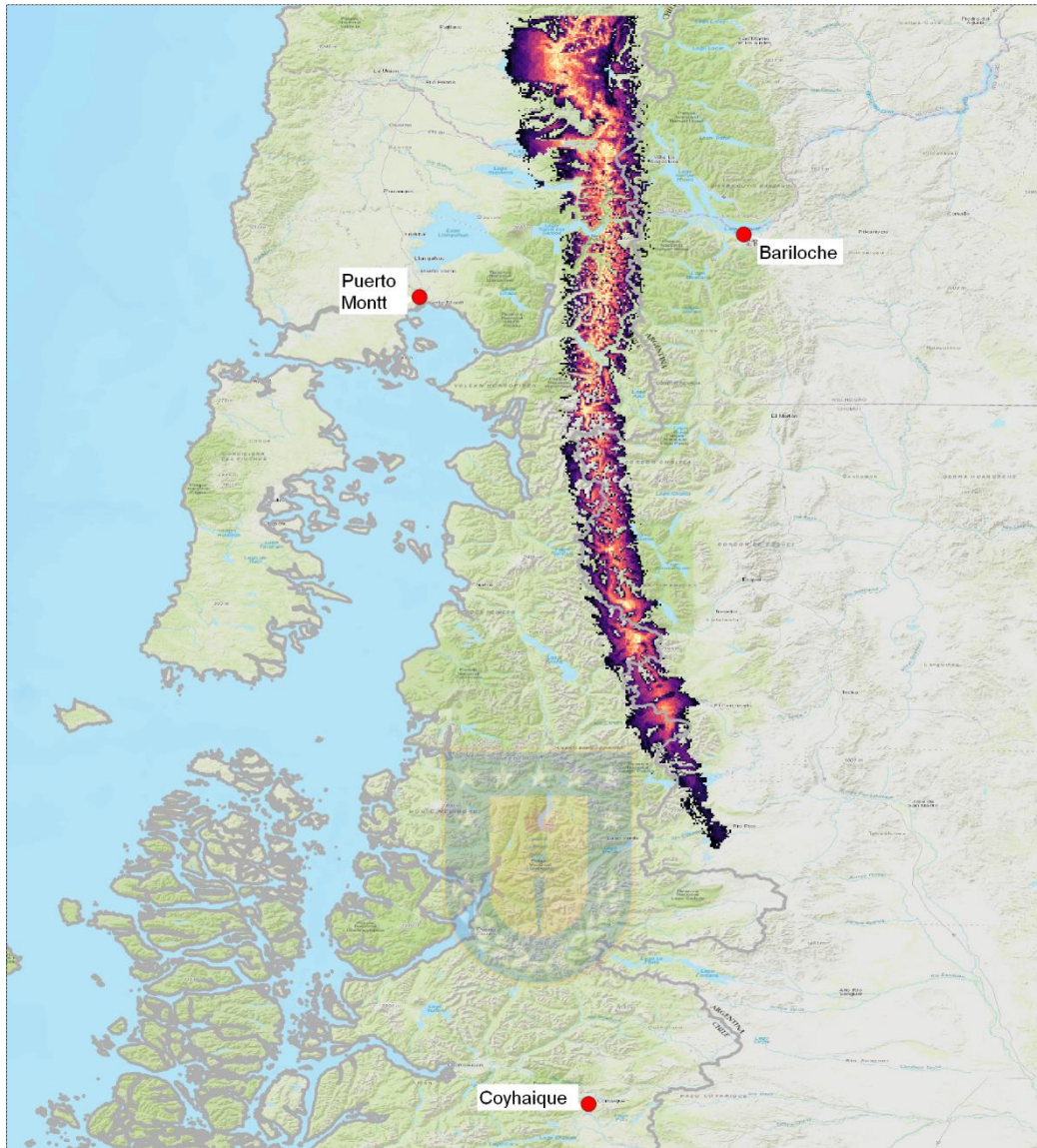


Figure 11. Offset genético de cada una de las 4 zonas adaptativas respecto del área de acogida potencial. Zonas más claras representan un potencial mayor para la colonización asistida de *A. araucana*, para la zona adaptativa correspondiente



0.53



0.63

Figure 12. Zona candidata de acogida para la zona adaptativa 4 (4 zonas). Se muestran los lugares donde el offset es menor a 0.63, que es el offset in-situ promedio de la zona 4. Colores más claros representan un mayor potencial de acogida.

4.7 DISCUSIÓN

Al igual que para el flujo genético asistido, se observan pocas posibilidades para la colonización asistida de *A. araucana*. El criterio para que un área fuera de la distribución de una especie sea considerada como zona de acogida potencial, es que el offset genético dado el cambio climático sea menor que el offset que experimentará la especie en su hábitat natural. Sin embargo, no existe un patrón que pueda ser denominado como tendencia para asumir que podremos encontrar zonas de acogida fuera de la distribución natural de *A. araucana*, que provean de condiciones más adecuadas para la supervivencia de la especie dado el cambio climático. Solamente para la zona 4, la de mayor riesgo de desadaptación, existe una franja de acogida potencial que recorre las zonas altas de Los Andes desde Futrono en Chile (-40,00° latitud sur), hasta Río Pico en Argentina (-44,27° latitud sur), predominantemente en las regiones de Los Ríos y Los Lagos, y que no incluye a la región de Aysén, cuya aptitud parece más adecuada para la zona 3 de la Costa (Fig. 11).

En este sentido, parece lógico pensar que el único camino para la migración de *A. araucana* es más al sur y más alto. Esto está apoyado por la evidencia de que solamente para la zona 4, la de mayor offset y la de menor altitud, pudimos encontrar una pequeña franja cordillerana potencialmente apta para la migración asistida. Esto también concuerda con los resultados del capítulo anterior, donde el único intercambio genético beneficioso observado fue desde la zona 5 a la zona 7 para el escenario de 8 zonas adaptativas. Ambas zonas (5 y 7), forman parte de la zona 4 en el escenario de 4 zonas adaptativas. La dirección del movimiento fue desde la zona de menor altitud hacia la de mayor altitud.

Para las zonas de Mayor altitud y la zona costera, aparentemente no hay lugares con la suficiente altitud para ser consideradas como zonas de acogida. Esto puede darse por 2 factores principales. Primero, el grado de adaptación de *A. araucana* al ambiente es muy alto, y esto provoca que la especie sea más sensible a los cambios climáticos. Segundo, la magnitud del cambio climático, incluso en este escenario moderado que estamos considerando, es muy grande. Esto provoca que la MA sea solamente posible para los sitios de baja altitud, pero a la vez los de mayor riesgo futuro.

Bibliografía

Alarcón, D., Cavieres, L.A., 2015. In the right place at the right time: habitat representation in protected areas of South American *Nothofagus*-dominated plants after a dispersal constrained climate change scenario. PloS One 10, e0119952.



5 CONCLUSION GENERAL

La estructura genética poblacional en *A. araucana* confirmó la existencia de 2 grupos ancestrales, uno presente mayormente en la cordillera de la Costa, y el segundo fue predominante en las poblaciones andinas. Los resultados del análisis de asociación ambiental demostraron la existencia de variación genética adaptativa en el genoma de *A. araucana*, correspondiente aproximadamente al 2% del genoma muestreado, y que, según el análisis de redundancia, fue explicada en gran parte por los efectos confundidos del ambiente, la ancestría genética y la geografía.

El modelo no lineal espacialmente explícito de divergencia adaptativa explicó el 85% de la distancia genómica adaptativa individual, siendo la variable rango anual de temperatura (°C) la más influyente. Esto implica un alto grado de adaptación local de *A. araucana* al ambiente, especialmente influido por el rango anual de temperatura, cuya variación futura en un escenario de cambio climático implicará un gran riesgo de desadaptación, concentrado principalmente en los piedemontes de la distribución andina sur de la especie. Por el contrario, las poblaciones ubicadas en las zonas altas de la cordillera de los Andes sufrirán un menor grado de desadaptación producida por el cambio climático.

La migración asistida, en la forma de flujo genético asistido, con el objetivo mejorar la aptitud promedio de las poblaciones frente al cambio climático, será difícil de implementar por la carencia de poblaciones donantes que contengan genotipos pre-adaptados a las condiciones futuras de las poblaciones receptoras. Respecto de la colonización asistida, se investigó si existían zonas fuera de la distribución de la especie, que pudieran presentar condiciones climáticas futuras adecuadas para la traslocación de *A. araucana*, y se

concluyó que solamente para la zona de mayor riesgo de desadaptación, las cumbres andinas entre Futrono y Río Pico en Argentina, pueden representar zonas de acogida candidatas. Para el resto de las zonas adaptativas, no existen sitios cuya configuración climática futura representen un menor riesgo de desadaptación que al que estarán sometidas sus poblaciones naturales *in-situ*.

Como estrategia de conservación *in-situ*, es recomendable intercambiar material genético dentro de las zonas adaptativas, de forma de aumentar la diversidad genética sin mayor riesgo de desadaptación. Como estrategia de Conservación *ex-situ*, se recomienda probar material genético de la zona de mayor riesgo de desadaptación en los sectores altos andinos al sur de la distribución de *A. araucana*, según lo indicado por los mapas incluidos en este trabajo.



Finalmente, es necesario señalar que esta aproximación genómica para la modelación de la divergencia adaptativa es aplicada por primera vez en Sudamérica en este trabajo, sin embargo, no está exenta de desafíos metodológicos. La validación de las funciones de los genes involucrados en la adaptación es uno de ellos, para lo cual el contar con genomas de referencia debiera ser un objetivo de investigación a corto plazo. Por otro lado, los patrones de distribución espacial de las frecuencias de alelos adaptativos en el paisaje es el resultado de largos períodos de selección natural y procesos neutrales, por lo que su validación sólo sería posible en un período suficientemente largo como para que dichas influencias se plasmen en la variación de las frecuencias alélicas adaptativas. Sin embargo, la base teórica de la genómica del paisaje es lo suficientemente robusta como para aprovechar las oportunidades que brinda, principalmente la ventaja de no requerir un largo tiempo de evaluación.

