

“Estudio de Señales de Adaptación Biológica en Poblaciones Descendientes de grupos Cazadores Recolectores Marinos de la Patagonia”

Tesis

**Entregada A La
Universidad De Chile
En Cumplimiento Parcial De Los Requisitos
Para Optar Al Grado De**

**Doctor en Ciencias Biológicas
Mención Ecología y Biología Evolutiva**

Facultad De Ciencias

Por

Patricio Alejandro Pezo Valderrama

Julio, 2019

**Director de Tesis:
Dr. Mauricio Moraga Vergara**

FACULTAD DE CIENCIAS
UNIVERSIDAD DE CHILE
INFORME DE APROBACION
TESIS DE DOCTORADO

Se informa a la Escuela de Postgrado de la Facultad de Ciencias que la Tesis de Doctorado presentada por el candidato.

Patricio Alejandro Pezo Valderrama

Ha sido aprobada por la comisión de Evaluación de la tesis como requisito para optar al grado de Doctor en Ciencias Biológicas mención Ecología y Biología Evolutiva, en el examen de Defensa Privada de Tesis rendido el día

Director de Tesis:

Dr. Mauricio Moraga

Co-Directores de Tesis

Dr. Ricardo Verdugo

Dr. Elie Poulin

Comisión de Evaluación de la Tesis

Dr. Marco Méndez (P)

Dr. Luis Felipe Hinojosa

Dr. David Véliz

Dr. Michelle de Saint Pierre

“A mi esposa Sandra, mi hermana Cata
y mi hermano Manuel, quien cree que la evolución es un mito”



Biólogo egresado de la Universidad de Concepción el año 2010. El 2012 ingresé al Programa de Magíster en Ciencias Biológicas de la Facultad de Ciencias en la Universidad de Chile, egresando el año 2014. Ese mismo año ingresé al programa de Doctorado en Ciencias Biológicas, mención Ecología y Biología Evolutiva en la Facultad de Ciencias de la Universidad de Chile. Mi área de especialización abarca diferentes aspectos en la genética de poblaciones, como el estudio de linajes uniparentales (ADN mitocondrial y Cromosoma Y), así como también la estructura poblacional y procesos de adaptaciones biológicas en poblaciones humanas.

AGRADECIMIENTOS

A mi tutor, el Dr. Mauricio Moraga el cual me recibió hace 10 años atrás, abriéndome las puertas de su laboratorio y por pasar estos largos años llenos de conocimiento, críticas, ciencia, e incluso cariño.

A mis cotutores el Dr. Elie Poulin por sus siempre acertivos comentarios y críticas a este trabajo. Así como también, al Dr. Ricardo Verdugo por facilitar su laboratorio, equipo humano y por sus constantes intentos de mejorar a su equipo.

A la comisión evaluadora por sus críticas constructivas y acertadas correcciones durante todo el proceso de elaboración de la tesis.

A mis padres, por todo ese gasto energético, monetario, pero por sobretodo amoroso. Como lo he dicho muchas veces, todo lo que he logrado hasta ahora es gracias a ellos.

A mis hermanos Manuel y Catalina, mis primeros compañeros de vida. En cada momento que tengo que defender estamos en planos distintos, hoy cada uno empezando sus propios caminos que estarán llenos de desafíos, angustias, pero también éxitos y alegrías. Los amo.

Por último, a Sandra, mi compañera de vida y con quien forme el proyecto mas importante y el que probablemente requiere mas dedicación y amor que cinco años (casi seis) de tesis. Gracias por ser el apoyo incondicional y por esta hermosa familia junto a la Pumara.

Índice de Materias

1. Resumen	1
2. Abstract	4
3. Introducción	7
3.1. <i>Antecedentes Generales</i>	7
3.2. <i>Poblamiento de América</i>	8
3.2.1. Poblamiento del Continente Americano	8
3.2.2. Poblamiento de la Patagonia	8
3.2.3. Registros Arqueológicos del Poblamiento de la Patagonia	9
3.3. <i>Grupos Cazadores Recolectores Marinos en la Patagonia</i>	11
3.4. <i>Paleoclimatología</i>	12
3.5. <i>Señales de Adaptación</i>	12
3.5.1. Señales de Adaptación en América	14
3.5.2. Genes bajo Selección en Respuesta a Frío y Dieta	15
3.6. <i>Señales de Adaptación post-contacto Europeo</i>	17
3.7. <i>Abordaje Genómico al Estudio de Poblaciones</i>	18
3.7.1. Estructuración Poblacional	19
3.7.2. Análisis de Ancestría	19
3.7.3. Análisis de Poblaciones basados en Ancestría	20
3.7.4. Análisis basados en Simulaciones	22
3.7.5. Análisis de Selección	23
3.8. <i>Hipótesis de Trabajo</i>	28
3.9. <i>Objetivos</i>	30
4. Metodología	31
4.1. <i>Poblaciones Analizadas y Genotipificación de las muestras</i>	31
4.2. <i>Control de Calidad</i>	34
4.2.1. Filtro de Individuos y de Variantes	34
4.2.2. Comparación Plataforma Axiom y MEGA	34
4.3. <i>Análisis de Estructuración Poblacional mediado por Ancestrías</i>	35
4.3.1. Análisis de Ancestría Global	35
4.3.2. Análisis de Componentes Principales	36
4.3.3. Análisis de Ancestría Local	36
4.3.4. TRACTs	39
4.3.5. Análisis de Componentes Principales específico de Ancestría	¡Error!
Marcador no definido.	
4.4. <i>Enmascaramiento Regiones del Genoma No Nativos</i>	40
4.5. <i>Análisis de Selección</i>	42
4.5.1. Populations Branch Statistic (PBS)	42
4.5.2. integrated Haplotype Score (iHS)	44
4.5.3. Detección de Loci bajo Selección mediante Ancestría	45

4.6.	<i>Identificación de Regiones bajo Selección</i>	46
4.6.1.	Marcadores y Regiones bajo Selección	46
4.6.2.	Identificación de Genes bajo Selección	46
4.6.3.	Enriquecimiento de Procesos Biológicos	47
4.6.4.	Comparación de Regiones bajo Selección con Genes asociados a Frío y Dieta	48
5.	Resultados	49
5.1.	<i>Control de Calidad y Depuración de los Datos</i>	49
5.1.1.	Patagonia (Axiom)	49
5.1.2.	Patagonia Austral (MEGA)	51
5.2.	<i>Comparación Plataformas Axiom y MEGA</i>	53
5.3.	<i>Análisis de Estructuración Poblacional mediante Ancestrías</i>	55
5.3.1.	Análisis de Ancestría Global	55
5.3.2.	Análisis de Componentes Principales	59
5.3.3.	Análisis de Ancestría Local	62
5.3.4.	Modelamiento de los procesos de mestizaje en la Patagonia	67
5.4.	<i>Enmascaramiento del Genoma no Nativo</i>	72
5.5.	<i>Análisis de Selección</i>	75
5.5.1.	Señales de Selección pre-contacto Europeo	75
5.5.2.	Señales de Selección post-contacto Europeo	89
5.6.	<i>Enriquecimiento de Procesos Biológicos</i>	92
5.6.1.	Population Branch Statistic (PBS)	93
5.6.2.	integrated Haplotype Score (iHS)	98
5.6.3.	Detección de Loci bajo Selección mediante Ancestría	100
5.7.	<i>Genes Candidatos a Selección Positiva en la Patagonia</i>	106
5.7.1.	Comparación de Regiones bajo Selección con Genes asociados a Frío y Dieta	110
5.7.2.	Procesos y Rutas Metabólicas en la Patagonia	117
5.8.	<i>Análisis sobre Cazadores Recolectores Marítimos en la Patagonia Austral</i>	121
5.8.1.	Análisis de Estructuración Poblacional mediante Ancestrías	121
5.8.2.	Análisis de Selección	134
5.8.3.	Enriquecimiento de procesos biológicos	145
5.8.4.	Genes Candidatos a Selección Positiva en la Patagonia Austral	150
6.	Discusión	160
6.1.	<i>Estructuración Poblacional en Individuos de la Patagonia a través de Ancestrías</i>	160
6.2.	<i>Procesos de Mestizaje en las Poblaciones de la Patagonia post-contacto Europeo</i>	163
6.3.	<i>Análisis de Selección</i>	167
6.3.1.	Señales de Selección pre-contacto Europeo	167
6.3.2.	Señales de Selección post-contacto Europeo	177
6.4.	<i>Historia de los Procesos de Adaptación Biológica en las Poblaciones de la Patagonia: Señales pre y post-contacto Europeo.</i>	179
6.5.	<i>Consideraciones Finales</i>	181
7.	Conclusiones	184
8.	Bibliografía	186

Lista de Tablas

<i>Tabla I. Lista de genes asociados a señales de selección en respuesta a frío, dieta y buceo.</i> _____	17
<i>Tabla II. Descripción de las poblaciones utilizadas en este estudio.</i> _____	33
<i>Tabla III. Descripción de las poblaciones de referencia utilizadas para este estudio..</i> _____	37
<i>Tabla IV. Comparación entre promedios de ancestrías por población para ADMIXTURE con filtro por parentesco.</i> _____	58
<i>Tabla V. Cuadro con los valores de Likelihood para cada modelo de migración.</i> _____	68
<i>Tabla VI. Número de SNPs, ventanas y Genes analizadas para los escenarios A y B junto con los marcadores sobre el 1% de la distribución de los datos.</i> _____	75
<i>Tabla VII. Top 10 de los valores de PBS sobre el 1% de los datos observados para el escenario A por SNPs.</i> _____	77
<i>Tabla VIII. Top 10 de los valores de PBS sobre el 1% de los datos observados para el escenario A mediante ventanas de 200Kb.</i> _____	79
<i>Tabla IX. Top 10 de los valores de PBS sobre el 1% de los datos observados para el escenario B mediante el análisis de SNPs.</i> _____	80
<i>Tabla X. Top 10 de los valores de PBS sobre el 1% de los datos observados para el escenario B mediante el análisis de ventanas de 200Kb</i> _____	82
<i>Tabla XI. Número de SNPs, Ventanas y Genes analizados para iHS en poblaciones cazadores recolectores de la Patagonia.</i> _____	86
<i>Tabla XII. Top 10 de los valores de iHS sobre el 1% de los datos observados mediante el análisis de SNPs.</i> _____	87
<i>Tabla XIII. Top 10 de los valores de iHS sobre el 1% de los datos observados mediante análisis de ventanas de 200Kb.</i> _____	88
<i>Tabla XIV. Resumen de los promedios de ancestría a nivel genómico para los grupos cazadores recolectores de la Patagonia y sus respectivas medidas de dispersión.</i> _____	89
<i>Tabla XV. Resumen con el número de SNPs y genes para el análisis de Detección de Loci bajo Selección mediante Ancestrías.</i> _____	90
<i>Tabla XVI. Enriquecimiento de procesos biológicos para los grupos cazadores recolectores de la Patagonia a partir de los genes por sobre el 1% identificados a partir de PBS.</i> _____	94
<i>Tabla XVII. Enriquecimiento de procesos biológicos en ventanas de 200 kb para los grupos cazadores recolectores de la Patagonia, a partir de los genes por sobre el 1% identificados a partir de PBS.</i> ____	95

<i>Tabla XVIII. Enriquecimiento de procesos biológicos para los grupos cazadores recolectores marítimos de la Patagonia a partir de los genes por sobre el 1% identificados a partir de PBS.</i>	96
<i>Tabla XIX. Enriquecimiento de procesos biológicos en ventanas de 200 kb para los grupos cazadores recolectores marítimos de la Patagonia, a partir de los genes por sobre el 1% identificados a partir de PBS.</i>	97
<i>Tabla XX. Enriquecimiento de procesos biológicos para los grupos cazadores recolectores de la Patagonia a partir de los genes con valores de $iHS > 2.5$ por SNPs.</i>	98
<i>Tabla XXI. Enriquecimiento de procesos biológicos en ventanas de 200kb para los grupos cazadores recolectores de la Patagonia, a partir de los genes por sobre el 1% identificados a partir de iHS.</i>	99
<i>Tabla XXII. Enriquecimiento de procesos biológicos para los grupos cazadores recolectores de la Patagonia a partir de los genes por sobre el 1% identificados en aquellos Loci bajo selección mediante Ancestría.</i>	101
<i>Tabla XXIII. Enriquecimiento de procesos biológicos para los grupos cazadores recolectores de la Patagonia a partir de los genes por sobre el 1% identificados en aquellos Loci bajo selección mediante Ancestría Europea.</i>	103
<i>Tabla XXIV. Enriquecimiento de procesos biológicos para los grupos cazadores recolectores marítimos de la Patagonia a partir de los genes por sobre el 1% identificados en aquellos Loci bajo selección a través de Ancestría.</i>	104
<i>Tabla XXV. Enriquecimiento de procesos biológicos para los grupos cazadores recolectores marítimos de la Patagonia a partir de los genes por sobre el 1% identificados en aquellos Loci bajo selección a través de Ancestría Europea.</i>	106
<i>Tabla XXVI. Descripción de los genes candidatos para frío y dieta en la Patagonia con los análisis realizados con PBS e iHS (SNPs y ventenas de 200Kb).</i>	107
<i>Tabla XXVII. Genes involucrados en la ruta metabolica de síntesis y secreción de cortisol en los individuos de la Patagonia.</i>	118
<i>Tabla XXVIII. Genes involucrados en la ruta metabolica de termogénesis en los individuos de la Patagonia.</i>	120
<i>Tabla XXIX. Comparación entre promedios de ancestrías por población para ADMIXTURE con filtro por parentesco . En paréntesis se encuentran los errores estándares para cada ancestría.</i>	124
<i>Tabla XXX. Número de SNPs, ventanas y genes analizadas para los cazadores recolectores marítimos de la Patagonia Austral, junto con los marcadores sobre el 1% de la distribución de los datos para PBS.</i>	134
<i>Tabla XXXI. Top 10 de los valores de PBS sobre el 1% de los datos observados para los grupos cazadores recolectores marítimos mediante el análisis de SNPs.</i>	135
<i>Tabla XXXII. Top 10 de los valores de PBS sobre el 1% de los datos observados para los grupos cazadores recolectores marítimos mediante el análisis de ventanas de 200Kb.</i>	137
<i>Tabla XXXIII. Número de SNPs, Ventanas y Genes analizados para iHS en poblaciones cazadores recolectores marítimas de la Patagonia.</i>	140

<i>Tabla XXXIV. Top 10 de los valores de iHS sobre el 1% de los datos observados para los grupos cazadores recolectores marítimos de la Patagonia mediante el análisis por SNPs.</i>	141
<i>Tabla XXXV. Top 10 de los valores de iHS sobre el 1% de los datos observados para los grupos cazadores recolectores marítimos de la Patagonia mediante el análisis por ventanas de 200Kb.</i>	142
<i>Tabla XXXVI. Promedio, desviación y error estándar para las ancestrías Amerindias, Europeas y Africanas para las poblaciones Kawéskar y Yámana en la Patagonia Austral.</i>	143
<i>Tabla XXXVII. Resumen con el número de SNPs y genes para el análisis de Detección de Loci bajo Selección para las ancestrías Amerindias y Europeas.</i>	144
<i>Tabla XXXVIII. Enriquecimiento de procesos biológicos para los grupos cazadores recolectores de la Patagonia Austral a partir de los genes por sobre el 1% identificados a partir de PBS.</i>	145
<i>Tabla XXXIX. Enriquecimiento de procesos biológicos en ventanas de 200 kb para los grupos cazadores recolectores marítimos de la Patagonia Austral, a partir de los genes por sobre el 1% identificados a partir de PBS.</i>	147
<i>Tabla XL. Enriquecimiento de procesos biológicos para los grupos cazadores recolectores de la Patagonia Austral a partir de los genes con valores de iHS > 2.5 por SNPs.</i>	147
<i>Tabla XLI. Enriquecimiento de procesos biológicos en ventanas de 200kb para los grupos cazadores recolectores de la Patagonia, a partir de los genes por sobre el 1% identificados a partir de iHS.</i>	148
<i>Tabla XLII. Enriquecimiento de procesos biológicos para los grupos cazadores recolectores marítimos de la Patagonia a partir de los genes por sobre el 1% identificados en aquellos Loci bajo selección mediante Ancestría Amerindia.</i>	149
<i>Tabla XLIII. Descripción de los genes candidatos dentro del top 10 para frío y dieta en la Patagonia con los análisis realizados con PBS e iHS (SNPs y ventanas de 200Kb).</i>	151
<i>Tabla XLIV. Genes involucrados en la ruta metabólica de síntesis y secreción del cortisol en los individuos de la Patagonia.</i>	158
<i>Tabla XLV. Genes involucrados en la ruta metabólica de señalamiento de apelin en los individuos de la Patagonia Austral.</i>	158

Lista de Figuras

<i>Figura 1. Esquema temporal con los sitios arqueológicos en la Patagonia (modificado de González 2003).</i>	10
<i>Figura 2. Adaptación a diversos ambientes durante el curso de la migración humana a lo largo del mundo. Extraído a partir de Tishkoff, 2015.</i>	14
<i>Figura 3. Homocigocidad de haplotipos extendido en el genoma.</i>	25
<i>Figura 4. Señales Genómicas de Selección Positiva.</i>	27
<i>Figura 5. Análisis de Ancestría local para el Genoma.</i>	28
<i>Figura 6. Proceso de enmascaramiento de los segmentos con ancestría local “No Nativos” para los datos imputados de la Patagonia.</i>	41
<i>Figura 7. Modelos de comparación utilizados en este estudio con sus respectivo cálculo de PBS.</i>	43
<i>Figura 8. Análisis de Control de Calidad para las muestras de la Patagonia.</i>	50
<i>Figura 9. Esquema con el número de SNPs e individuos utilizados para los principales análisis en este estudio.</i>	52
<i>Figura 11. Dendograma basado en la matriz de distancias Manhattan entre las plataformas de genotipificación de Axiom y MEGA.</i>	54
<i>Figura 12. Comparación del error de genotipificación para las plataformas de Axiom y MEGA.</i>	55
<i>Figura 13. Error de validación cruzada para el análisis de ADMIXTURE (cross-validation, CV).</i>	57
<i>Figura 14. Análisis de Ancestría Global sobre las poblaciones de la Patagonia.</i>	58
<i>Figura 15. Comparación de Ancestría Global para las muestras de la Patagonia..</i>	59
<i>Figura 16. Análisis de Componentes Principales sobre la muestra de la Patagonia.</i>	61
<i>Figura 17. Análisis de Componentes Principales sobre las muestras de la Patagonia.</i>	62
<i>Figura 18. Mapeo Genómico de la Ancestría Local en tres individuos de la Patagonia.</i>	64
<i>Figura 19. Test de Correlación para la inferencia de Ancestría Global realizado por ADMIXTURE y RFMix en las poblaciones de la Patagonia.</i>	66
<i>Figura 20. Test de Correlación para la inferencia de Ancestría Global realizado por ADMIXTURE y la ancestría local realizado por RFMix en las poblaciones de la Patagonia.</i>	67
<i>Figura 21. Análisis de TRACTs y tiempos de mestizaje en las poblaciones de la Patagonia.</i>	70

<i>Figura 22. Análisis de TRACTs y tiempos de mestizaje en las poblaciones de la Patagonia para las ancestrías Aymara, Mapuche y Europa.</i>	72
<i>Figura 23. Esquema del proceso de enmascaramiento de regiones No Nativas del genoma.</i>	74
<i>Figura 24. Intersección de genes encontrados bajo selección (top 1%) para los análisis realizados por SNPs (izquierda) y por ventanas de 200Kb (derecha).</i>	76
<i>Figura 25. Manhattan Plot para los valores de PBS sobre las poblaciones de la Patagonia para el escenario A (IBS; Pehuenche/Huilliche; Cazadores Recolectores).</i>	77
<i>Figura 26. Proporción de SNPs con valores de PBS sobre el 1% en ventanas de 200kb sobre las poblaciones de la Patagonia para el escenario A (IBS; Pehuenche/Huilliche; Cazadores Recolectores).</i>	78
<i>Figura 27. Manhattan Plot para los valores de PBS sobre las poblaciones de la Patagonia para el escenario B (IBS; Tehuelche; Cazadores Recolectores Marítimos).</i>	80
<i>Figura 28. Proporción de SNPs con valores de PBS sobre el 1% en ventanas de 200kb sobre las poblaciones de la Patagonia para el escenario B (IBS; Tehuelche; Cazadores Recolectores Marítimos).</i>	81
<i>Figura 29. Representación del proceso de filtro de individuos y SNPs de origen Amerindio para los análisis de iHS.</i>	83
<i>Figura 30. Número de individuos y SNPs para cada rango de componente Amerindio en grupos cazadores recolectores de la Patagonia.</i>	84
<i>Figura 31. Frecuencia de SNPs por ventana de 200 Kb para diferentes rangos de Ancestría Amerindia.</i>	85
<i>Figura 32. Valores de iHS por SNPs en poblaciones cazadores recolectores de la Patagonia.</i>	87
<i>Figura 33. Proporción de SNPs con valores de $iHS > 2.5$ dentro de cada ventana de 200Kb en poblaciones cazadoras recolectoras de la Patagonia.</i>	88
<i>Figura 34. Detección de loci bajo selección a lo largo del Genoma basado en ancestrías para los grupos cazadores recolectores de la Patagonia (Chono, Tehuelche, Kawéskar y Yámana).</i>	91
<i>Figura 35. Detección de loci bajo selección a lo largo del Genoma basado en ancestrías para los grupos cazadores recolectores marítimos de la Patagonia (Chono, Kawéskar y Yámana).</i>	92
<i>Figura 36. Esquema con la representación de rutas biológicas enriquecidas para los análisis de PBS sobre los SNPs en los individuos de la Patagonia.</i>	95
<i>Figura 37. Esquema con la representación de rutas biológicas enriquecidas para los análisis de iHS sobre las ventanas de 200 Kb en los individuos de la Patagonia.</i>	100
<i>Figura 38. Esquema con la representación de rutas biológicas enriquecidas en grupos cazadores recolectores de la Patagonia.</i>	102
<i>Figura 39. Esquema con la representación de rutas biológicas enriquecidas en grupos cazadores recolectores marítimos de la Patagonia.</i>	105

<i>Figura 40. Intersección de genes bajo selección entre el Top 10 de los cazadores recolectores de la Patagonia (azules, lado izquierdo de cada cromosoma) y el 1% de los cazadores recolectores marítimos de la Patagonia (naranjos, lado derecho de cada cromosoma).</i>	108
<i>Figura 41. Intersección de genes bajo selección entre el Top 10 de los cazadores recolectores de la Patagonia (azules, lado izquierdo de cada cromosoma) y el 1% de los cazadores recolectores marítimos de la Patagonia (naranjos, lado derecho de cada cromosoma).</i>	109
<i>Figura 42. Intersección de genes bajo selección entre el Top 10 de los cazadores recolectores marítimos de la Patagonia (naranja, lado derecho de cada cromosoma) y el 1% de los cazadores recolectores de la Patagonia (azules, lado izquierdo de cada cromosoma).</i>	110
<i>Figura 43. Comparación entre los genes asociados previamente a frío y dieta con las ventanas sobre el 1% de los valores de PBS. En azul se encuentran los genes asociados a frío y dieta descritos previamente.</i>	112
<i>Figura 44. Distribución de los valores de PBS por SNPs entre las regiones 40-46Mb del cromosoma 2.</i>	113
<i>Figura 45. Distribución de los valores de PBS por SNPs entre las regiones 95-97Mb del cromosoma 14.</i>	113
<i>Figura 46. Comparación de valores de PBS por SNPs entre los cazadores recolectores de la Patagonia en regiones cromosómicas de genes descritos previamente en literatura para otras poblaciones nativas.</i>	115
<i>Figura 47. Comparación de valores de PBS por SNPs entre los cazadores recolectores marítimos de la Patagonia en regiones cromosómicas de genes descritos previamente en literatura para otras poblaciones nativas.</i>	116
<i>Figura 48. Análisis de iHS para SNPs en grupos Cazadores Recolectores de la Patagonia.</i>	117
<i>Figura 49. Ruta metabólica descrita para la síntesis y secreción del cortisol (hsa04927) extraído a partir del Mapa de rutas KEGG (www.kegg.jp).</i>	119
<i>Figura 50. Ruta metabólica descrita para la termogénesis (hsa04714) extraído a partir del Mapa de rutas KEGG (www.kegg.jp).</i>	120
<i>Figura 51. Error de validación cruzada para el análisis de ADMIXTURE (cross-validation, CV).</i>	122
<i>Figura 52. Análisis de Ancestría Global sobre las poblaciones de la Patagonia genotipificados en MEGA.</i>	123
<i>Figura 53. Análisis de Componentes Principales para las muestras de la Patagonia Austral.</i>	126
<i>Figura 54. Análisis de Componentes Principales para las muestras de la Patagonia Austral.</i>	127
<i>Figura 55. Análisis de Ancestría Local sobre el genoma de un individuo.</i>	128
<i>Figura 56. Test de Correlación para la inferencia de Ancestría Global y Local en las poblaciones de la Patagonia Austral.</i>	129

<i>Figura 57. Test de Correlación para la inferencia de Ancestría Global y Local en las poblaciones de la Patagonia Austral.</i>	130
<i>Figura 58. Análisis de TRACTs y tiempos de mestizaje en las poblaciones de la Patagonia para las ancestrías Europeas, Amerindias y Africanas.</i>	132
<i>Figura 59. Análisis de TRACTs y tiempos de mestizaje en las poblaciones de la Patagonia para las ancestrías Aymara, Centro Sur de Chile y Europea.</i>	133
<i>Figura 60. Manhattan Plot para los valores de PBS sobre los cazadores recolectores marítimos de la Patagonia Austral (IBS; Pehuenche/Huilliches; Cazadores Recolectores Marítimos).</i>	135
<i>Figura 61. Proporción de SNPs con valores de PBS sobre el 1% en ventanas de 200kb sobre las poblaciones de la Patagonia Austral (IBS; Pehuenche/Huilliches; Cazadores Recolectores Marítimos).</i>	136
<i>Figura 62. Número de individuos y SNPs para cada rango de componente Amerindio en grupos cazadores recolectores marítimos de la Patagonia.</i>	138
<i>Figura 63. Frecuencia de SNPs por ventana de 200 Kb para diferentes rangos de Ancestría Amerindia.</i>	139
<i>Figura 64. Valores de iHS por SNPs en poblaciones cazadores recolectores marítimas de la Patagonia.</i>	140
<i>Figura 65. Proporción de SNPs con valores de $iHS > 2.5$ dentro de cada ventana de 200Kb en poblaciones cazadoras recolectoras de la Patagonia.</i>	142
<i>Figura 66. Detección de loci bajo selección a lo largo del Genoma basado en ancestrías para las poblaciones Kawéskar y Yámana de la Patagonia Austral.</i>	144
<i>Figura 67. Esquema con la representación de rutas biológicas enriquecidas para los análisis de PBS sobre los SNPs en los individuos de la Patagonia Austral.</i>	146
<i>Figura 68. Esquema con la representación de rutas biológicas enriquecidas para la detección de Loci bajo selección mediante ancestría en los individuos de la Patagonia Austral.</i>	150
<i>Figura 69. Señal de Adaptación para la región del Cromosoma 11 (123,6-123,8Mb) para las comparaciones de Axiom y MEGA.</i>	153
<i>Figura 70. Comparación entre los genes asociados previamente a frío y dieta con las ventanas sobre el 1% de los valores de PBS. En azul se encuentran los genes asociados a frío y dieta descritos previamente.</i>	154
<i>Figura 71. Distribución de los valores de PNS por SNPs entre las regiones 61-70Mb en el cromosoma 11.</i>	155
<i>Figura 72. Comparación de valores de PBS por SNPs entre los cazadores recolectores marítimos de la Patagonia en regiones cromosómicas de genes descritos previamente en literatura para otras poblaciones nativas.</i>	156
<i>Figura 73. Análisis de iHS para SNPs en la Patagonia Austral.</i>	157

Figura 74. Ruta metabólica descrita para el señalamiento de la apelina (hsa04371) extraído a partir del Mapa de rutas KEGG (www.kegg.jp). _____ 159

Figura 75. Esquema con la historia de los procesos evolutivos por parte de las poblaciones de la Patagonia. _____ 181

1. Resumen

Antecedentes. La Patagonia, situada al sur de Suramérica corresponde a uno de los últimos lugares en ser poblados por la especie humana hace aproximadamente 13,000 años. Desde entonces, esta zona ha experimentado un proceso de deglaciación con variaciones en las condiciones climáticas y en la distribución de la biota, siendo actualmente la región mas fría de Suramérica. Las primeras poblaciones humanas colonizadoras de la Patagonia enfrentaron estos paisajes y desarrollaron modos de vida acorde a los climas fríos. Los registros de sitios tempranos indican que los primeros pobladores fueron cazadores-recolectores terrestres que ocuparon principalmente la Patagonia oriental. El proceso de adaptación a la vida marina en la Patagonia se inicia a partir de los 6,000 cal años AP desde grupos de cazadores-recolectores terrestres en la costa de la Patagonia occidental. Esta situación consideró una serie de respuestas culturales asociadas a la caza y recolección de productos marinos y a un modo de vida estrechamente ligado con condiciones climáticas adversas. Por otro lado, la llegada de los españoles generó una drástica reducción de las poblaciones Nativas, así como también la introducción de patógenos. **Objetivo.** En este marco surge nuestro interés de determinar si las adaptaciones culturales que generaron estas poblaciones fueron suficientes para enfrentar las condiciones climáticas adversas y el modo de vida marino, o experimentaron además procesos de adaptación biológica que les permitieron optimizar los recursos y enfrentar de mejor manera el clima de la Patagonia. De esta manera, hipotetizamos que los primeros pobladores de la Patagonia presentaron adaptaciones biológicas asociadas a temperatura y dieta marina y resistencia a enfermedades. **Metodología.** Generamos datos de genotipos con cobertura genómica en descendientes de cazadores recolectores de la Patagonia mediante microarreglos de ADN. Evaluamos la presencia de regiones genómicas con evidencia de selección positiva que pueda haber ocurrido previo al contacto con europeos mediante el análisis de diferenciación poblacional y extensión de haplotipos. Por

otro lado, investigamos el genoma por regiones bajo selección post-contacto europeo mediante el enriquecimiento local de ancestría. **Resultados.** Los individuos de la Patagonia Chilena y Argentina muestran un componente Nativo Americano común entre sí que no fue capturado por ninguna nacestría amerindia conocida sino por una aparente mezcla entre los componentes Mapuche y Aymara. La distribución de los largos de fragmentos de cada ancestría permitió datar el proceso de mestizaje en la Patagonia argentina hace 11 generaciones y en los grupos más australes de la Patagonia chilena, hace 9 generaciones. Además, se encontró evidencia de señales de selección pre-contacto alrededor de genes asociados con metabolismo de lípidos (**ASXL2, LRP2, ADAM17**), mecanismos de respuesta a frío (**TRPM8**) y a regulación del sistema cardiovascular en ambientes de hipoxia (**UCN3**). Además, encontramos regiones bajo selección en genes previamente descritos en grupos Nativos de Siberia (**FADS, CPT1A, THADA**, etc.). La detección de loci bajo selección mediante enriquecimiento ancestral de una ancestría identificó un significativo aumento de la ancestría amerindia en la región HLA del cromosoma 6 y en una región proximal del cromosoma 21, ambos abarcando genes asociados a respuesta inmune y metabolismo de lípidos, entre otros. **Discusión y Conclusiones.** El análisis de ancestría reveló un componente amerindio común entre los grupos patagónicos que no pudo ser completamente caracterizado por ninguno de los paneles de referencia usados, sugiriendo la presencia de una historia ancestral particular de estos grupos. El mestizaje en Patagonia ocurrió hace 9 a 11 generaciones, al menos tres generaciones más tardío que en centro de Chile. Se encontraron señales de selección en común para los genes **ASXL2, LRP2** y **ZNF202** en todas las poblaciones estudiadas de la Patagonia, los cuales están asociados con el metabolismo de lípidos. También, encontramos señales de selección en el genoma de descendientes de cazadores recolectores marítimos de la Patagonia los cuales involucran a los genes **ZNF202, GRAMD1B** y **UCN3** que participarían en los procesos de adaptación de dieta terrestre a marina y en mecanismos de buceo. El gen **TRPM8**, ya había sido asociado con la regulación de la temperatura, y descrito como el principal

responsable de la respuesta adaptativa a ambientes fríos en varias poblaciones de Europa. El enriquecimiento de ancestría amerindia en los cromosoma 6 y 21 sugieren una selección posmestizaje de haplotipos con adaptaciones a patógenos y otras condiciones locales. Los resultados de este trabajo permitieron encontrar múltiples señales de selección positiva en poblaciones mestizas de la Patagonia que ocurrieron tanto antes del encuentro con europeos, posiblemente en el holoceno temprano, como posteriores al mestizaje ocurrido.

2. Abstract

Background. Patagonia is located south of South America and corresponds to one of the last places to be populated by the human species, approximately 13,000 years ago. Since then, this area has undergone a deglaciation process with variations in climatic conditions and biota distribution, currently being the coldest region in South America. The first colonizing human populations of Patagonia faced these landscapes and developed ways of life according to cold climates. Early site records indicate that the first settlers were land hunter-gatherers who mainly occupied eastern Patagonia. The process of adaptation to marine life in Patagonia starts from 6,000 cal years AP from groups of land hunter-gatherers on the coast of western Patagonia. This situation considered a series of cultural responses associated with hunting and collecting marine products and a way of life closely linked to adverse weather conditions. On the other hand, the arrival of the Spaniards generated a drastic reduction of Native populations, as well as the introduction of pathogens.

Objective. Within this framework, our interest arises to determine if the cultural adaptations generated by these populations were sufficient to face adverse climatic conditions and marine life, or also experienced biological adaptation processes that allowed them to optimize resources and to better face Patagonia's climate. In this way, we hypothesized that the first inhabitants of Patagonia presented biological adaptations associated with temperature, marine diet and resistance to diseases.

Methodology. We generated genotype data with genomic coverage in descendants of hunter-gatherers from Patagonia using DNA microarrays. We screened the genome for evidence of positive selection that may have occurred prior to contact with Europeans by analyses of genetic divergence and of extended haplotypes. On the other hand, we investigated the genome for selection signatures occurred after admixture by testing for local enrichment of one ancestry.

Results. The individuals of Chilean and Argentinean Patagonia showed a common Native American component that was not represented by known amerindian ancestry but by an apparent mixture of the Mapuche and Aymara components. The distribution of the

lengths of fragments of each ancestry allowed us dating the process of miscegenation in the Argentine Patagonia at 11 generations ago and in the southernmost groups of the Chilean Patagonia at 9 generations ago. In addition, we found evidence of pre-contact selection that included genes associated with lipid metabolism (ASXL2, LRP2, ADAM17), cold response mechanisms (TRPM8) and to regulation of the cardiovascular system in hypoxia environments (UCN3). In addition, we found regions under selection in genes previously described in Siberian Native groups (FADS, CPT1A, THADA, etc.). The detection of loci under selection by ancestral enrichment of an ancestry identified a significant increase in Native American ancestry in the HLA region of chromosome 6 and in a proximal region of chromosome 21, both encompassing genes associated with immune response and lipid metabolism, among others. **Discussion and Conclusions.** The analysis of ancestry revealed a common Native American component among the Patagonian groups that could not be fully characterized by any of the reference panels used, suggesting the presence of a particular ancestral history of these groups. Miscegenation in Patagonia occurred 9 to 11 generations ago, at least three generations later than in central Chile. Common selection signals were found for the ASXL2, LRP2 and ZNF202 genes in all the populations studied in Patagonia, which are associated with lipid metabolism. Also, we found signs of selection in the genome of descendants of marine gatherer hunters from Patagonia, which involve the genes ZNF202, GRAMD1B and UCN3 that would participate in the processes of adaptation of land-to-marine diet and diving mechanisms. The TRPM8 gene had already been associated with temperature regulation, and described as the main responsible for the adaptive response to cold environments in several populations in Europe. The enrichment of Amerindian ancestry on chromosomes 6 and 21 suggests a post-admixture selection of haplotypes with adaptations to pathogens and other local conditions. The results of this work allowed to find multiple signs of positive selection in admixed populations in Patagonia that occurred both before the

encounter with Europeans, possibly in the early Holocene, and after the miscegenation occurred.

3. Introducción

3.1. Antecedentes Generales

La Patagonia es una región extensa y diversa en el sur de Suramérica que se extiende desde aproximadamente 40°S hasta el extremo sur del continente (55°S), incluyendo Tierra del Fuego y la sección meridional de la Cordillera de los Andes. El oeste de la Patagonia es una porción de tierra angosta que se extiende desde la intrincada costa del Pacífico hasta la crestas de los Andes que se elevan por sobre los 1,500 metros sobre el nivel del mar en estas latitudes (Garreaud, Lopez, Minvielle, & Rojas, 2013), en cambio el este de la Patagonia se caracteriza por ser un semi-desierto con hierbas y arbustos resistentes a la sequedad (Scheinsohn, 2003). Los modos de vida de las poblaciones nativas de la Patagonia están caracterizadas por el desarrollo cultural de dos estilos: los cazadores recolectores terrestres situados de preferencia en las áreas esteparias al oriente de la cordillera de los Andes y los cazadores recolectores marinos o canoeros, habitantes del medio boscoso y húmedo de los archipiélagos occidentales. Cada modo de vida implica un proceso cultural distinto asociado a las diferentes condiciones climáticas de la Patagonia (Massone & Gusinde, 1987). En términos de adaptación climática, los grupos de la Patagonia ocuparon una posición en el mundo que es comparable con aquellas poblaciones adaptadas al frío en el hemisferio norte (Pearson & Millones, 2005). Las poblaciones canoeras marinas poseían una baja estatura, eran robustos y de extremidades cortas (Barros Valenzuela, 1975a, 1975b, 1975c). Las medidas antropométricas sugieren que las características físicas de estas poblaciones son semejantes a las de las poblaciones del Ártico adaptadas al frío en el hemisferio norte (Pearson & Millones, 2005). Para el caso de los cazadores recolectores marinos hay otra exigencia, y es que un ambiente frío, con viento y lluvioso impone una demanda metabólica alta. La arqueología indica que su dieta era rica en grasas provenientes en gran medida de los mamíferos marinos que cazaban (Orquera, Legoupil, & Piana, 2011). En este estudio se investigó mediante una aproximación genómica, la ocurrencia de eventos

de adaptación biológica frente a las condiciones de frío o al tipo de dieta de cazadores recolectores marinos en poblaciones originarias de la Patagonia.

3.2. Poblamiento de América

3.2.1. Poblamiento del Continente Americano

El poblamiento de América se llevó a cabo vía Beringia y por una única población proveniente de Siberia (Goebel, Waters, & O'Rourke, 2008). El tiempo de entrada de estos cazadores recolectores se estima entre 15,000 y 19,000 años atrás, a finales del último máximo glacial y habrían seguido una ruta costera por el oeste de Norteamérica (Goebel et al., 2008; Gravel et al., 2013; Perego et al., 2009; Tamm et al., 2007). Suramérica fue colonizado por grupos cazadores-recolectores entre los 15,000 y los 13,500 años atrás por el istmo de Panamá (Rothhammer & Dillehay, 2009). Una vez en Suramérica, estas poblaciones pudieron haber tomado diferentes rutas migratorias: (1) por el altiplano Andino mediante los valles del río Cauca y Magdalena en Colombia, (2) algunos grupos migraron hacia el este siguiendo el borde caribeño (Venezuela, las Guayanas, y el noreste de Brasil) y hacia el interior de Venezuela, (3) otros grupos también pudieron haber migrado a lo largo de la costa del Pacífico hasta Chile, aprovechando los recursos marinos.

3.2.2. Poblamiento de la Patagonia

El poblamiento temprano de la Patagonia se sostiene en base a dos rutas: una costera por el pacífico y una oriental. Distintas evidencias interdisciplinarias (arqueológicas, moleculares) señalan que el poblamiento de la Patagonia se llevó a cabo por la ruta oriental a través de las Pampas Argentinas (Rothhammer & Dillehay, 2009), mientras que de Saint Pierre et al., (2012) revela a través de análisis de ADN mitocondrial que las poblaciones nativas que habitan al sur de los 40°S en Chile y Argentina comparten un origen en común y pertenecen al mismo substrato

poblacional indicando una colonización trans-cordillerana por grupos de la ruta costera del Pacífico.

3.2.3. Registros Arqueológicos del Poblamiento de la Patagonia

El sitio arqueológico Monte Verde II, ubicado en el noroeste de la Patagonia (40-43°S), corresponde al sitio más antiguo de ocupación temprana registrado en Chile y Suramérica. El datado mediante radiocarbono de restos orgánicos en Monte Verde II señalan intervalos de 13,980-15,000 años cal AP (Dillehay, 1997; Dillehay et al., 2008). A partir de los 12,500 años cal AP es posible encontrar una mayor cantidad de sitios arqueológicos mayoritariamente en la Patagonia Argentina, los cuales se caracterizan por estar asociados a grupos de cazadores recolectores terrestres (González-José, 2003; Dominique Legoupil, 2005; Rivas, Ocampo, & Aspillaga, 1999). Los cazadores terrestres ocuparon la estepa patagónica meridional y las áreas norte y sur de la isla de Tierra del Fuego. Estos grupos nómades se sucedieron durante cerca de 11,500 años hasta constituir dos etnias independientes: la etnia tehuelche meridional, los “aonikenk”, asentada en la pampa continental, actual Patagonia Argentina; y los onas o “selknam”, que poblaron las estepas septentrionales y los bosques meridionales de Tierra del Fuego (Massone & Gusinde, 1987).

De acuerdo a diversos estudios, los primeros indicios de ocupación humana marina en el oeste de la Patagonia se concentran en la ribera norte del Estrecho de Magallanes (Punta Santa Ana) y en el sur del canal de Beagle (Imiwaia I y Lancha Packewaia), límite entre pampa y archipiélagos y un territorio de transición donde la caza de especies terrestres y marinas es igualmente posible (D Legoupil & Fontugne, 1997). Los fechados muestran que el poblamiento de este territorio comenzó aproximadamente hace 6,400 años AP (7,300 cal AP) hacia el sector más austral del continente (Canal de Beagle) (Orquera et al., 2011; Orquera & Piana, 2009). Sin embargo, se han abierto otras hipótesis de inicios de vida marina en Puente Quilo 1 en el noreste de la Isla de Chiloé, Golfo de Quetalmahue, y en las islas

Guaitecas hace 6,500 años cal AP (Rivas et al., 1999) y bahía Chamiza en la costa occidental del Seno de Reloncaví al sur de Puerto Montt hace 6,500 años cal AP (Gaete et al., 2004; Ocampo & Rivas, 2004). Algunos restos óseos humanos muestran un excesivo desgaste dentario, la presencia de osteoma en el conducto auditivo del cráneo y el desarrollo de algunas inserciones y lesiones en los huesos largos los que sugieren que estos individuos habrían desarrollado estrategias de subsistencia relacionadas con la explotación de recursos marinos, el buceo y el uso de embarcaciones (Rivas et al., 1999) (Figura 1).

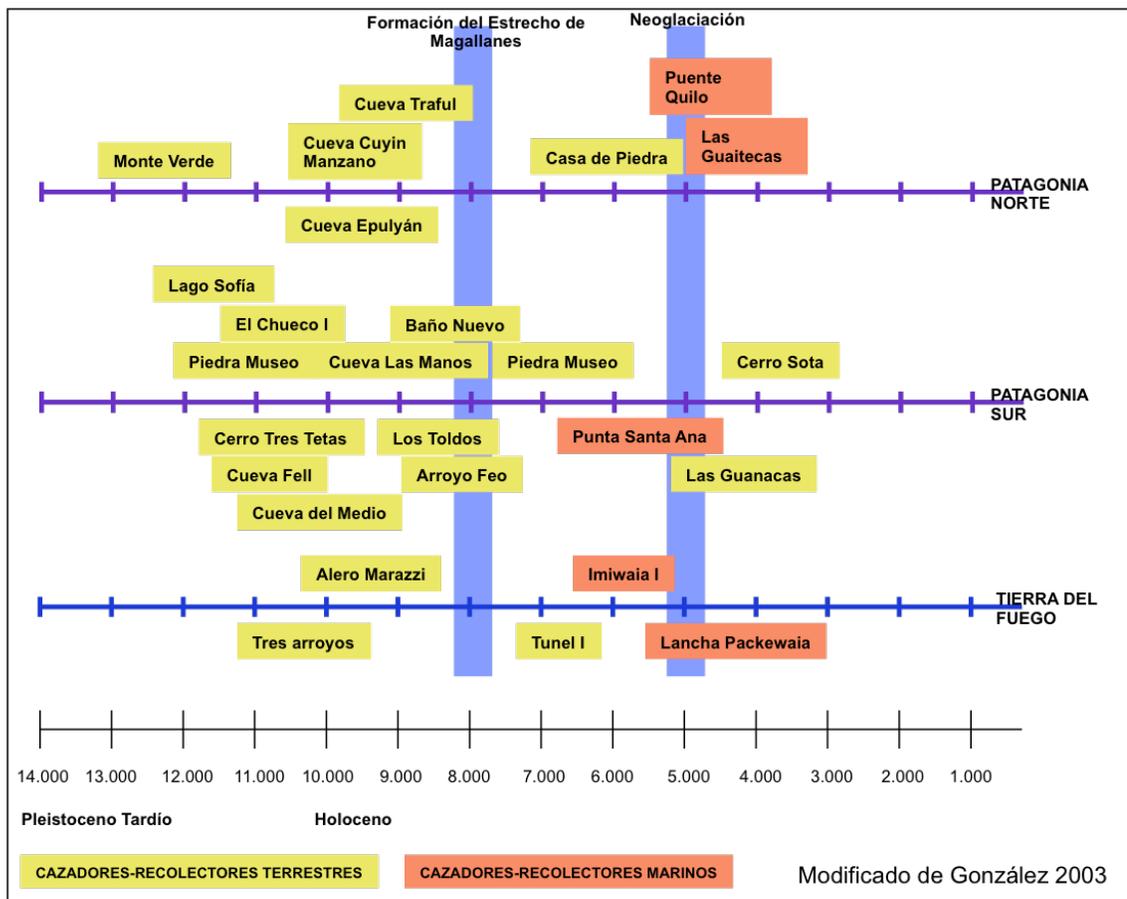


Figura 1. Esquema temporal con los sitios arqueológicos en la Patagonia (modificado de González 2003).

3.3. Grupos Cazadores Recolectores Marinos en la Patagonia

Los grupos canoeros nómades eran cazadores marinos, pescadores y mariscadores, cuya antigüedad se remonta por lo menos a 6,000 años. Estas poblaciones se desarrollaron culturalmente para formar en los últimos siglos a la etnia Kawéskar, habitantes de los archipiélagos entre el Golfo de Penas y la península de Brecknock, y la etnia Yámana, situada desde las costas del canal de Beagle hasta el Cabo de Hornos (Massone & Gusinde, 1987). Sobre este periodo de 6000 años, los patrones de asentamiento y los set de herramientas fueron simples. De hecho, las viviendas, vestimenta, y organización social descritas por los europeos eran escasamente elaboradas (Orquera & Piana, 2009). En el área litoral de la costa oeste de la Patagonia muchos factores estimularon el desarrollo de un modo de vida litoral: el terreno montañoso adyacente a la costa, las dificultades para la movilidad humana, la abundancia de alimentos con altos valores energéticos a lo largo de la costa, y la escasez de recursos de plantas y animales en los bosques (Orquera & Piana, 2009). A lo largo de la costa era posible encontrar pinnípedos, aves marinas, peces, y mustélidos a corta distancia, eventualmente cetáceos varados, así como madera y materiales líticos para la construcción de herramientas (Orquera et al., 2011). El análisis a través de los isótopos estables de ^{13}C y ^{15}N sobre restos óseos (colágeno) permite inferir el tipo de dieta de las poblaciones cazadoras recolectoras. Los estudios realizados en muestras humanas pertenecientes a grupos canoeros del Archipiélago de Chonos y de los canales de la Tierra del Fuego revelan valores característicos de dietas ricas en recursos marinos, evidenciado por los altos valores de $\delta^{15}\text{N}$ propios de las extensas cadenas tróficas marinas (Báez, Moraga, Méndez, & Cherkinsky, 2015; Yesner, Figuerero Torres, Guichon, & Borrero, 2003). En cambio, la dieta de las poblaciones cazadoras recolectoras terrestres estaba basada principalmente en la caza de guanacos (*Lama guanicoe*), complementada con el consumo de cánidos, aves, roedores y algunos frutos y vegetales (Berón, Luna, & Barberena, 2009; Borrero, Barberena, Franco, Charlin, & Tykot, 2009; Méndez,

Barberena, Reyes, & Nuevo Delaunay, 2014; Panarello, Zangrando, Tessone, Kozameh, & Testa, 2006; Tessone et al., 2009), (Massone & Gusinde, 1987).

3.4. Paleoclimatología

Posterior al Último Máximo Glacial (UMG), climas fríos y secos predominaban en el este de los Andes y en Tierra del Fuego (Quattrocchio, Volkheimer, Borrromei, & Martínez, 2011). Hace 14,800 años cal AP una fase de enfriamiento empezó a nivel global en ambos hemisferios, la que duró hasta principios del Holoceno y presentó una atenuación en sus condiciones hace 12,900-11,700 años cal AP (Sagredo et al., 2011). Durante el Holoceno Temprano (11,500-8,000 años), el sur de la Patagonia andina y Tierra del Fuego presentaron condiciones más cálidas y secas que en la actualidad, con un reemplazo de gramíneas por bosques (producto de la retirada de los hielos) y una elevada incidencia de incendios (Kilian & Lamy, 2012; Quattrocchio et al., 2011). En el Holoceno medio (8,000-4,000) se registran en la Patagonia condiciones más húmedas y frías que durante el Holoceno temprano. La Transgresión Marina del Holoceno (formación del estrecho de Magallanes) fue el evento más importante, seguido por tres períodos de neoglaciaciones, siendo el más reciente hace 5,000 años cal AP (Quattrocchio et al. 2011). Posteriormente, en la región de Magallanes y Tierra del Fuego las variaciones de la temperatura superficial del aire (SAT) se han mantenido similares desde hace 5,000 años cal AP (Kilian & Lamy, 2012). Para el noroeste de la Patagonia se registró un aumento en las precipitaciones y subsecuentes eventos de enfriamiento después de los 7,000 años cal AP que alcanzaron su máximo hace 5,000 años cal AP (Moreno, 2004; Moreno & León, 2003; Villa-Martínez & Moreno, 2007).

3.5. Señales de Adaptación

Dentro de los últimos 100,000 años, el hombre anatómicamente moderno se ha esparcido desde África sub-sahariana para colonizar la mayoría de las masas de

tierra alrededor del mundo. Las poblaciones humanas viven en una extraordinaria variedad de diferentes hábitats: frío y cálido; húmedo y seco; bosques, praderas y tundras (Pritchard, Pickrell, & Coop, 2010). La travesía del *Homo sapiens* durante el proceso de colonización por el mundo estuvo envuelto estrechamente por las distintas condiciones climáticas que tuvieron que enfrentar mediante innovaciones tecnológicas, incluyendo el fuego, vestimentas, vivienda, avances en la caza, y métodos para el almacenamiento de agua y comida. Sin embargo, también pudieron haber experimentado adaptaciones biológicas, donde las poblaciones humanas sufrieron cambios fisiológicos y genéticos que les permitieron sobrevivir a estas condiciones (M. Ilardo & Nielsen, 2018). En la Figura 2 podemos apreciar las distintas señales de selección que se han descrito alrededor del mundo. Un estudio que identificó los SNPs con las correlaciones más fuertes entre las frecuencias alélicas y nueve variables climáticas sobre un total de 61 poblaciones a nivel global estableció que las señales más extremas de adaptación provienen de SNPs asociados con pigmentación y enfermedades autoinmunes y con rutas metabólicas relacionados con la radiación UV, infección e inmunidad, y cáncer (A. M. Hancock et al., 2011). Sin lugar a duda, la variación de la incidencia de calor a lo largo del mundo desencadena una serie de efectos climáticos muy diversos, por lo tanto representa una de las restricciones más importantes por las cuales el ser humano se ha tenido que enfrentar. En términos generales, el exceso de calor favorece mecanismos de reducción a la producción de calor, mientras que en condiciones frías se favorece la producción de calor a través de diferentes rutas metabólicas (Leonard, Snodgrass, & Sorensen, 2005a). El ambiente ártico es quizás uno de los ambientes más inhóspitos sobre la tierra. Sin embargo, diferentes culturas se han adaptado a este ambiente, incluyendo los Siberianos como los Chukchi y los Evenks, Europeos como los Sami, y Norte Americanos Nativos como los Inuit (M. Ilardo & Nielsen, 2018). En Siberia específicamente, las poblaciones que llegaron tuvieron que soportar las temperaturas más frías del mundo (con una temperatura mínima registrada de -71,2°C en invierno) (Cardona et al., 2014a).



Figura 2. Adaptación a diversos ambientes durante el curso de la migración humana a lo largo del mundo. Extraído a partir de Tishkoff, 2015.

3.5.1. Señales de Adaptación en América

En América, el sector de los Andes ha conferido diferentes presiones selectivas para las poblaciones nativo americanas que se establecieron hace miles de años atrás. Uno de estos factores es la altitud, donde varios genes relacionados a la homeostasis del oxígeno presentan evidencia preliminar de estar expuestos a selección positiva frente a la baja concentración de oxígeno (Hipoxia) (A. W. Bigham et al., 2009). También se han encontrado tres variantes presentes en los genes USP7, LIF y MDM2 que fueron seleccionados para ambientes de alta altitud (>2500 msnm). Estos genes participan en la ruta del gen TP53, el cual codifica para un factor de transcripción (p53) que participa en varios procesos metabólicos, dentro de los cuales se encuentran la hipoxia (Vanessa C Jacovas et al., 2018; Vanessa Cristina

Jacovas, Rovaris, Pérez, & Azevedo, 2015). Otro factor un poco más localizado ha sido la exposición constante a elevados niveles de Arsénico en el agua de consumo por parte de diferentes poblaciones alto andinas durante miles de años (Schlebusch et al., 2013). En base a estos, se han detectado señales de selección positiva sobre el gen AS3MT, asociado con la metabolización del arsénico en poblaciones andinas expuestas a fuentes de agua con elevados niveles de Arsénico (Eichstaedt, Antao, et al., 2015; Schlebusch et al., 2015, 2013). En algunas poblaciones mestizas de México se ha detectado una variante en el gen transportador de colesterol (ABCA1) que es exclusivamente propia de los grupos nativos americanos y que favorece el colesterol intracelular y el almacenamiento de energía. Esto permite compensar las fluctuaciones de la distribución energética en caso de hambruna. Sin embargo, en las poblaciones mestizas actuales esta condición genera severos problemas asociados con problemas al colesterol, obesidad y diabetes tipo 2 (Acuna-Alonzo et al., 2010).

3.5.2. Genes bajo Selección en Respuesta a Frío y Dieta

Como se ha mencionado anteriormente, las poblaciones de Siberia se encuentran en una de las regiones más extremas del mundo, teniendo que adaptarse ya sea culturalmente, o mediante la selección de genotipos que permitieron conferir una mayor adecuación biológica que les permita soportar las condiciones de frío extremo y baja disponibilidad de recursos que allí ocurre (Cardona et al., 2014a; a. M. Hancock, Alkorta-Aranburu, Witonsky, & Di Rienzo, 2010; M. Ilardo & Nielsen, 2018; Leonard, Snodgrass, & Sorensen, 2005b). En la Tabla I se muestra una lista con los genes bajo selección que se han encontrado en estas poblaciones. Se han detectado varios genes candidatos a respuestas al frío, de los cuales CPT1A y LRP5 muestran una mejor señal de adaptación en grupos del noreste de Siberia relacionados con la regulación energética y el metabolismo. También aparecen otros genes como THADA el cual está involucrado con la diabetes tipo 2 y regulación metabólica y PRKG1 asociado con la contracción muscular, el cual incluye vasodilatación y vasoconstricción. Además se ha encontrado el gen POLD3, el cual

tiene elevadas señales en poblaciones del sur de Siberia y está envuelto en la respuesta al daño del ADN producido por la radiación UV (Cardona et al., 2014a; Clemente, Cardona, Inchley, Peter, Jacobs, Pagani, et al., 2014). Respecto a la dieta, podemos afirmar que también se han encontrado señales de adaptación asociadas a genes FAD1, FAD2, FAD3 y CPT1A en poblaciones Inuit. Estos genes participan en la metabolización de ácidos grasos poli-insaturados (Omega 3), los cuales se encuentran en una elevada concentración en una dieta de tipo marítima, dado el alto consumo de recursos marinos (mariscos, pescados, pinnípedos). Cabe mencionar que estas poblaciones Siberianas guardan importante relación con el poblamiento americano, puesto que a partir de estos salieron los grupos que cruzaron Beringia y colonizaron América (Achilli et al., 2013; Dulik et al., 2012), pudiendo significar que los mecanismos adaptativos a climas fríos de las poblaciones de Siberia tuvieron una gran influencia en la capacidad de respuesta de las poblaciones nativas americanas en el poblamiento de América (Cardona et al., 2014a). Estos resultados pueden ser interesantes bajo el marco del poblamiento de la Patagonia, ya que el temprano avance y ocupación del continente Americano desde el Estrecho de Bering pudo haber evitado que las señales de selección a frío y dieta se disolvieran, y por lo tanto, hayan servido como respuestas adaptativas de las poblaciones de la Patagonia frente a las condiciones ambientales que se presentaron en esta.

No solo se han encontrado señales de selección en respuesta a frío y dieta que podrían estar jugando un rol importante en poblaciones de la Patagonia, sino que también se han descrito una serie de genes involucrados en el buceo en poblaciones nativas de Indonesia. A pesar de que estos grupos divergen hace mucho más tiempo que con las poblaciones de la Patagonia, estos compartían un modo de vida asociado a la caza y recolección de productos marítimos, lo cual involucraba periodos constantes de buceo. El gen receptor B2 Bradikinina (BDKRB2), un péptido asociado con la vasodilatación y vasoconstricción, fue sugerido como principal causante de la vasoconstricción periférica inducida por respuesta a buceo (Mitrofanova et al., 2017). Una señal de selección fue encontrada cerca de este gen en la población nativa

de Bajau, un grupo de “Nómades marítimos” al Sureste de Asia, por lo cual podría estar ayudando a la oxigenación preferencial de importantes órganos como el cerebro, corazón y pulmones, incrementando potencialmente el tiempo de buceo (M. A. Ilardo et al., 2018). En este mismo estudio menciona genes candidatos asociados a buceo como FAM178B, que codifica para una proteína formando un complejo estable con la anhidrasa carbonica, responsable de mantener el balance dióxido de carbono/bicarbonato ayudando a mantener el pH de la sangre; CACNA1A, el cual esta envuelto en la regulación de la liberación de glutamato y la respuesta a condiciones hipóxicas; y PDE10A, un gen envuelto en la regulación de la contracción del músculo liso, incluyendo el músculo que rodea el bazo (M. A. Ilardo et al., 2018; M. Ilardo & Nielsen, 2018).

Tabla I. Lista de genes asociados a señales de selección en respuesta a frío, dieta y buceo.

chr	Gene	adaptation	population	source
2	FAM178B	diving	Bahau, Indonesia	Llardo et al., 2018
2	THADA	cold	Siberians	Cardona et al., 2014
6	PDE10A	diving	Bahau, Indonesia	Llardo et al., 2018
9	ABCA1	type 2 diabetes	Mexican Mestizos, Native American	Acuña-Alonso et al., 2010
10	PRKG1	cold	Siberians	Cardona et al., 2014
11	CPT1A	high-fat diet	Siberian	Clemente et al., 2014
11	FADS	high-fat diet	Siberian	Fumagalli et al., 2015
11	CPT1A	cold	North-East Siberian	Cardona et al., 2014
11	LRP5	cold	North-East Siberian	Cardona et al., 2014
14	BDKRB2(upstream)	diving	Bahau, Indonesia	Llardo et al., 2018
19	CACNA1A	diving	Bahau, Indonesia	Llardo et al., 2018

3.6. Señales de Adaptación post-contacto Europeo

Los patrones de selección en las poblaciones humanas no solo están modelados por factores como el clima, sino que se ven fuertemente afectados también por la diversidad local de patógenos en un área determinada (Fumagalli et al., 2011). Esto se debe a que los patógenos y las enfermedades infecto-contagiosas han impuesto

una presión de selección fuerte a lo largo de la historia humana (Fumagalli & Sironi, 2014). Después de la llegada de los españoles, se ha asociado una fuerte disminución de las poblaciones Nativas Americanas relacionada a muchos factores, incluyendo guerras, cambios en la estructura social y una masiva introducción de patógenos provenientes de Europa (K. Patterson & Runge, 2002; R. Thornton, 1997). De acuerdo a esto, muchas poblaciones Nativas han mostrado evidencias de selección positiva frente a la resistencia a los patógenos posterior al contacto Europeo asociado con la región HLA (Lindo et al., 2018; Nakagome et al., 2016). El escenario que envuelve el contacto Europeo en la Patagonia puede haber sido diferente del resto del continente, dado que esta interacción fue esporádica entre barcos españoles y piratas durante el siglo XVI (Gleisner & Sara Montt, 2014). Sin embargo, independiente del tiempo de migración, la susceptibilidad pudo haber jugado un rol importante en presiones selectivas sobre los grupos Nativos de la Patagonia, y posteriormente el inicio del proceso del mestizaje con los grupos Europeos.

3.7. Abordaje Genómico al Estudio de Poblaciones

Hoy en día, existen múltiples herramientas moleculares que nos permiten abordar la genética de las poblaciones a través de una mirada genómica debido al desarrollo de nuevas tecnologías que permiten secuenciar genomas completos, u obtener un set de miles o millones de SNPs a través de microarreglos. Esta gran cantidad de información trae consigo el desafío de implementar nuevas metodologías que permitan interpretar de forma óptima los datos (Schraiber & Akey, 2015). A continuación mostraremos algunas de las aproximaciones que ya sea, se han estado utilizado comúnmente, o se están recién implementando. Abordaremos esto en el contexto del estudio de la estructura de las poblaciones, así como también de la determinación de los eventos de selección que pueden haber experimentado las poblaciones consideradas en nuestro estudio.

3.7.1. Estructuración Poblacional

La inferencia de la historia demográfica de las poblaciones en estudio juega un rol vital para la determinación de eventos de selección, ya que un modelo apropiado inferido podría ofrecer una hipótesis nula bajo neutralidad (Nielsen, Hellmann, Hubisz, Bustamante, & Clark, 2007). De otra manera, los eventos demográficos (e.g. migración, expansiones y cuellos de botella) pueden a menudo imitar a señales de selección (Vitti, Grossman, & Sabeti, 2013a).

3.7.2. Análisis de Ancestría

Las aproximaciones más comúnmente utilizadas corresponden al análisis de componentes principales (PCA), el cual permite maximizar los patrones de similitud genética observada; cuando existen clúster de individuos se pueden interpretar como poblaciones genéticas (N. Patterson, Price, & Reich, 2006). La otra aproximación corresponde a la estimación de ancestrías individuales basada en modelos de maximum likelihood como ADMIXTURE (Alexander & Lange, 2011) y STRUCTURE (Pritchard, Stephens, & Donnelly, 2000), el cual asume que el genoma de un individuo –potencialmente mestizo– puede provenir de ciertas poblaciones ancestrales que confluyeron en el pasado mediante migración y que originaron la población actual.

Estos dos análisis permiten inferir la proporción total para cada ancestría en el genoma de un individuo, también denominada ancestría global. Sin embargo, esto no permite obtener información acerca del origen ancestral sobre cada locus en el genoma de un individuo. Esto último puede resultar altamente relevante si queremos evaluar la diversidad presente dentro de cada componente ancestral sobre genomas mestizos. Una forma de afrontar esto es mediante la inferencia de la ancestría local, la cual permite identificar la ancestría de diferentes segmentos cromosómicos (haplotipos) sobre el genoma de un individuo (Maples, Gravel, Kenny, & Bustamante, 2013a).

3.7.3. Análisis de Poblaciones basados en Ancestría

TRACTs: este método analiza las ancestrías de cada segmento cromosómico (a partir de la ancestría local), y permite inferir la demografía de una población mestiza, mediante la predicción de la distribución de las longitudes de los diversos fragmentos cromosómicos de acuerdo a un modelo demográfico. Estos fragmentos – o “tracts”- corresponden a un trozo del cromosoma con una ancestría específica (calculado mediante ancestría local) y su longitud se ajusta a la dinámica de los procesos estocásticos de recombinación, los cuales tienden a romper los fragmentos en trozos más pequeños. Como resultado de este proceso, los fragmentos tienden a ser más cortos cuando la mezcla es más antigua, permitiendo inferir el tiempo del primer evento de mezcla. La construcción de los modelos responde a la historia demográfica de las poblaciones mestizas, los cuales describen un evento de mezcla inicial entre dos grupos ancestrales, y pulsos posteriores al primer proceso de mezcla (Gravel, 2012).

PCA específico de Ancestría: como se mencionó anteriormente, estimar la variabilidad dentro de cada componente ancestral puede ser de gran importancia dentro de la misma población, puesto que permite descomponer o dar a conocer el grado de heterogeneidad genética, y por lo tanto, identificar distintos procesos evolutivos como procesos de migración. El asPCA (ancestry-specific PCA) permite justamente estimar la ancestría continental específica para cada locus (inferencia de ancestría local), seguido por un PCA asociado únicamente a las porciones del genoma con una ancestría específica. Esto combinado con distintos paneles subcontinentales para cada ancestría que puedan dar cuenta de la variabilidad presente en la población mestiza (Gravel, Zakharia, Moreno-Estrada, et al., 2013; Homburger et al., 2015).

Métodos Basados en Haplotipos: estos análisis permiten inferir la estructura genética de una población en el pasado respecto a otras poblaciones mediante afinidad o cercanía genética de haplotipos. Esto tomando en cuenta la reconstrucción genealógica, los diferentes eventos de recombinación y el desequilibrio de ligamiento entre los distintos marcadores (Lawson, Hellenthal, Myers, & Falush, 2012). ChromoPainter es un programa el cual tiene como función principal la reconstrucción de los cromosomas de un individuo (también llamado “recipiente”) tomando distintos trozos de ADN (“chunks”) donados por otros individuos (“donadores”). Este “Chromosome Painting” se puede resumir en una matriz de co-ancestría la cual revela información acerca de las diferentes relaciones ancestrales entre los individuos (Lawson et al., 2012). Existen dos análisis que toman como input los datos entregados por ChromoPainter; a) fineSTRUCTURE es un programa que permite construir clúster basados en modelos Bayesianos. Cada clúster esta compuesto por un grupo de individuos que poseen similitudes en la proporción de haplotipos que recibieron de cada población donadora, así como también cuanto aportó a cada población. De esta manera se pueden inferir procesos de sub-estructuración genética sin definir a priori el agrupamiento de las poblaciones (Lawson et al., 2012). Por otro lado, b) GLOBETROTTER permite identificar, datar y describir eventos de mestizaje que ocurrieron en la historia ancestral de una población “target” dentro de los últimos 4,500 años. Este análisis establece una población “target” y múltiples grupos “surrogates” los cuales pueden o no estar asociados con las fuentes ancestrales de la población target. GLOBETROTTER utiliza estos grupos para identificar si la población target proviene de algún evento(s) de mestizaje, y si es así, determinar precisamente cuando estos eventos ocurrieron y los grupos que participaron en la mezcla (Hellenthal et al., 2014).

3.7.4. Análisis basados en Simulaciones

Los procesos coalescentes o de difusión pueden ser aplicados para trazar la historia de las especies (Wu, Zheng, Hu, & Wei, 2014a). En este sentido una serie de parámetros que describan la historia podrían ser inferidos mediante máximo de verosimilitud, aproximaciones Bayesianas o Cadenas de Markov Monte Carlo (MCMC) (Crisci, Poh, Bean, Simkin, & Jensen, 2012). Para realizar inferencias acerca de la historia de una población, dos pasos se necesitan. Primero hay que formular la historia de la población con cierto modelo matemático, involucrando un proceso de evolución durante la historia de esta. Segundo, incorporar los métodos de inferencia estadística que podrían estar implicadas sobre los modelos establecidos (Wu, Zheng, Hu, & Wei, 2014b). El modelo de aislamiento con migración es uno de los más utilizados para inferir escenarios demográficos, sobre el cual diferentes métodos pueden ser usados para trazar la historia evolutiva de la variación genética. El modelo asume que dos poblaciones desciendan de una población ancestral panmíctica en algún momento del pasado y donde pueden estar conectadas (o no) por migración (Nielsen & Wakeley, 2001). La genealogía de los alelos puede ser trazada hacia atrás en el tiempo bajo los procesos de coalescencia, durante el cual los parámetros del modelo demográfico pueden ser estimados, incluyendo el cambio en el tamaño efectivo poblacional y el punto en el tiempo de eventos de cuellos de botella y crecimiento exponencial (Wu et al., 2014b).

El método de proceso de difusión es un modelo clásico en la genética de poblaciones puesto que permite predecir la probabilidad de cambios en las frecuencias génicas bajo un proceso markoviano (Kimura, 1955). Los polimorfismos que se analizan también pueden ser organizados dentro de un Espectro de Frecuencias de Sitios (SFS), que es la distribución de las frecuencias alélicas en un set de datos. En el caso de múltiples poblaciones, un SFS conjunto (JSFS) puede ser utilizado, y la evolución de los datos polimórficos entre poblaciones podría ser descrito como el cambio en el tiempo de la distribución alélica en el SFS/JSFS. En el contexto de la teoría neutral, el cambio podría deberse a procesos de difusión (Wu

et al., 2014b). Una implementación denominada **δaδi** se ha llevado a cabo últimamente, el cual basado en la distribución conjunta de las frecuencias alélicas a lo largo de variantes bi-alélicas de múltiples poblaciones, el programa usa un método de likelihood para inferir el SFS esperado bajo un modelo demográfico específico vía un modelo evolutivo simulado por difusión (Gutenkunst, Hernandez, Williamson, & Bustamante, 2009).

3.7.5. Análisis de Selección

La identificación de variantes genéticas que confieren una ventaja al organismo, y que se han esparcido por otras fuerzas evolutivas más que el azar, corresponden a una de las preguntas más importantes en biología evolutiva. A parte de obtener información acerca de los procesos evolutivos que hayan ocurrido, también se tienen implicancias que van desde lo clínico hasta lo ecológico (Bank, Ewing, Ferrer-admettla, Foll, & Jensen, 2014). A continuación, mencionaremos las principales aproximaciones sobre el genoma que permiten obtener señales de selección.

3.7.5.1. Métodos basados en Desequilibrio de Ligamiento

Cuando una mutación ventajosa aparece en la población y subsecuentemente se selecciona, esta tiende a aumentar en frecuencia dentro de la población. Los alelos de los genes circundantes presentes en el cromosoma sobre el cual se encuentra la mutación ventajosa tienden a incrementar en sus frecuencias también (Figura 3). Este proceso se conoce como el “efecto Hitch-hicking”, dado que un alelo puede tener un aumento en su frecuencia a partir de una selección actuando sobre un alelo vecino (Smith & Haigh, 1974). Estos barridos selectivos pueden ser: a) fuertes, donde nuevas mutaciones aparecen en un solo haplotipo que rápidamente se fija en la población, mostrando una variación nucleotídica muy reducida, o b) suaves, donde

la mutación beneficiosa proviene de variación pre-existente, y por lo tanto este presenta una distribución de haplotipos esperados bajo selección neutral, generando múltiples haplotipos de fondo al momento que actúa la selección (Hermisson & Pennings, 2005; Olson-Manning, Wagner, & Mitchell-Olds, 2012). Los alelos provenientes de variación pre-existente tienen más ventajas sobre las nuevas mutaciones. Primero, las poblaciones no tienen que esperar por nuevas mutaciones dado que los alelos ya están presentes. Segundo, los alelos pueden estar en elevadas frecuencias debido a evolución neutral, disminuyendo la probabilidad que se pierda por deriva. Esto les permite responder rápidamente cuando las presiones selectivas cambian debido a cambios ambientales o colonización de nuevos hábitats (A. M. Hancock, Alkorta-aranburu, Witonsky, & Rienzo, 2010; Olson-Manning et al., 2012).

Actualmente existen muchos estadísticos que tienen como objetivo detectar señales de barrido selectivo (Pritchard & Di Rienzo, 2010). Estos métodos se centran alrededor del estadístico del haplotipo homocigoto extendido (EHH), en donde uno define el EHH desde una región core (e.g. un alelo seleccionado putativo) hasta una distancia específica en ambas direcciones y calcula la probabilidad que dos cromosomas cualquiera elegidos al azar dentro de la población que tengan la región core, sean idénticos por descendencia para la región entera (Figura 3). De esta manera, a medida que uno viaja desde la región core, el EHH disminuye, reflejando la acción de la recombinación moldeando el haplotipo dentro de la población (Sabeti et al., 2002). El haplotipo de largo alcance (LRH) compara la frecuencia haplotípica relativo al EHH a varias distancias, buscando por haplotipos que son propagados rápidamente, sugiriendo que la elevada prevalencia fue lo suficientemente rápido como para que la recombinación no haya tenido tiempo para romper el haplotipo (Zhang et al., 2006). El score del haplotipo integrado (iHS) es una variación del EHH. Este estadístico compara el área bajo la curva definida por el EHH para las variantes derivadas y ancestrales a medida que uno viaja en distancias génicas desde la región core. Mediante el cálculo del área bajo la curva definida por el EHH, este test captura valores extremos de EHH sobre distancias cortas y valores moderados de EHH para

distancias mas grandes, indicadores de selección positiva (Voight, Kudaravalli, Wen, & Pritchard, 2006b). Otra variación del EHH corresponde al XP-EHH (cross-population extended haplotype homozygosity), el cual compara longitudes de haplotipos entre poblaciones para controlar la variación local en la tasa de recombinación (Sabeti et al., 2007).

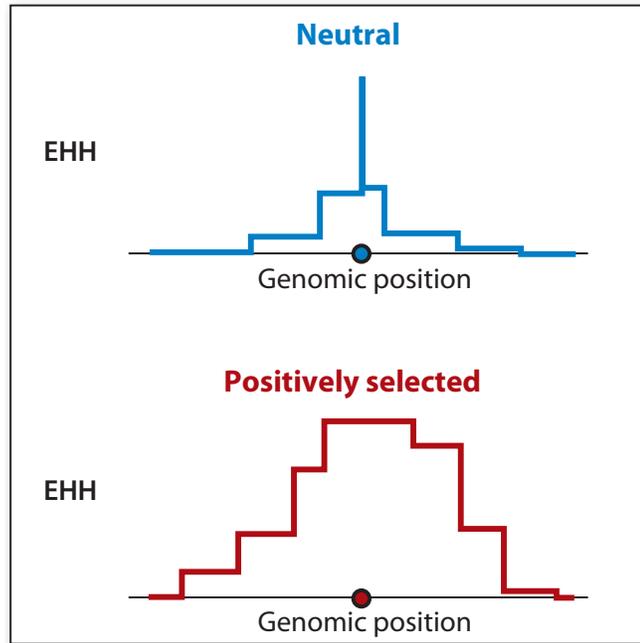


Figura 3. Homocigocidad de haplotipos extendido en el genoma. En azul se aprecia una variante neutral (circulo) y lo esperado en la distribución de haplotipos. En rojo vemos una variante ventajosa y el cambio en la homocigocidad de haplotipos. Extraído de Lachance & Tishkoff, 2013.

3.7.5.2. Métodos basados en Diferenciación Poblacional

Estos análisis se basan en cambios bruscos en las frecuencias alélicas de algunos loci al momento de comparar dos o mas poblaciones. Comparativamente, elevados valores de F_{st} sobre un locus indican grandes diferencias entre poblaciones, lo cual podría sugerir selección direccional, mientras que bajos valores de F_{st} indican que las poblaciones comparadas son homogéneas, que podría ser indicativo de selección balanceadora o direccional en ambas poblaciones (Excoffier, Hofer, & Foll, 2009; Hofer, Ray, Wegmann, & Excoffier, 2009; Vitti, Grossman, &

Sabeti, 2013). A diferencia de las otras aproximaciones, este método permite detectar varios tipos de selección, incluyendo barridos selectivos, barridos sobre variación estándar y selección negativa. Si asumimos que, distintas poblaciones están sujetas a diferentes presiones ambientales, entonces, los rasgos adaptativos en cada población también pueden ser diferentes. Si la selección está actuando sobre un locus dentro de una población pero no dentro de otra población relacionada, entonces las frecuencias alélicas en ese locus entre las poblaciones pueden diferir significativamente (Vitti et al., 2013). El índice más comúnmente utilizado para medir diferenciación poblacional es el índice de fijación (F_{st}), el cual compara la varianza de las frecuencias alélicas dentro y entre poblaciones (Holsinger & Weir, 2009). Varios test de diferenciación poblacional han sido utilizados en base al índice F_{st} , como el test LKT que permite evaluar la heterogeneidad en múltiples loci (Bonhomme et al., 2010; Lewontin & Krakauer, 1973). Sin embargo, cuando se evalúan señales de selección entre dos poblaciones mediante el cambio en frecuencias alélicas, no es posible poder identificar cual es la población que está bajo selección. El análisis de PBS (Population Branch Statistic) no obstante, permite afrontar esta situación a través de la medición del largo de rama para cada población, pudiendo identificar loci con grandes cambios en las frecuencias alélicas para una población en estudio (Yi et al., 2010). Este modelo funciona en base a tres poblaciones relacionadas: la primera corresponde a la población bajo selección; la segunda, es una filogenéticamente más cercana, pero no expuesta a presiones selectivas o al menos a distintas; y por último, la tercera, corresponde a una población ancestral filogenéticamente más lejana que cumple la función de outgroup (Figura 4).

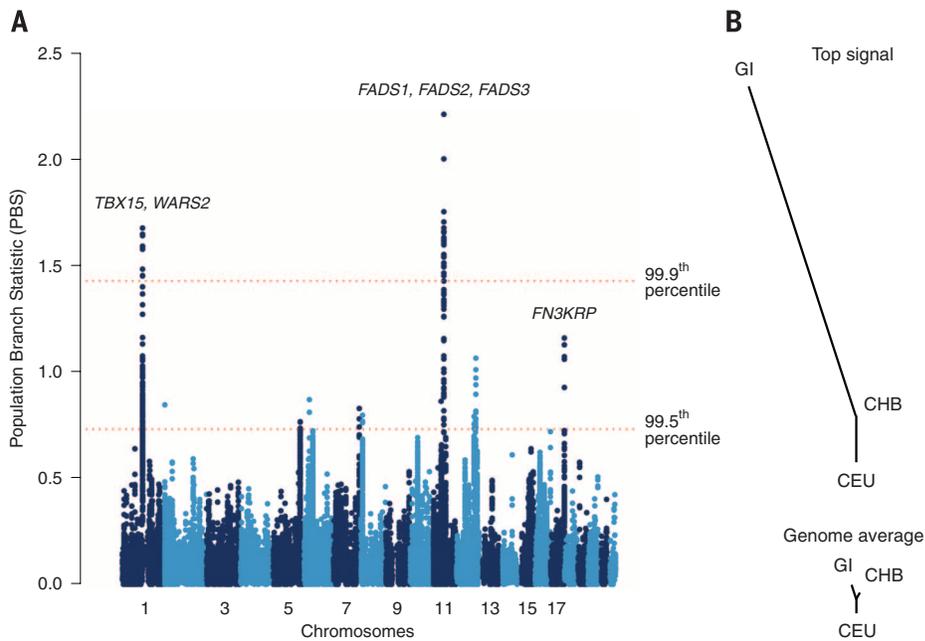


Figura 4. Señales Genómicas de Selección Positiva. (A) valores de PBS por sobre el percentil 99.5 y 99.9 se asocian a regiones bajo selección. (B) Árbol evolutivo que muestra la señal de selección mas fuerte a través del largo de rama. En la parte inferior vemos el largo de ramas promedio del genoma entre estas tres poblaciones. Extraído a partir de Fumagalli et al, 2015.

3.7.5.3. Mapeos Genómicos de Ancestría

Esta aproximación guarda estrecha relación con la inferencia de componentes ancestrales. Es posible rescatar señales de adaptación a partir de las poblaciones mestizas, como las de América Latina, que están conformadas por el mestizaje de poblaciones ancestrales genéticamente distintas como las europeas, africanas y nativas americanas (Zhou, Zhao, & Guan, 2016). Bajo la ausencia de selección, las ancestrías asociadas a cada grupo ancestral en una población mestiza deberían distribuirse de manera relativamente homogénea a lo largo del genoma (Pierron et al., 2018); sin embargo, se espera que aquellos alelos que confieran alguna ventaja para estas poblaciones aumenten sus frecuencias después de la mezcla, causando que las ancestrías locales se desvíen del promedio del genoma (Figura 5). Estas desviaciones pueden ser interpretadas como una señal de selección después del mestizaje (Bhatia et al., 2014), aunque también pueden significar procesos de

adaptación previos al mestizaje y que debido a la acción de agentes selectivos siguen manteniéndose en elevadas frecuencias (Jin et al., 2011).

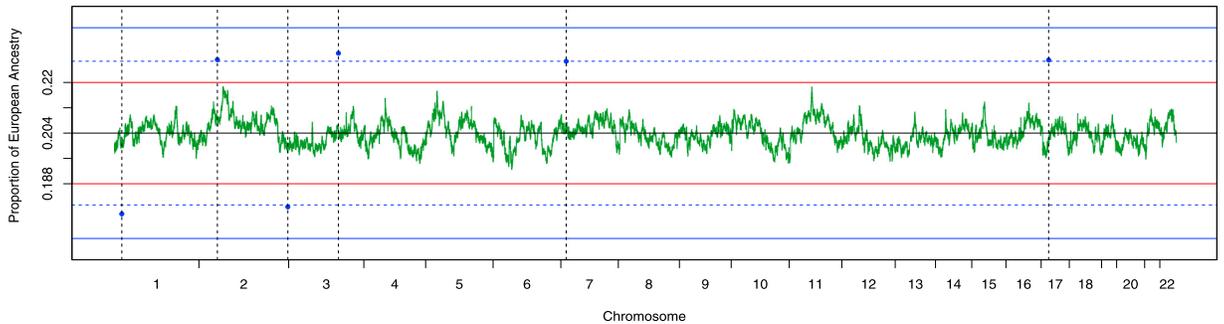


Figura 5. Análisis de Ancestría local para el Genoma. En esta figura se muestra la proporción de la ascendencia Europea para cada SNPs distribuidos en el genoma. La línea negra representa el promedio de a ascendencia. Las líneas rojas y azules representan los puntos de corte para las desviaciones en cada proporción. Extraído a partir de Bhatia et al., 2014.

3.8. Hipótesis de Trabajo

Hipótesis 1: Las poblaciones de cazadores recolectores de la Patagonia experimentaron procesos de adaptación biológica en respuesta al frío durante la transición Pleistoceno-Holoceno similares a las ocurridas previamente en los primeros pobladores del continente Americano durante su paso por Beringia.

Predicciones:

-Poblaciones descendientes de grupos humanos que han habitado la Patagonia por miles de años presentarán señales de selección en regiones de su genoma donde se localizan genes asociados a procesos metabólicos relacionados con el control de la temperatura y la termogénesis.

-Las principales señales de selección en el genoma de estos individuos se encontrarán sobre las mismas regiones que presentaron los primeros colonizadores de América expuestas al frío en latitudes similares en el hemisferio norte.

-Estas regiones comprenderán a los genes CPT1A, LRP5, THADA y PRKG1 asociados a procesos metabólicos relacionados con el control de la temperatura y la termogénesis.

-Estas señales de selección en regiones bajo selección por frío no se encontrarán presentes en poblaciones genéticamente cercanas pero que no han estado expuestas a las condiciones de temperatura de la Patagonia.

Hipótesis 2: Las poblaciones de cazadores recolectores marinos de la Patagonia occidental experimentaron procesos de adaptación biológica en respuesta a dieta marina a partir del Holoceno medio similares a las ocurridas previamente en los primeros pobladores del continente Americano durante su paso por Beringia.

Predicciones:

-Poblaciones descendientes de cazadores recolectores marinos de la Patagonia occidental presentarán señales de elección en regiones de su genoma donde se localizan genes asociados a procesos metabólicos relacionados con dieta marítima.

-Las principales señales de selección en el genoma de estos individuos se encontrarán sobre las mismas regiones que presentaron los primeros colonizadores de América expuestas a condiciones de dieta marina en latitudes similares en el hemisferio norte.

-Estas regiones comprenderán a los genes CPT1A, FADS1, FADS2 y FADS3 asociados a procesos metabólicos relacionados con dieta elevada en ácidos grasos de origen marino.

-Estas señales de selección en regiones bajo selección por frío no se encontrarán presentes en poblaciones genéticamente cercanas pero que no han estado expuestas a las condiciones de temperatura de la Patagonia.

3.9. Objetivos

Objetivos Principales

- Evaluar adaptación biológica a través de señales de selección asociadas a procesos metabólicos en respuesta a condiciones de frío en poblaciones descendientes de grupos cazadores recolectores de la Patagonia.
- Evaluar adaptación biológica a través de señales de selección asociadas a procesos metabólicos en respuesta a la dieta en poblaciones descendientes de grupos cazadores recolectores marinos de la Patagonia occidental.

Objetivos Específicos

1. Determinar la contribución de los diferentes componentes ancestrales en las poblaciones de la Patagonia con ancestría Chono, Tehuelche, Kawéskar y Yámana, así como también estimar su cercanía genética con otros grupos nativos de América mediante una aproximación genómica de ancestrías.
2. Identificar señales de selección en el genoma y determinar su posible función sobre procesos biológicos en poblaciones descendientes de cazadores recolectores con ancestría Chono, Tehuelche, Kawéskar y Yámana.
3. Detectar señales de adaptación local en el genoma presentes únicamente en poblaciones de la Patagonia en comparación con aquellos grupos cercanos genéticamente pero no expuestos a condiciones de frío (Mapuches), y con grupo bajo condiciones ambientales similares de frío (Siberia).
4. Detectar señales de adaptación local en el genoma presentes únicamente en grupos descendientes de cazadores recolectores marítimos de la Patagonia en comparación con poblaciones descendientes de cazadores recolectores terrestres (Tehuelches).

4. Metodología

4.1. Poblaciones Analizadas y Genotipificación de las muestras

Las poblaciones analizadas en este estudio correspondieron a un total de 119 individuos que se reconocieron de pertenencia o ascendencia Yámana (n=20), Kawésqar (n=4), e individuos de la Provincia del Chubut en la Patagonia Argentina con alta ancestría amerindia confirmada por un panel de 46 AIMs (n=20). También, se incluyeron a habitantes del sector de Laitec (n=30) en la Región de los Lagos. Esta localidad se ubica al sur de Chiloé y son posibles descendientes del pueblo originario Chono considerado extinto. Además, se incluyeron 27 individuos de ascendencia Pehuenche de la localidad de Trapa Trapa, Región del BioBio y 18 de ascendencia Huilliche de la localidad de San Juan de la Costa, Región de Los Lagos.

Las muestras de las primeras dos localidades, así como Laitec fueron tomados en la década del 80 y los participantes aceptaron mediante un consentimiento informado su participación en un estudio para determinar su genotipo. El uso de estas muestras en los proyectos fue aprobado por el Comité de Ética de la Facultad de Medicina, de la Universidad de Chile y forman parte del proyecto CONICYT USA2013-0015: “Genomic investigation of the human biodiversity in the pre-Columbian and contemporary Chilean Patagonia.”. Las muestras de Tehuelche fueron recolectadas por el equipo de Antropología Biológica de la Facultad de Filosofía y Letras, Universidad de Buenos Aires en el marco del proyecto “Diversidad genética y mestizaje en poblaciones cosmopolitas de la Patagonia Argentina: Análisis de las particularidades regionales” (PICT 2013-2414, Agencia Nacional de Promoción Científica y Tecnológica) facilitadas para este trabajo por la investigadora María Laura Parolín. La toma de muestras y los consentimientos informados fueron aprobados por el Comité de Ética y Docencia del Hospital sub-zonal de Esquel, Hospital regional de Comodoro Rivadavia y Hospital Zonal Andrés Bello de Puerto Madryn.

Las muestras fueron genotipificadas con el microarreglo Axiom Latino (World Array 4, Affymetrix Inc, Santa Clara, CA) en el Institute for Human Genetics, University of California, San Francisco, California, para un set total de 817.810 SNPs, distribuidos homogéneamente a lo largo del genoma. La genotipificación fue financiada por el proyecto CONICYT USA2013-0015: “Genomic investigation of the human biodiversity in the pre-Columbian and contemporary Chilean Patagonia” dirigido por el Dr. Ricardo Verdugo.

A partir de este set se extrajo un sub-set de 38 individuos y se genotipificaron nuevamente con el microarreglo MEGA (Infinium Multi-Ethnic Global arrays, Illumina) en los Servicios genómicos CINVESTAV-LANGEBIO, Guanajuato, México, para un total de 1.779.819 SNPs distribuidos a lo largo del genoma. Este nuevo microarreglo contiene además a 19 individuos que se reconocen de pertenencia o ascendencia Kaweskar (10 provenientes de Punta Arenas y 9 de Puerto Edén, Región de Magallanes). El caso de Puerto Eden es particularmente especial comparado al resto de las poblaciones en este estudio, puesto este pueblo se fundó en 1969 en el Parque Nacional Bernardo O’Higgins (XII Región). Esta localidad se originó exclusivamente para las comunidades Kawéskar que residían en chozas en las costas cercanas a Puerto Edén. Hasta el 2014 residían 15 Kawéskar, nueve de ellos aún conservan su lengua originaria, muchos de ellos (Gleisner & Sara Montt, 2014) con algún grado de parentesco evidente al momento de tomar la muestra. Estas muestras fueron colectadas por el Dr. Mauricio Moraga como parte del desarrollo del proyecto FONDECYT 1140544 “Señales recientes de adaptación genética a climas extremos en poblaciones de cazadores recolectores marítimos de Patagonia y Tierra del Fuego”. El criterio de selección de estos individuos para ser genotipificados fue tener un elevado componente genético de ancestría nativa. La descripción en detalle del total de los individuos utilizados para este estudio se encuentran en la Tabla II.

Tabla II. Descripción de las poblaciones utilizadas en este estudio.

Localidad	Población Nativa	Sitio Muestreo	Coordenadas	Número Individuos	Plataforma Secuenciación ¹
Trapa-Trapa	Pehuenche	Alto Bío Bío, Región del Bío Bío, Chile	37°43'55"S 71°13'12"O	27	Axiom Latino (13)*
La Misión	Huilliche	San Juan de la Costa, Región de los Lagos, Chile	40°29'31"S 73°24'15"O	18	Axiom Latino (9)*
			73°32'44"O		
Laitec	Chono	Isla Laitec, Chiloé, Región de los Lagos, Chile	43°12'13"S 73°36'28"O	30	Axiom Latino (6)*
Puerto Edén	Kawésqar	Región de Magallanes y la Antártica Chilena, Chile	49°07'34"S 74°24'48"O	9	MEGA
Punta Arenas	Kawésqar	Región de Magallanes y la Antártica Chilena, Chile	53°09'45"S 70°54'29"O	4	Axiom Latino (1)*
				10	MEGA
Ukika	Yámana	Pto. Williams, Isla Navarino, Región de Magallanes y la Antártica Chilena, Chile	55°10'09"S 67°29'35"O	20	Axiom Latino
Trelew	Tehuelche	Trelew, Provincia del Chubut, Argentina	43°15'12"S 65°18'34"O	3	Axiom Latino
Puerto Madryn	Tehuelche	Provincia del Chubut, Argentina	42°49'46"S 65°04'56"O	3	Axiom Latino
Esquel	Tehuelche	Provincia del Chubut, Argentina	42°54'00"S 71°19'00"O	12	Axiom Latino (2)*
Comodoro Rivadavia	Tehuelche	Provincia del Chubut, Argentina	45°52'00"S 67°30'00"O	2	Axiom Latino
Número Individuos				138	

1. Los asteriscos en la Tabla 1 representan al número de individuos que se genotipificaron nuevamente con el microarreglo de MEGA.

4.2. Control de Calidad

4.2.1. Filtro de Individuos y de Variantes

Los análisis de control de calidad para los resultados de la genotipificación para cada plataforma (Axiom y Mega) fueron realizados con PLINK (Purcell et al., 2007), donde se incluyeron distintos parámetros: a) se evaluó la tasa de error de genotipificación por individuo y por SNPs para detectar la calidad de genotipado sobre las muestras y sobre cada sitio del microarreglo, b) la identidad por descendencia (IBD) que comparó similitudes en el genoma entre los individuos, detectando si hay pares de individuos que son más similares respecto a lo esperado por el azar, c) la distribución del alelo en menor frecuencia (MAF), d) Hardy-Weinberg para evaluar aquellos SNPs que se escapan del equilibrio (se utilizó un p-value de 0.0000001 para la detección de marcadores fuera del equilibrio HW), y e) Control de Sexo, mediante la comparación del sexo de los individuos de acuerdo al pedigree con el calculado a partir del cromosoma X. Se eliminaron aquellos individuos que poseían una pérdida de datos superior al 5%, así como también aquellos SNPs que poseían una pérdida sobre el 5% en la muestra y aquellos que no estaban en equilibrio de Hardy Weinberg (p.value 1×10^{-6}) con el fin de evitar errores en la interpretación de los datos.

4.2.2. Comparación Plataforma Axiom y MEGA

Como medida de control, se evaluó la proporción del error de genotipificación para aquellas muestras que se encuentran genotipificadas para ambas plataformas. Para esto, calculamos la proporción de SNPs genotipados diferente entre ambas plataformas. Luego, se evaluó su similitud genotípica a través de una matriz de distancias por genotipos realizado mediante PLINK (Purcell et al., 2007). En este último caso, se fusionaron ambas plataformas solo con aquellos individuos duplicados y con el número de SNPs en común. El dendograma se realizó con el

lenguaje de programación R (R Core Team, 2017) utilizando el método “Manhattan” como medida de distancia para la matriz.

4.3. Análisis de Estructuración Poblacional mediado por Ancestrías

4.3.1. Análisis de Ancestría Global

Se determinó la composición y diversidad de los componentes ancestrales sobre los individuos pertenecientes a las poblaciones de la Patagonia mediante ADMIXTURE (Alexander & Lange, 2011). Estos análisis se realizaron por separado para la plataforma de Axiom y MEGA. Como poblaciones de referencia se incorporaron datos del Proyecto 1000 Genomas (The Genomes Project Consortium, 2012): 30 individuos Africanos de ancestría Yoruba de Nigeria (YRI), 30 individuos europeos provenientes de España y Portugal (IBS), 30 individuos de Beijing, China (CHB), y 30 individuos provenientes de Puno en Perú, facilitadas por el Dr. Andrés Moreno, investigador del Laboratorio Nacional de Genómica para la Biodiversidad (LANGEBIO, CINVESTAV, Irapuato, Gto, México). Además se utilizaron como referencia 34 individuos provenientes de Trapa Trapa (Pehuenche) y San Juan de la Costa (Huiliche). Estos individuos, desde ahora Mapuches, poseen un componente Nativo superior al 99%. El número de componentes ancestrales se determinó mediante un test de Cross-Validation, el cual permite identificar mediante likelihood el mejor valor de K (número de poblaciones ancestrales) para explicar los datos. Posterior al análisis de ancestría global, se estimó el número de individuos que estén relacionados entre sí mediante REAP (Relatedness Estimation in Admixed Populations) el cual estima coeficientes de parentesco entre pares de individuos y las probabilidades de compartir un alelo mediante IBD (Identidad por Descendencia) (T. Thornton et al., 2012). Este análisis funciona en base a ADMIXTURE, por lo que se realizó en base al valor de K que mejor se ajustó a los datos mediante el *cross-validation*. Se tomo como puntaje de corte un coeficiente de parentesco del 12.5, de manera que todos aquellos pares de individuos que hayan

presentado un valor superior a este fueron evaluados, y se procedió a remover uno de estos. El criterio para esto fue de eliminar aquellos individuo con menor aporte del componente Nativo. Los gráficos se realizaron en el lenguaje de programación de R (R Core Team, 2017).

4.3.2. Análisis de Componentes Principales

Se realizó un análisis de componentes principales sobre ambas plataformas por separado (Axiom y Mega) con la finalidad de explorar el componente genético de los individuos mediante smartPCA, el cual viene asociado al programa de EIGENSOFT (N. Patterson et al., 2006) y que permite analizar las diferencias de ancestrías en una población, permitiendo evaluar por estratificación. Para esto se incluyeron las siguientes poblaciones de referencia: 30 individuos de Nigeria (YRI, África), 30 individuos de España (IBS, Europa), 30 individuos de China (CHB, Asia), 30 individuos del Altiplano Andino (Puno, América) y 34 individuos Pehuenche (18) y Huilliche (16). Los datos de las primeras cuatro poblaciones de referencia fueron obtenidas a partir de 1000 Genomas. Dado que este análisis requiere la independencia de los datos, se removieron aquellos SNPs que estuviesen en desequilibrio de ligamiento, además se removieron aquellos SNPs que fueran monomórficos (mediante minor allele frequency). Los gráficos se realizaron en el lenguaje de programación de R (R Core Team, 2017).

4.3.3. Análisis de Ancestría Local

4.3.3.1. Ajuste de Fase

Este proceso consistió en la reconstrucción de las combinaciones alélicas existentes en cada cromosoma a partir de los datos de genotipificación. Esto se debió en un principio a que no es posible asignar alelos de diferentes loci heterocigotos a un cromosoma homólogo (materno o paterno) durante el proceso de genotipificación. Para poder realizar este ajuste de fases se utilizó el programa

SHAPEIT2 (Delaneau, Marchini, & Zagury, 2012) que permitió inferir la combinación alélica sobre los cromosomas. El ajuste de fase se realizó por separado para las muestras de Axiom y Mega. Se utilizaron 2504 individuos de la base de datos del Proyecto 1000 Genomas (The Genomes Project Consortium, 2012) en la fase 3, 110 individuos de Puno genotipificados por Axiom y 210 genotipificados por Mega, Perú, 35 TRIOS correspondientes a individuos provenientes de sectores urbanos (Proyecto Chile Genómico), 34 individuos provenientes de poblaciones Pehuenche y Huilliche (Tabla 2), y los individuos de nuestro estudio (Tabla III).

Tabla III. Descripción de las poblaciones de referencia utilizadas para este estudio. En el género, M = Masculino, F = Femenino y SE = Sin Especificar.

Continente	Región	Individuos	Género			Referencia
			M	F	SE	
Africa	Afro-Caribeños	96	47	49	0	1000 Genomas
	Afro-Americanos SW	61	26	35	0	1000 Genomas
	Esan	99	53	46	0	1000 Genomas
	Gambia	113	55	58	0	1000 Genomas
	Luhya	99	44	55	0	1000 Genomas
	Mende	85	42	43	0	1000 Genomas
	Yoruba	108	52	56	0	1000 Genomas
América	Colombia	94	43	51	0	1000 Genomas
	México	64	32	32	0	1000 Genomas
	Perú	85	41	44	0	1000 Genomas
	Puerto Rico	104	54	50	0	1000 Genomas
Este Asia	Chinos Dai	93	44	49	0	1000 Genomas
	Chinos Han	103	46	57	0	1000 Genomas
	Chinos Han Sur	105	52	53	0	1000 Genomas
	Japón	104	56	48	0	1000 Genomas
	Vietnam Kinh	99	46	53	0	1000 Genomas
Europa	Utah	99	49	50	0	1000 Genomas
	Finlandia	99	38	61	0	1000 Genomas
	Gran Bretaña	91	46	45	0	1000 Genomas
	España	107	54	53	0	1000 Genomas
	Toscanos	107	53	54	0	1000 Genomas
Sur Asia	Bengali	86	42	44	0	1000 Genomas
	Gujarati	103	56	47	0	1000 Genomas

	India	102	59	43	0	1000 Genomas
	Punjabi	96	48	48	0	1000 Genomas
	Sri Lanka	102	55	47	0	1000 Genomas
America	Puno	30	15	15	0	Este Estudio
	Pehuenche	27	13	13	1	Este Estudio
	Huilliche	18	8	10	0	Este Estudio
	Caremapu	6	0	5	1	Este Estudio
	Quetalmahue	8	3	10	5	Este Estudio
	Detif	20	9	11	1	Este Estudio
	Chono	30	9	20	1	Este Estudio
	Tehuelche	20	13	7	0	Este Estudio
	Kawésqar	23	3	1	0	Este Estudio
	Yámana	20	7	13	0	Este Estudio

4.3.3.2. Análisis de Ancestría Local

Se utilizó el programa RFMix que mediante una aproximación “random forest” logró clasificar el origen ancestral de los haplotipos sobre el genoma (Maples, Gravel, Kenny, & Bustamante, 2013b). Este análisis se realizó con los cromosomas faseados a partir del paso anterior. La inferencia de las ancestrías locales sobre los genomas fueron clasificados como: Europeo, Africano y Nativo. Para esto se tomaron como poblaciones de referencia a la Península Ibérica (IBS, España) y Yoruba (YRI, Nigeria) (1000 Genomas) para las dos primeras categorías, mientras que para los Nativos se utilizaron los 34 individuos Mapuches con una alta ancestría amerindia (>99%). La selección de estos grupos ancestrales se determinó a priori como las principales fuentes de contribución genética a las poblaciones de Chile y Suramérica. El tamaño de la muestra para los grupos de referencia Europeo y Africano fue el mismo al número de individuos de origen nativo que se tomaron como referencia. Por otro lado, bajo la aproximación de poder diferenciar las ancestrías de los segmentos cromosómicos entre el componente nativo, se realizó un segundo análisis de ancestría local, esta vez reemplazando la ancestría africana en las muestras de Patagonia por el componente Aymara, representado por Puno.

Como una medida de control y precisión de las inferencias globales y locales en la muestra, se compararon mediante un test de correlación. Para esto, se calculó el porcentaje de ancestría Europeo, Nativo y Africano sobre cada individuo (a partir de las inferencias de ancestría local) y se contrastaron con los porcentajes obtenidos mediante ADMIXTURE tomando en cuenta un total de $K=3$ componentes ancestrales. Lo mismo se llevó a cabo con la inferencia de los componentes Aymara, Europeo y Mapuche, pero tomando un cuenta un total de $K=4$ componentes ancestrales. Los gráficos se realizaron en el lenguaje de programación de R (R Core Team, 2017).

4.3.4. TRACTs

4.3.4.1. *Componente Europeo, Nativo y Africano*

Se evaluó la distribución de los fragmentos de ancestría local para los individuos de la Patagonia en ambas plataformas con la finalidad de determinar el tiempo (en generaciones) del primer evento de mezcla y posteriores pulsos migratorios. Esto se realizó de acuerdo a dos modelos de migración: a) el primero considera una mezcla inicial entre Europeos y Nativos Americanos, seguido por un pulso migratorio Africano (EUR, NAT + AFR), el segundo modelo b) incluye un segundo pulso migratorio (EUR,NAT + AFR + EUR). Cada modelo entregó para cada población un valor de likelihood que señaló la precisión para predecir la distribución de las longitudes de los fragmentos. La elección del mejor modelo se basó en el mayor valor de likelihood. Los gráficos se realizaron en el lenguaje de programación de R (R Core Team, 2017), utilizando el paquete ggplot2 (Wickham, 2016).

4.3.4.2. *Componentes Aymara, Nativo y Europeo*

Se realizó un segundo análisis de TRACTs con los componentes Aymara (Puno), Nativo y europeo con la finalidad de determinar los eventos de mezcla dentro de las ancestrías americanas (Aymara y Mapuche). Para esto se asumieron los mismo modelos de mestizaje: a) mezcra inicial entre el componente Aymara y Mapuche

seguido por un pulso Europeo (PUNO, NAT + EUR), mientras que b) el segundo modelo incluye un segundo pulso de origen europeo (PUNO, NAT + EUR + EUR). Se escogió aquel modelo con un menor valor de Likelihood. Los gráficos se realizaron en el lenguaje de programación de R (R Core Team, 2017), utilizando el paquete ggplot2 (Wickham, 2016).

4.4. Enmascaramiento Regiones del Genoma No Nativos

Este paso consistió en enmascarar todos aquellos SNPs que tuvieran una ancestría “No Amerindia”, es decir, que estén etiquetadas con ancestría “Africana”, “Europea” o “Desconocida” de acuerdo a la inferencia de ancestría local que se realizó a través de RFMix. De esta manera tuvimos para nuestros datos solo los haplotipos asociados con la ancestría Nativa. El proceso para realizar esta etapa se encuentra descrito en la **Figura 6** y comienza con los genotipos faseados en un archivo en formato VCF (*variant call format*) el cual se puede obtener mediante vcftools (Danecek et al., 2011a) tomando como input los archivos faseados generados por SHAPEIT2. Luego se generó un archivo para cada individuo mediante *VCF -subset* en vcftools. Estos archivos van a servir posteriormente como datos de entrada para el enmascaramiento que se realizó en R (R Core Team, 2017). Para continuar con el proceso se separaron los cromosomas materno y paterno de cada individuo y se realizó un *overlap join*, el cual consiste en tomar la posición física de cada SNP y añadirle una etiqueta con su respectiva ancestría local, tomando para esto los archivos de salida (Outputs) de RFMix. Estos archivos de RFMix (Outputs) entregan las ancestrías a lo largo del cromosoma con sus respectivas longitudes (posición de inicio y de término). Posteriormente, se procedió a enmascarar reemplazando con valores perdidos (“.”) todos aquellos SNPs que no tengan una etiqueta “Nativa”. Finalmente, se procedió a juntar los cromosomas paterno y materno para cada individuo y se generó una tabla con todos los otros individuos analizados con sus genotipos faseados y enmascarados.

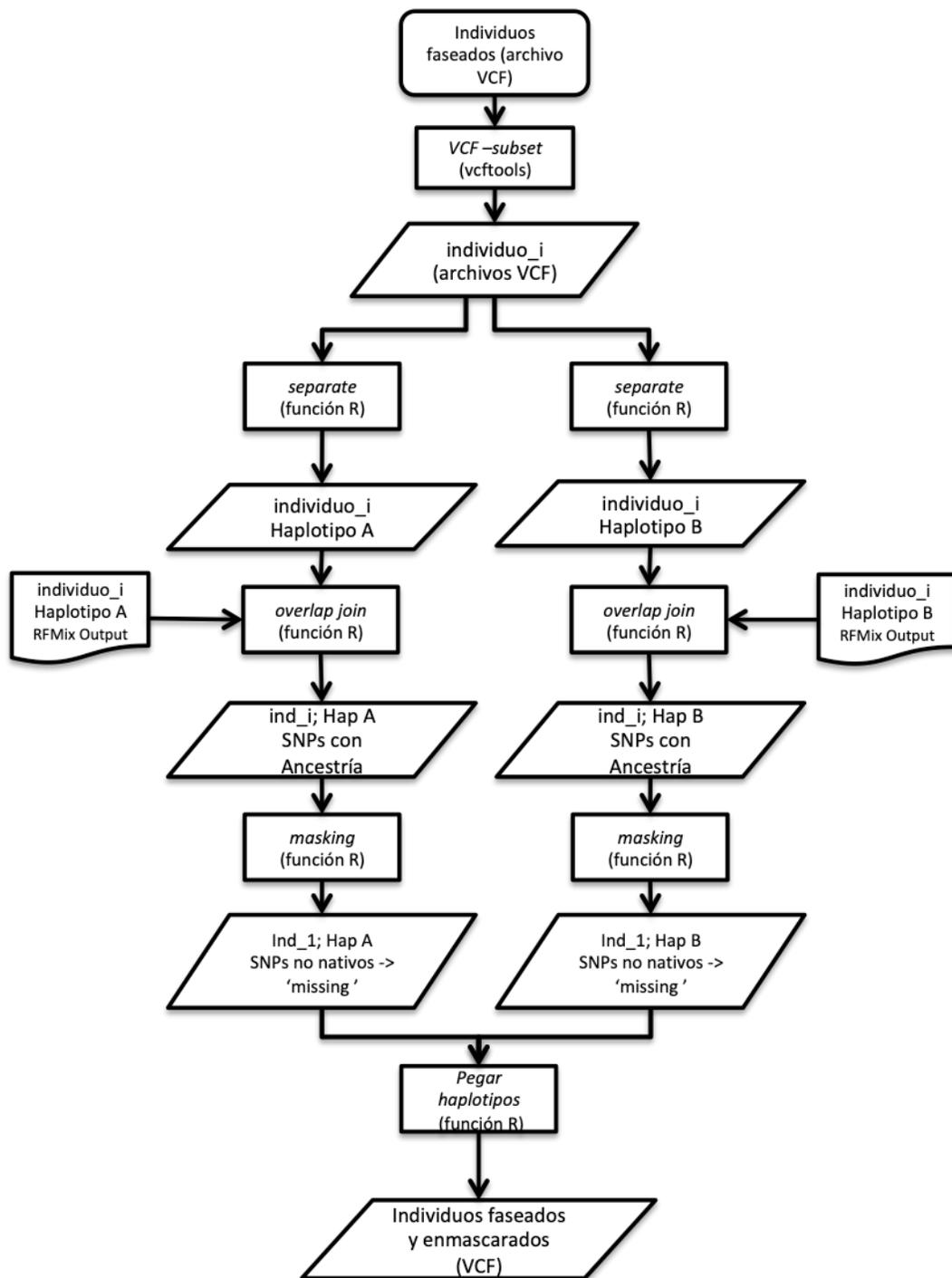


Figura 6. Proceso de enmascaramiento de los segmentos con ancestría local “No Nativos” para los datos imputados de la Patagonia. En la parte superior se muestra como input a los archivos VCF faseados para todos los individuos de la Patagonia. Luego, mediante una función se separaron los cromosomas materno y paterno para cada individuo. Posteriormente, para cada cromosoma se realizó una intersección con los outputs dados por los resultados de la ancestría local. Como resultado, obtuvimos para cada cromosoma

la información de ancestría para cada SNPs y se enmascaró con "." aquellos con ancestría No Nativa. Por último, se juntaron estos cromosomas para cada individuo y se generó un nuevo archivo VCF faseados y enmascarados para todo el genoma en las poblaciones de la Patagonia.

4.5. Análisis de Selección

4.5.1. Populations Branch Statistic (PBS)

Esta metodología analizó las muestras que se generaron como output del paso del enmascaramiento de regiones del genoma no nativos (Figura 6), de manera que se incluyeron aquellos individuos faseados y enmascarados presentes en el archivo VCF. Para el PBS se consideró como outgroup a los Han de China (CHB) disponibles en la base de datos de 1000G (Tabla III). Se usaron los 34 individuos con ascendencia Pehuenche y Huilliche como población cercana a los grupos de la Patagonia, pero considerados *a priori* como no expuestas a las señales de selección que estamos evaluando. Dado que estamos probando regiones bajo adaptación biológica en la Patagonia, realizamos las siguientes agrupaciones: 1) para probar adaptación al frío, agrupamos todas las poblaciones con ascendencia de grupos Patagónicos (Chono, Tehuelche, Kawésqar y Yámana); y 2) para probar adaptación a la dieta marina se removieron los individuos con ascendencia Tehuelche. Los inputs utilizados para este estudio correspondieron a los datos enmascarados que se produjeron en el proceso anterior, dejando aquellos SNPs con ancestría Nativa.

Para el cálculo de PBS se requiere de la comparación pareada entre cada grupo, mediante F_{st} genómico de Weir & Cockerheim (1984) en VCFtools, tomando en cuenta algunos filtros recomendados para el análisis, como la remoción de aquellos SNPs que presentaran datos perdidos y con valores de MAF menores a 1% (Danecek et al., 2011). Se transformó para cada valor de F_{st} un valor logístico de T , que permite estimar el tiempo de divergencia poblacional ajustado al tamaño poblacional (Cavalli-Sforza, 1969). A partir de estos valores se calcularon los PBS para cada SNP de acuerdo a Yi et al., (2010) mostrado en la Figura 7.

Los resultados del PBS se analizaron para cada SNPs y se evaluaron aquellos valores que estuvieron sobre el 1 y 5% de los datos. También se realizaron estos análisis en ventanas de 200Kb mediante bedtools (Quinlan & Hall, 2010), dado que las regiones bajo selección tienden a mostrar clústers de SNPs con elevados valores de selección, mientras que bajo un modelo neutro, SNPs con valores extremos tienden a estar dispersos de manera uniforme en el genoma (Voight et al., 2006). Es por esta razón que se abordaron ambas estrategias. Nos quedamos con aquellas ventanas que tuvieran más de 20 SNPs y que estuvieran por sobre el 5 y 1% de los datos. Por último, se generó un intersección para los valores de PBS entre dos poblaciones para evaluar si son señales únicas de una población, o mas bien, compartidas. Los gráficos se realizaron en el lenguaje de programación de R (R Core Team, 2017), utilizando el paquete ggplot2 (Wickham, 2016). Para el caso de los análisis realizados con el set de datos de MEGA se procedió a remover aquellos individuos que posean una relación directa, la cual esta dada por el pedigree registrado durante la toma de muestras específicamente para la localidad de Puerto Edén.

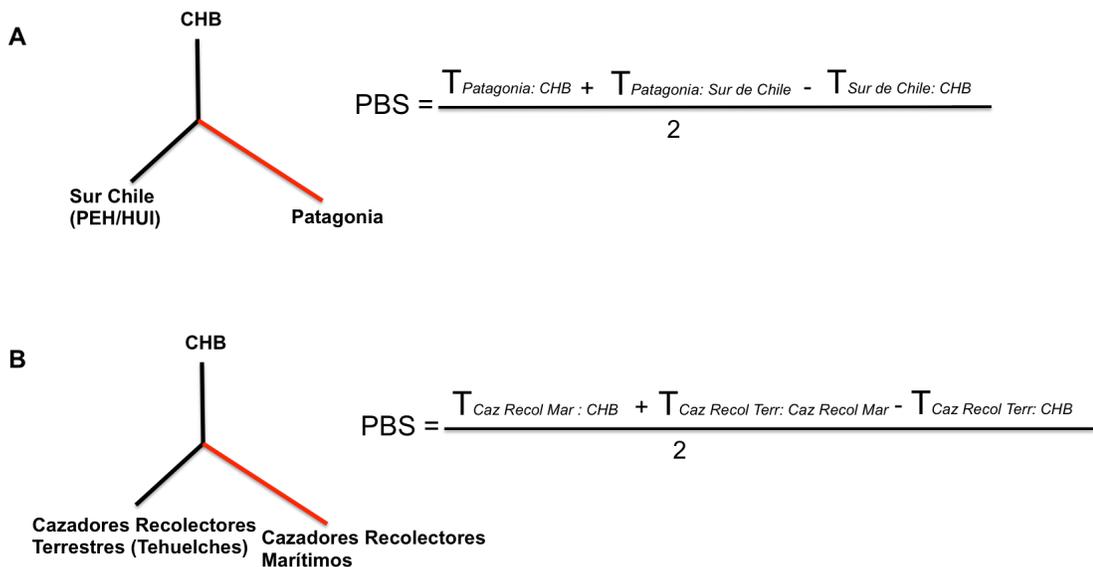


Figura 7. Modelos de comparación utilizados en este estudio con sus respectivo cálculo de PBS. En A) se usó a Han de Asia (CHB), Sur de Chile (PEH/HUI) y todas las poblaciones de Patagonia (grupo objetivo). En B) se usó a Han de Asia (CHB), los cazadores recolectores terrestres (Tehuelches) y los cazadores recolectores marítimos (grupo objetivo).

4.5.2. integrated Haplotype Score (iHS)

Se procedió a evaluar las señales de selección basado en la homocigocidad de haplotipos. Para esto, y al igual que anteriormente, se analizaron aquellos individuos faseados y luego de haber pasado por el enmascaramiento de regiones no nativas del genoma. Se hicieron estos análisis bajo dos agrupaciones principales: a) grupos cazadores recolectores de la Patagonia (Chono, Tehuelches, Kawésqar y Yámana), y b) grupos cazadores recolectores marítimos (Chono, Kawésqar y Yámana). Para cada comparación, se calculó el puntaje iHS para cada SNP y por ventanas de 200Kb de acuerdo a lo señalado por los autores de este método (Pickrell et al., 2009; Voight et al., 2006), mediante el programa Selscan v3.2 (Szpiech & Hernandez, 2014). El análisis de iHS no entrega una prueba estadística, puesto que no se conoce la distribución de los datos bajo condiciones de neutralidad. Por esta razón, se recomienda considerar valores de iHS absoluto ≥ 2.5 y que estén dentro del 1% de la distribución de los valores a nivel genómico (Pickrell et al., 2009). El tamaño de las ventanas (200Kb) se debe a que se ha estimado como suficiente para contener barridos selectivos entre 0.3-0.5 cM (Sabeti et al., 2007).

Dado que las regiones enmascaradas están configuradas como datos perdidos, es necesario filtrar los datos; por un lado remover aquellos individuos con poca ancestría amerindia, y por otro lado remover aquellos SNPs con datos perdidos en la muestra. Esto ultimo, se debe a que este análisis de selección no permite trabajar con información perdida. Lo primero consistió en convertir los archivos VCF en PLINK, luego se evaluó tasa de error de genotipificación (representado por los datos perdidos) en individuos a diferentes valores (5-50%), posteriormente, cada resultado se convirtió en una matriz traspuesta (*tped*) y se removieron todos aquellos SNPs con al menos un dato perdido por cromosoma a través de un script (*filter_genotype.py*). Con esto, se tomó una decisión que involucra elegir el número más propicio de individuos con tal de rescatar la mayor de cantidad de SNPs. Una vez obtenidos los datos filtrados, se calculó el iHS utilizando archivos en formato PLINK (mediante genotipos traspuestos), las posiciones genéticas se extrajeron a partir de

un mapa genético para cada cromosoma. Posterior a esto, los valores de *iHS* fueron normalizados a través de la función *-norm* implementado en *Selscan* (Szpiech & Hernandez, 2014). La normalización se hizo separando los marcadores en 100 grupos de acuerdo a sus frecuencias alélicas, usando la desviación estándar y el promedio por grupo, y se eliminaron los marcadores cuyo alelo de menor frecuencia fuera menor a 0,05. Los gráficos se realizaron en el lenguaje de programación de R (R Core Team, 2017), utilizando el paquete *ggplot2* (Wickham, 2016).

4.5.3. Detección de Loci bajo Selección mediante Ancestría

Una primera aproximación para poder explorar la ancestría local sobre el genoma de las poblaciones mestizas fue la realización de un conteo de ancestría para cada SNP del arreglo. Dado que la idea principal de *RFMix* es de entregar la ancestría local para el genoma de cada individuo, se creó una estrategia que nos permitiera visualizar las frecuencias de ancestría local para todo el genoma y para una población en particular. El conteo se realizó a partir de uno de los archivos de salida que genera *RFMix* (Viterbi), el cual posee para cada haplotipo faseado la ancestría de cada uno de sus SNP. Luego, se contó el número de veces que un SNP era de origen Europeo, Nativo o Africano sobre todos los haplotipos faseados. Mediante este análisis se obtuvieron las frecuencias relativas para cada población ancestral (Europeo, Nativo, Africano) sobre cada SNP del arreglo en todas nuestras poblaciones. Se seleccionaron todas aquellas regiones con ancestría nativa sobre el top 1 y 5% de los datos observados. Para evaluar las señales de selección en grupos cazadores recolectores, se realizaron dos agrupaciones: a) cazadores recolectores de la Patagonia (Chono, Tehuelches, Kawésqar y Yámana), y b) cazadores recolectores marítimos de la Patagonia (Chono, Kawésqar y Yámana). Los gráficos se realizaron en el lenguaje de programación de R (R Core Team, 2017), utilizando el paquete *ggplot2* (Wickham, 2016).

4.6. Identificación de Regiones bajo Selección

4.6.1. Marcadores y Regiones bajo Selección

Aquellos valores que se encontraron por sobre el 1% de la distribución de los datos para el Mapeo Genómico de Ancestría, PBS e iHS se consideraron locus candidatos a selección positiva. Para el caso de los análisis realizados por ventanas de 200Kb, se seleccionaron solo aquellas que contengan como mínimo 20 SNPs y por sobre el 1% de la distribución. Esto último se debe a que especialmente para iHS, existe mayor poder de detección de regiones bajo selección en ventanas con varios marcadores de elevado valor, en comparación a cada marcador por sí solo (Voight, Kudaravalli, Wen, & Pritchard, 2006). El umbral de discriminación del 1% es utilizado con frecuencia en análisis de selección en poblaciones humanas, y corresponde a lo recomendado por los autores de estos métodos (Sabeti et al., 2007; Voight et al., 2006b; Yi et al., 2010). Este valor, aunque conservador para algunos, permite establecer un puntaje de corte preciso e informativo para estudiar asociación entre marcadores o regiones en el genoma frente a algún fenotipo bajo selección (Bigham et al., 2010; Bigham et al., 2009).

4.6.2. Identificación de Genes bajo Selección

Una vez identificados los loci y ventanas sobre el 1% para cada análisis, se procedió a identificar los genes involucrados mediante la utilización del programa de Annovar (K. Wang, Li, & Hakonarson, 2010). Este software analizó las posiciones físicas de cada loci, o el punto de inicio y fin para el caso de las ventanas, permitiendo realizar una búsqueda por posición en la base de datos “Known Canonical Genes” (21.680 genes a nivel genómico) publicado en el Genome Browser de UCSC (F. Hsu et al., 2006). Como resultado se obtuvieron los genes asociados o cercanos a cada loci o ventana consultada. También se consultaron aquellos loci o regiones producto del intersección entre las dos agrupaciones analizadas en este estudio.

4.6.3. Enriquecimiento de Procesos Biológicos

Este método permitió identificar clases de genes que estén sobrerrepresentados (e.g genes bajo selección) en un set de datos de mayor volumen (e.g genes totales en un microarreglo), y que por consiguiente tuvieran una asociación con fenotipos complejos, enfermedades u otros procesos biológicos. Este análisis se llevó a cabo a través del software online GO term Enrichment Analysis (GORilla), el cual usa las categorías de procesos biológicos definido para genes conocido como Gene Ontology (GO) (<http://www.geneontology.org>). El procedimiento de enriquecimiento se realizó con la lista de genes de interés (top 1% de los genes en Mapeo Genómico de Ancestría, PBS e iHS) y una lista de todos los genes *background* asociados en cada agrupación (para todo el microarreglo). El funcionamiento de enriquecimiento toma en cuenta los procesos biológicos (GO) asociados en la lista de interés, definiendo un valor de significancia para la sobrerrepresentación de cada categoría GO respecto a la lista de los genes *background*. La significancia es corregida a posterior mediante múltiples evaluaciones simuladas para cada categoría GO a través de un valor de p-FDR (Yakhini, Eden, Navon, Steinfeld, & Lipson, 2009). Posteriormente, se procedió a evaluar las posibles rutas biológicas, metabólicas o enzimáticas sobre las cuales pudiesen estar participando diferentes procesos biológicos. Esto se realizó utilizando el sitio de WebGestalt (Web-based Gene SeT Analysis Toolkit)¹ el cual permite a través de una interfaz gráfica visualizar distintos parámetros para un set de genes (Liao et al., 2019). A través de esta página web se evaluaron aquellos genes que se encuentran sobrerrepresentados, y se evaluaron sus rutas a través de la base de datos de KEGG² proporcionada por la página y que proporciona distintos mapas de rutas biológicas.

¹ <http://www.webgestalt.org>

² <https://www.genome.jp/kegg/pathway.html>

4.6.4. Comparación de Regiones bajo Selección con Genes asociados a Frío y Dieta

Se compararon todas las ventanas (200Kb) bajo selección analizadas tanto por PBS como iHS con los genes que se describieron bajo selección en distintas poblaciones, especialmente Siberia. Estos genes se encuentran principalmente asociados con adaptación a frío y dieta (FADS, CPT1A, LRP5, PRKG1, THADA). Además, se compararon otros genes asociados con buceo en poblaciones de Indonesia (BDKRB2, CACNA1A, PDE10A). Por último, dado que los análisis por ventanas pasan por varios filtros que podrían originar pérdida de información, se analizaron también aquellos SNPs sobre el 1% de la distribución de los datos cercano a los genes mencionados anteriormente. Los gráficos se efectuaron en el lenguaje de programación de R (R Core Team, 2017) y con el paquete gráfico de chromplot (Verdugo & Orostica, 2018) y ggplot2 (Wickham, 2016).

5. Resultados

5.1. Control de Calidad y Depuración de los Datos

5.1.1. Patagonia (Axiom)

Se evaluó la calidad de genotipificación en la muestra de Axiom mediante PLINK (Purcell et al., 2007) para un total de 607.853 SNPs. La pérdida de datos por individuo refleja que solo una muestra proveniente de la población Tehuelche tiene un porcentaje superior al 5%, por lo que fue removida para todos los análisis posteriores (Figura 8A). La tasa de error de genotipificación para los marcadores muestra un total de 2.287 con un error superior al 5% (Figura 8B), los cuales también serán excluidos para posteriores análisis. Para el análisis de los alelos en menor frecuencia (MAF por sus siglas en inglés) se observa que la mayoría poseen un porcentaje inferior al 6% en la muestra, donde el resto de los alelos se distribuye de manera homogénea (Figura 8C). También se determinaron aquellos marcadores que se encontraban fuera del equilibrio de Hardy-Weinberg (eq. HW) con la finalidad de eliminarlos asumiendo que corresponden a probables errores en su genotipado (Figura 8D). Por último, se evaluó la identidad por descendencia (IBD) mediante múltiples comparaciones pareadas. Se encontró que una gran cantidad de comparaciones cayó en un rango inferior al 12.5% de IBD (6157 de 7503), mientras que 64 comparaciones fueron mayores a un 25% de IBD (Figura 8E). En la Figura 9 podemos apreciar un esquema con los filtros tanto de individuos como de marcadores utilizados para el control de calidad y para los posteriores análisis en este estudio.

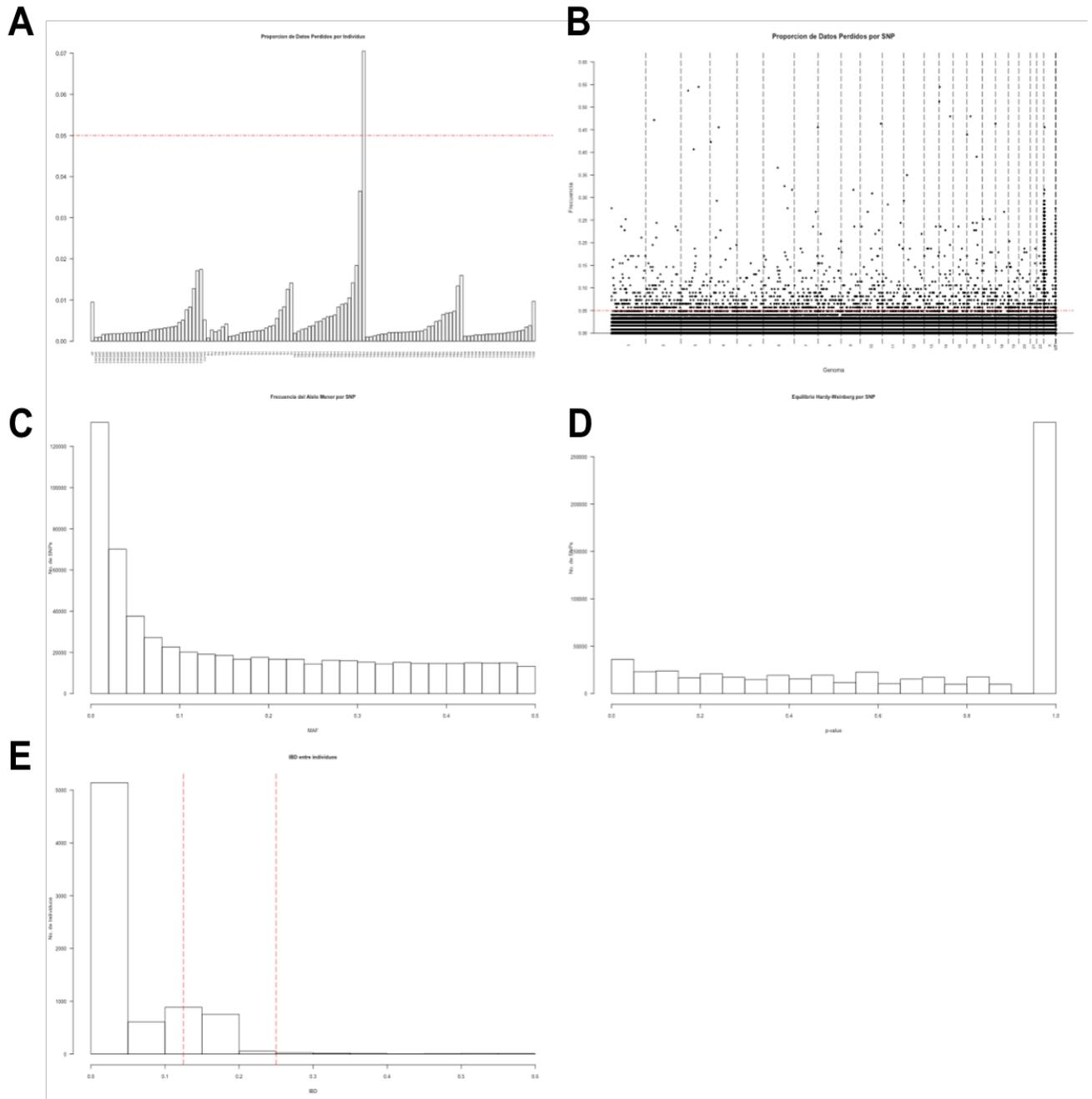


Figura 8. Análisis de Control de Calidad para las muestras de la Patagonia. En A se muestra la tasa de pérdida por individuo, la línea roja representa una tasa de error superior al 5%. En B la tasa de pérdida por marcador donde la línea roja indica aquellos SNPs con una tasa de pérdida superior al 5%. En C el MAF (“Minor Allele Frequency”) para todos los marcadores del microarray. En D se muestra la distribución del equilibrio Hardy Weinberg para todos los SNPs presentes en la muestra. En E las comparaciones pareadas de coeficiente de parentesco (IBD) para todos los individuos incluidos en este análisis, donde las línea vertical roja indica un coeficiente de 0.125 y 0.25 respectivamente.

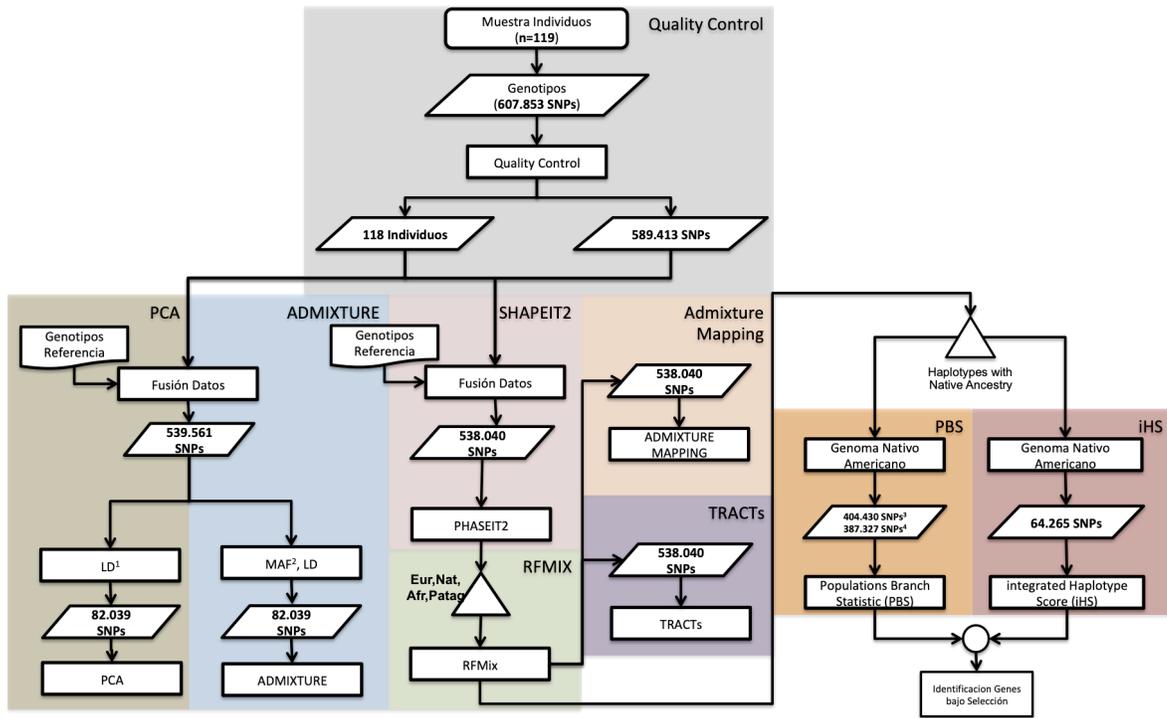


Figura 9. Esquema con el número de SNPs e individuos utilizados para los principales análisis en este estudio. Cada cuadro coloreado representa un análisis diferente. Los triángulos representan sub-set de elementos, en este caso de individuos específicos, o solo haplotipos con ancestría nativa. 1. LD: desequilibrio de ligamiento, 2. MAF: minor allele frequency, 3. Corresponde al número de SNPs para los cazadores recolectores de la Patagonia, 4. Corresponde al número de SNPs para los cazadores recolectores marítimos de la Patagonia.

5.1.2. Patagonia Austral (MEGA)

Se evaluó la calidad de genotipificación en las muestras de MEGA mediante PLINK (Purcell et al., 2007) para 1.476.122 SNPs. Para la tasa de genotipificación vemos que ningún individuo sobrepasa el 5% de error (Figura 10A), por otro lado, para los SNPs notamos que solo 21.641 sobrepasan el 5% de error de genotipado, por lo que serán removidos en los análisis posteriores (Figura 10B). De acuerdo a distribución del alelo en menor frecuencia (MAF) notamos que la mayoría de los marcadores poseen una frecuencia menor al 1% (Figura 10C). También se encontraron 62 SNPs fuera del equilibrio Hardy-Weinberg (P-value < 0.000001) por

lo que se eliminaron de los datos (Figura 10D). Finalmente, se evaluó el parentesco en los individuos de la muestra a través de IBD, para esto se removieron primero todos aquellos SNPs en estado monomórficos –solo para este análisis-, quedándonos con un total de 586.158 marcadores. Tendiendo esto en cuenta, podemos mencionar que hay un total de 67 observaciones con IBD mayor al 12.5%, mientras que 27 presentaron un IBD mayor al 25% (Figura 10E).

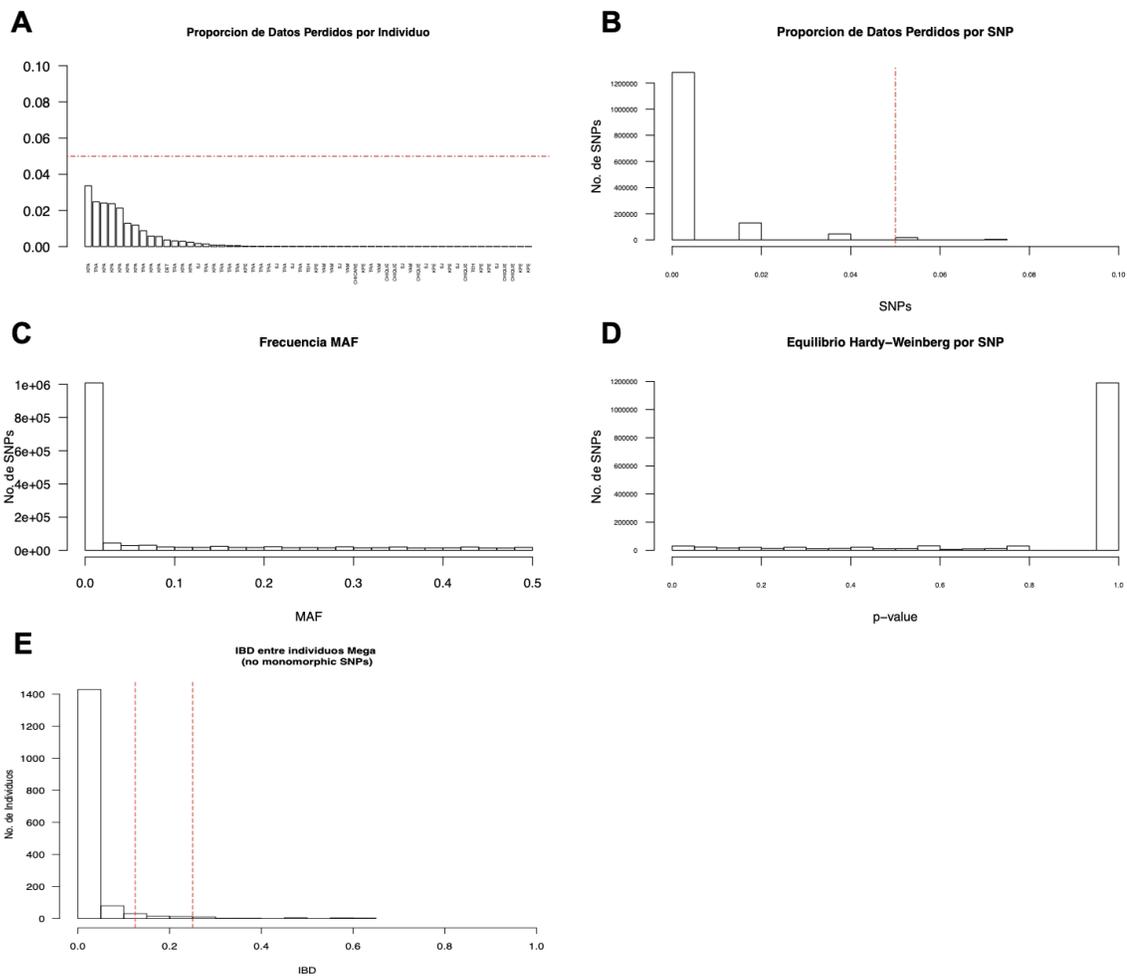


Figura 10. Análisis de Control de Calidad para las muestras de la Patagonia genotipificadas en MEGA (Illumina). En A se muestra la tasa de pérdida por individuo, la línea roja representa una tasa de error superior al 5%. En B la tasa de pérdida por marcador donde la línea roja indica aquellos SNPs con una tasa de pérdida superior al 5%. En C el MAF (“Minor Allele Frequency”) para todos los marcadores del microarray. En D se muestra la distribución del equilibrio Hardy Weinberg para todos los SNPs presentes

en la muestra. En E las comparaciones pareadas de coeficiente de parentesco (IBD) para todos los individuos incluidos en este análisis, donde las línea vertical roja indica un coeficiente de 0.125 y 0.25 respectivamente.

5.2. Comparación Plataformas Axiom y MEGA

Se comparó la calidad de genotipificación entre las plataformas de Axiom y MEGA utilizando aquellos individuos que fueron genotipificados en ambos sets. Esto nos permite controlar la calidad de genotipificación entre ambas plataformas para análisis posteriores. Para esto, se realizó una matriz de distancias basado en los genotipos determinados por cada plataforma de genotipificación a través de *PLINK* para un total de 109870 SNPs (comunes para Axiom y MEGA). En la Figura 11 podemos ver que los clúster con menor distancia corresponden a cada individuo genotipificado tanto por Axiom como por MEGA. Las diferencias de genotipificación de ambas plataformas no sobrepasa el 1% con excepción de dos individuos los cuales sobrepasan el 2%, estas discrepancias corresponden a errores en la genotipificación ya sea en las plataformas de Axiom o MEGA, no obstante, estos errores son muy bajos respecto a la totalidad de los datos (Figura 12).

Axiom vs Mega

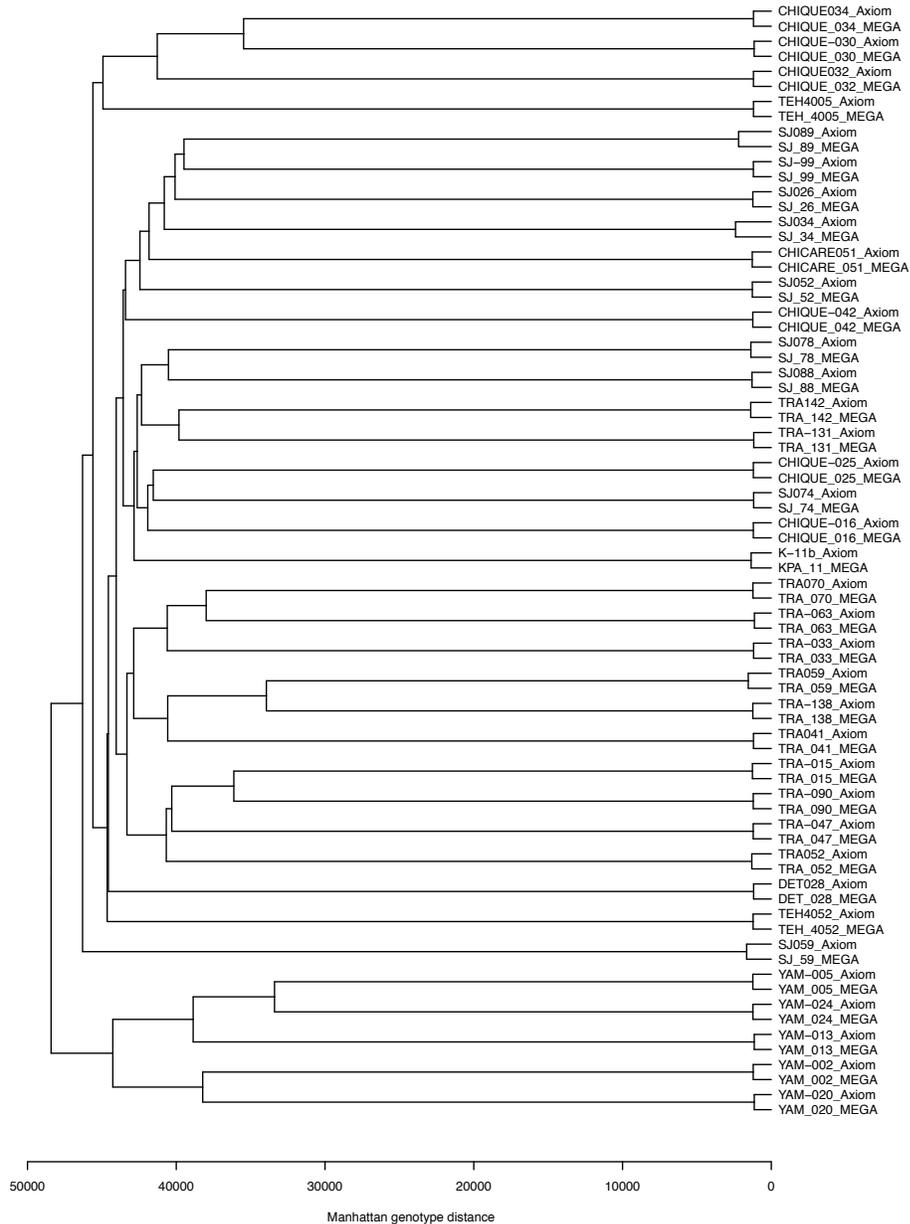


Figura 11. Dendrograma basado en la matriz de distancias Manhattan entre las plataformas de genotipificación de Axiom y MEGA. Se utilizaron para esto aquellos individuos genotipificados en ambas plataformas.

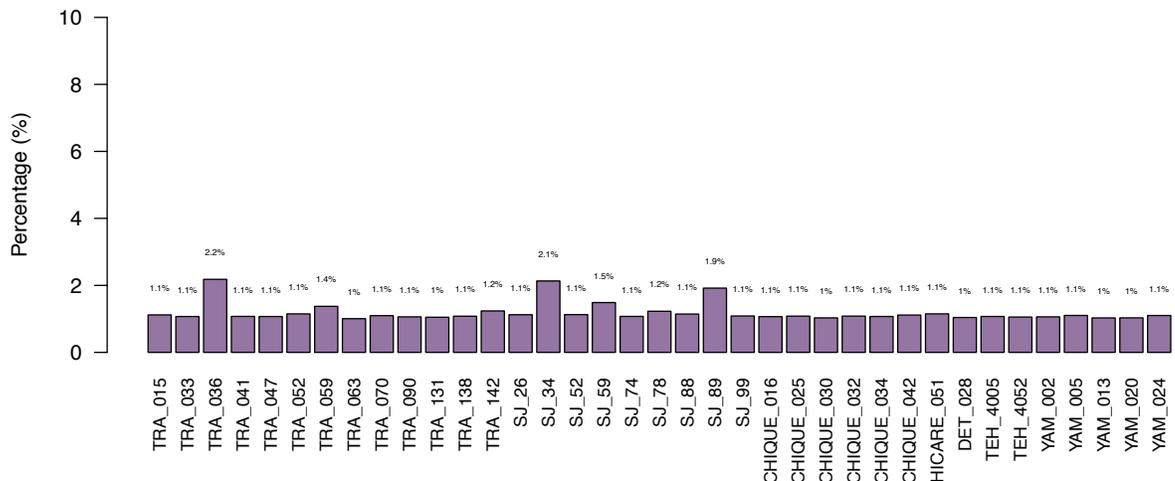


Figura 12. Comparación del error de genotipificación para las plataformas de Axiom y MEGA. Los individuos que aparecen en el eje horizontal corresponden a aquellos que se encuentran genotipificados en ambas plataformas. Los valores se encuentran representados en porcentajes.

5.3. Análisis de Estructuración Poblacional mediante Ancestrías

5.3.1. Análisis de Ancestría Global

Se estimó la ancestría global para distintos componentes ancestrales mediante ADMIXTURE, esto con el propósito de poder obtener la proporción de cada ancestría en la muestra. La estimación de las ancestrías fue calculada a partir de diferentes poblaciones de referencia. Estos grupos corresponden a Nigeria (África), Península Ibérica (Europa), China (Este de Asia), Puno (Altiplano Andino, Suramérica) y Chile (Pehuenche y Huilliche, Suramérica). El número de poblaciones ancestrales (K) que arrojó un menor valor de likelihood para el error de validación cruzada fue de K=4 (Figura 13A).

Descriptivamente, vemos que en las muestras de las poblaciones de la Patagonia se puede apreciar una mezcla principalmente de dos componentes ancestrales: Europeo y Nativo Americano (compuesto por Puno y Pehuenche-Huilliche). Sin embargo, es posible apreciar un componente africano, aunque comparativamente en una proporción baja. Cuantitativamente, la proporción del componente Europeo en las muestras de la Patagonia (Chono, Tehuelche, Kawéskar y Yámana) corresponde a un 28.2%, siendo los Chono con el valor más bajo (19,1%),

mientras que el resto de las poblaciones poseen valores superiores: 23,9% Tehuelche, 38,8% Kawésqar y 31% Yámana. El promedio de la ancestría Nativa para todas las poblaciones corresponde a un 70.5%. Dentro de esta ancestría, la población Chono con el valor mas elevado (80,6%), seguido por Tehuelche (73%), Yámana (68,1%) y Kawésqar (60,2%) (Figura 14).

Ahora bien, cuando analizamos el K=5 (Figura 14), nos encontramos con una diferenciación dentro del componente Amerindio, por un lado un componente asociado con el Altiplano Andino (Puno) y por otro, las poblaciones Pehuenche y Huilliche del Centro Sur de Chile. Las muestras de la Patagonia evidencian un 4,9% de componente Andino. Este componente está pobremente representado en la población Chono (0.5%), pero a medida que avanzamos hacia el extremo austral de la Patagonia, este componente empieza a aumentar, observando un 5,7% en la población Kawésqar y un 6,2% en los Yámana. Los Tehuelches por otro lado son los que presentan un mayor componente Aymara con un 10,5%. El segundo componente que podemos desprender de la ancestría Nativa corresponde al Centro Sur de Chile, el cual alcanza en la Patagonia un 69,2%. Al contrario de lo que podemos ver con el componente Andino, la población Chono posee un aporte del 80,2%, el cual va disminuyendo progresivamente hacia el Extremo Sur de la Patagonia con un 62,6% en la población Tehuelche, un 54,5% en la población Kawésqar y un 61,9% en la población Yámana. La ancestría africana, como se mencionó anteriormente, representa un aporte bajo en la población (0.8%). La población Tehuelche posee el aporte mas alto de un 2,2%. La ancestría de origen Asiático al igual que el Africano, se mantiene en frecuencias menores al 0,5% (Tabla IV).

Se realizó el análisis de parentesco mediante REAP tomando en cuenta los cinco componentes ancestrales que arrojó el test de *cross-validation* por ADMIXTURE y se realizó nuevamente el análisis de ancestría global. Se encontraron 22 pares de individuos en la población que presentan un coeficiente de parentesco mayor a 12.5. De cada par de individuos se escogió aquel que tuviera el mayor aporte de ancestría Nativa. De esta manera, se removió un sujeto de Trapa-Trapa

(Pehuenche), siete probablemente Chono (Laitec) y seis de Ukika (Yámana). El valor de likelihood que mas se ajusta a los datos con la muestra filtrada por REAP corresponde a un $K=4$ (Figura 13B). En relación al análisis de ancestría global sin el filtro de parentesco, se puede observar que las variaciones de los principales componentes ancestrales (Europeo, Aymara, Mapuche) varían en no más del 1,5%. El componente Africano y Asiático se mantienen en bajas frecuencias y sin grandes modificaciones (Tabla IV). Además, para los cambios en las ancestrías en las diversas localidades de la Patagonia podemos observar los mismos patrones mencionados en el análisis anterior (Figura 15).

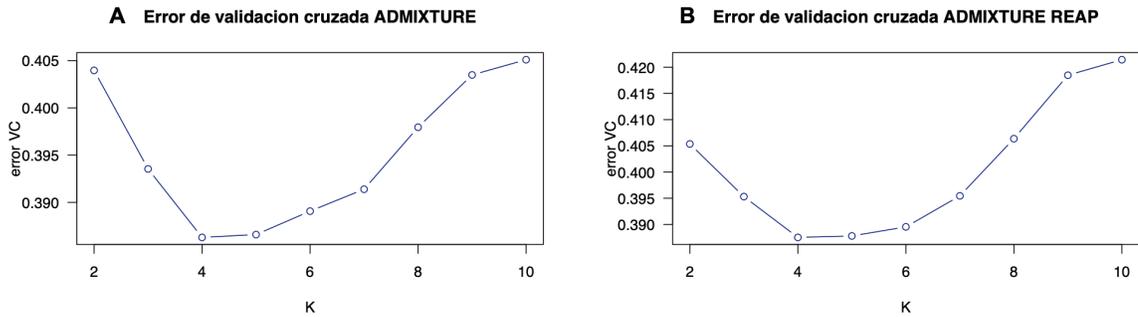


Figura 13. Error de validación cruzada para el análisis de ADMIXTURE (cross-validation, CV). En A se muestra el CV para todos los individuos de la muestra, mientras que en B se muestra el CV posterior al filtrado de individuos con un coeficiente de parentesco mayor a 0.125 mediante REAP.

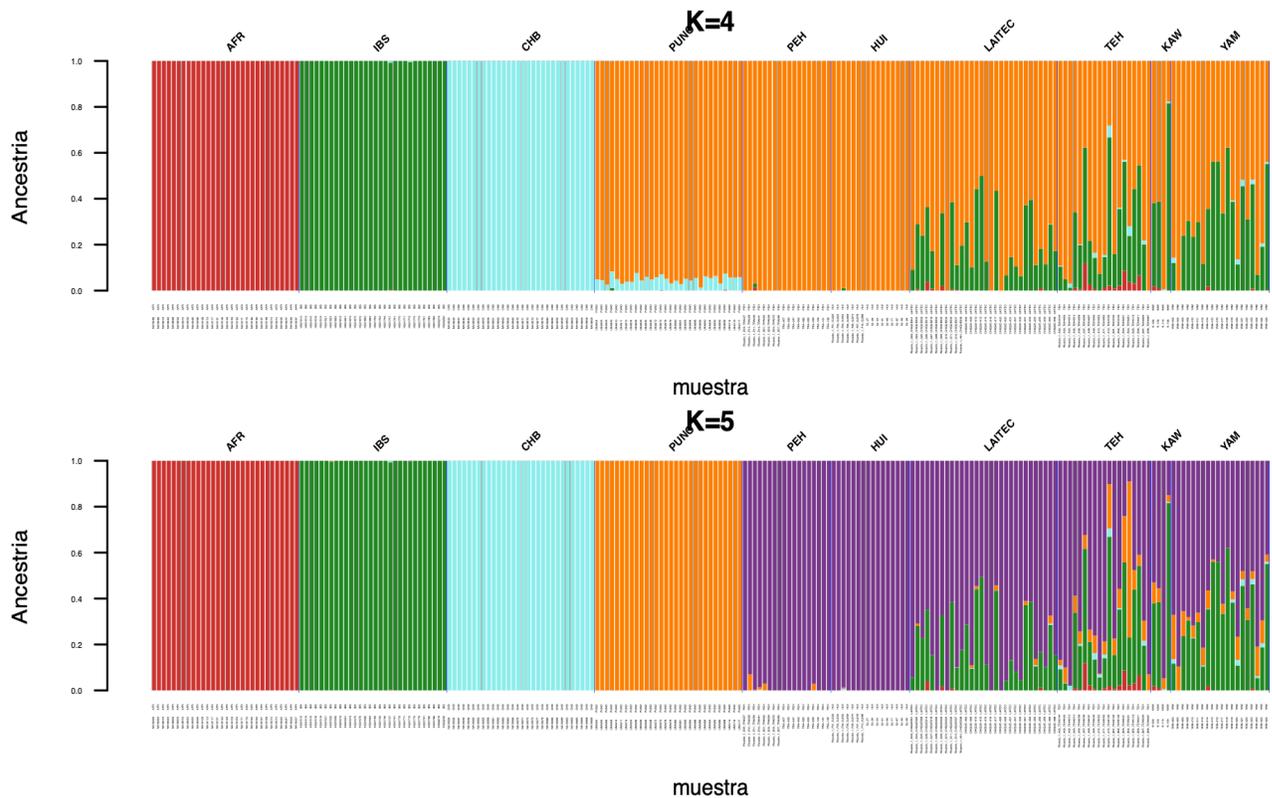


Figura 14. Análisis de Ancestría Global sobre las poblaciones de la Patagonia. En la figura se observan solo las ancestrías con un K=4 y 5. Los paneles de referencia corresponden a Africa (YRI), Europa (IBS), Este de Asia (CHB), Aymara (Puno), Pehuenche (PEH) y Huilliche (HUI).

Tabla IV. Comparación entre promedios de ancestrías por población para ADMIXTURE con filtro por parentesco*.

	ADMIXTURE, K=5					ADMIXTURE (con REAP), K=5				
	África	Europa	Asia	Puno	Sur	África	Europa	Asia	Puno	Sur
Chono	0.0026	0.1905	0.0005	0.0045	0.8019	0.0036	0.2003	0.0008	0.0056	0.7897
	<i>0.0080</i>	<i>0.1488</i>	<i>0.0014</i>	<i>0.0093</i>	<i>0.1539</i>	<i>0.0096</i>	<i>0.1500</i>	<i>0.0019</i>	<i>0.0120</i>	<i>0.1578</i>
Tehuelche	0.0225	0.2385	0.0087	0.1046	0.6257	0.0206	0.2413	0.0081	0.0982	0.6317
	<i>0.0330</i>	<i>0.1890</i>	<i>0.0110</i>	<i>0.1482</i>	<i>0.2805</i>	<i>0.0326</i>	<i>0.1911</i>	<i>0.0102</i>	<i>0.1499</i>	<i>0.2848</i>
Kawéskar	0.0074	0.3877	0.0028	0.0565	0.5455	0.0067	0.3901	0.0026	0.0561	0.5446
	<i>0.0091</i>	<i>0.3336</i>	<i>0.0032</i>	<i>0.0255</i>	<i>0.3261</i>	<i>0.0083</i>	<i>0.3343</i>	<i>0.0035</i>	<i>0.0278</i>	<i>0.3199</i>
Yámana	0.0014	0.3099	0.0074	0.0619	0.6194	0.0026	0.3077	0.0134	0.1078	0.5685
	<i>0.0042</i>	<i>0.1831</i>	<i>0.0097</i>	<i>0.0492</i>	<i>0.1472</i>	<i>0.0054</i>	<i>0.1571</i>	<i>0.0106</i>	<i>0.0689</i>	<i>0.0973</i>
Total	0.0085	0.2817	0.0049	0.0569	0.6481	0.0335	0.2849	0.0249	0.0670	0.6336

* En negrita se muestra los promedios, mientras que en cursiva se incluyó la desviación estándar.

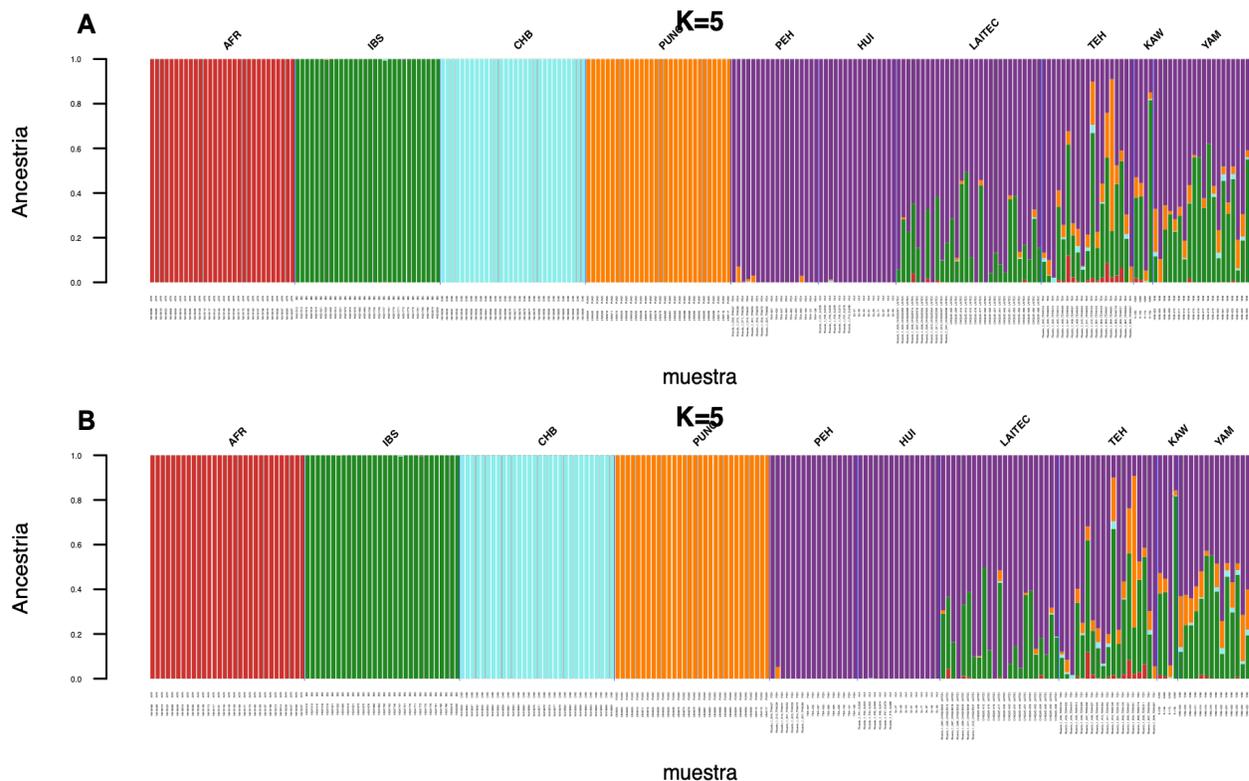


Figura 15. Comparación de Ancestría Global para las muestras de la Patagonia. En A, se puede apreciar el ADMIXTURE para las muestras sin filtrar por REAP (parentesco > 12,5), mientras que en B, se puede apreciar la ancestría global después del filtrado de individuos por parentesco.

5.3.2. Análisis de Componentes Principales

Se realizó un Análisis de Componentes Principales (PCA) para evaluar las diferencias en las ancestrías de las poblaciones de ascendencia Chono, Tehuelche, Yámana y Kawéskar (Figura 16). Para esto se compararon con diferentes poblaciones de referencia: Nigeria (África), España (Europa), China (Asia), Puno (Altiplano Andino, Suramérica), Pehuenche y Huilliche (Sur de Suramérica). El primer y segundo componente principal explican el 6.2% y 3.95% del total de la varianza respectivamente, y nos permiten diferenciar claramente a las poblaciones de África, Europa y Este de Asia. Dentro del componente Amerindio no es posible diferenciar la población del altiplano andino (Puno) con las del Centro Sur de Chile

(Pehuenche y Huilliche). Bajo estos componentes principales podemos ver que las poblaciones de la Patagonia se encuentran asociadas en un estado intermedio, compartiendo las ancestrías entre dos grupos: la población Europea y las poblaciones nativas de Sudamérica.

Dado que en una primera instancia no hay una participación importante del componente Africano y el Asiático, se procedieron a remover de este análisis, y se realizó nuevamente el PCA (**Figura 17**) con la finalidad de remover la variabilidad explicada por las grandes diferencias entre grupos continentales. Mientras que el primer componente principal (6,79%) permite separar el componente Europeo con el Amerindio, vemos que el tercer componente (1,47%) resuelve el componente Aymara del componente centro sur de Chile (**Figura 17A**). Esto genera que los individuos de la Patagonia se encuentren en una primera instancia con un importante componente Europeo, y por otro lado asociados principalmente con los grupos del centro sur de Chile, con excepción de algunos individuos principalmente Tehuelche. Cabe destacar que cuando observamos el cuarto (1,32%) con el quinto (1,23%) componentes principales, se puede distinguir una diferenciación mediada en su mayoría por localidades asociadas a la Patagonia Chilena: por un lado los individuos con ascendencia Chono, y por otro lado individuos pertenecientes a la etnia Yámana. Los individuos Kawésqar se ubican en un punto medio a estos grupos (**Figura 17B**).

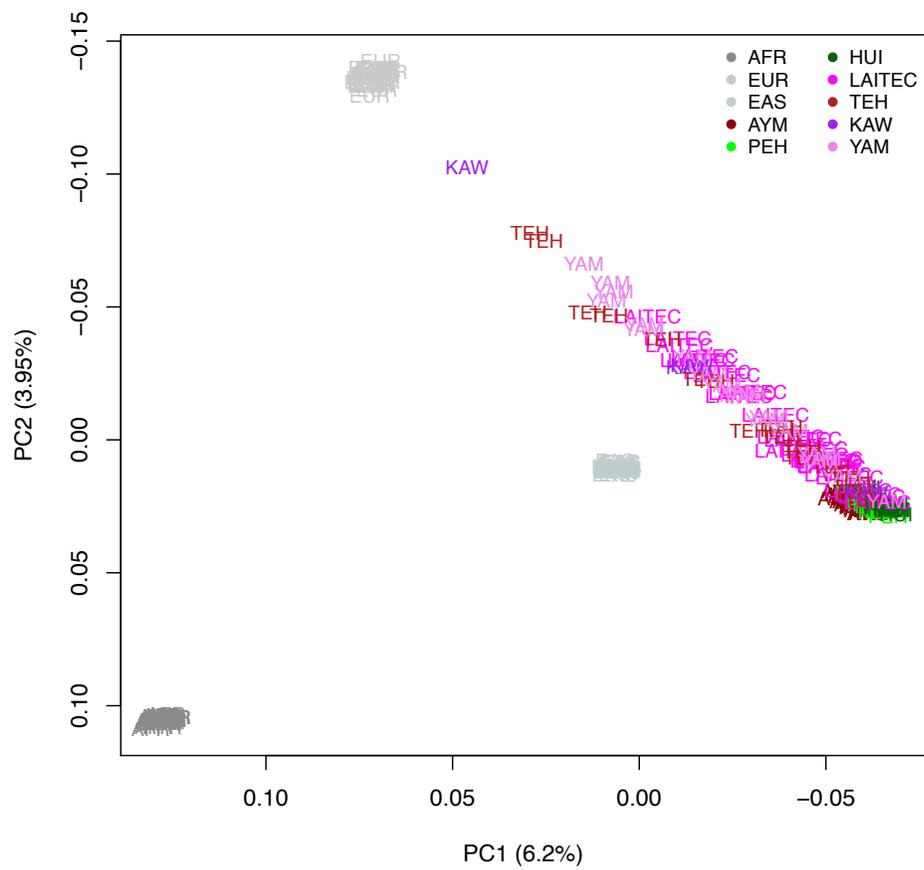


Figura 16. Análisis de Componentes Principales sobre la muestra de la Patagonia. Se utilizaron las poblaciones de referencia Africana (AFR), Europea (EUR), Asiática (EAS), Aymara (AYM), Pehuenche (PEH) y Huilliche (HUI). Las poblaciones utilizadas en la Patagonia están señaladas como: LAITEC (Chono), TEH (Tehuelches), KAW (Kawésqar) y YAM (Yámana).

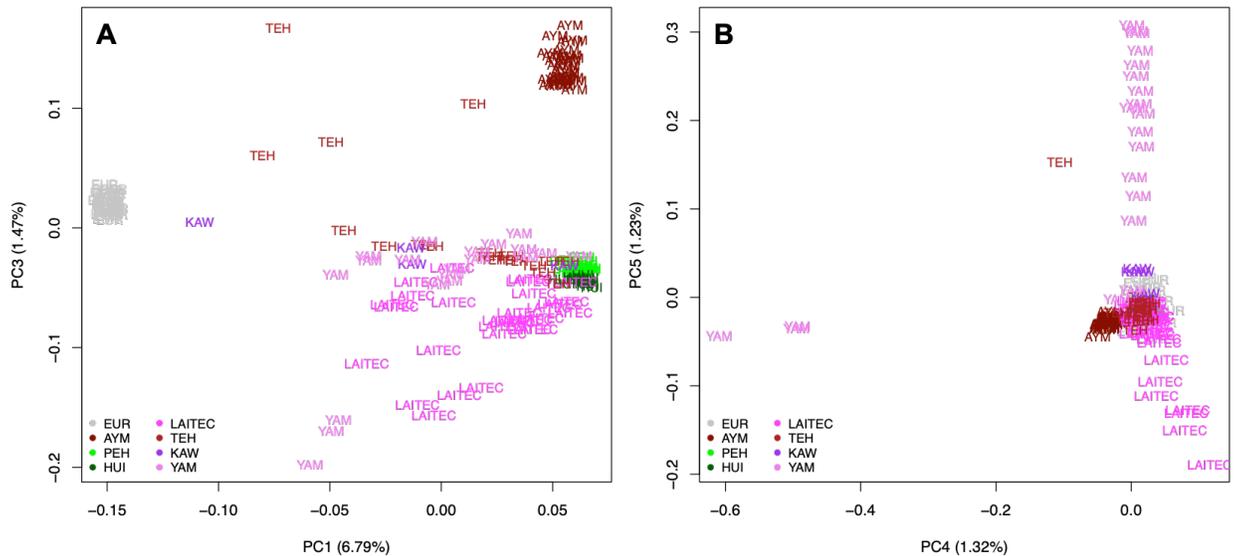


Figura 17. Análisis de Componentes Principales sobre las muestras de la Patagonia. Se removieron las poblaciones de referencia Africana y Asiática. En A, se muestra el primer componente principal (6,79%) y el tercer componente principal (1,47%). En B, el análisis para el cuarto componente principal (1,32%) con el quinto componente principal (1,23%).

5.3.3. Análisis de Ancestría Local

Luego de haber estimado la contribución de los diferentes componentes ancestrales en las poblaciones de la Patagonia se realizó el análisis de ancestría local mediante RfMix. Previo a esto se requirió de un ajuste de fases, el cual consiste en la reconstrucción de los haplotipos materno y paterno en un individuo. Esta inferencia se realizó con SHAPEIT2 como se mencionó en la metodología y con 2504 individuos del Proyecto 1000 Genomas (The Genomes Project Consortium, 2012). En adición a esto, se incluyeron individuos pertenecientes al sur de Suramérica (Tabla III) con la finalidad de obtener un resultado mas preciso en torno a esta región del continente. A partir del ajuste de fase se realizó la inferencia de ancestría local, la cual consiste en la determinación de la ancestría de segmentos cromosómicos específicos. Las ancestrías que se definieron para nuestras poblaciones fueron la Africana (YRI, Yoruba), Europea (IBS, Península Ibérica) y Mapuche (Pehuenche-Huilliche). La

selección de este último grupo de referencia se llevó a cabo a partir del análisis de Ancestría Global, permitiéndonos elegir solo aquellos individuos con un elevado aporte Nativo (>99%). Dentro de estas, 38 individuos cumplen con esta condición, de los cuales 18 pertenecen a la población Pehuenche, 16 a la población Huilliche y 4 con ascendencia Chono. El resultado de la Ancestría Local se traduce finalmente en la disposición de todos los cromosomas de un individuo coloreados de acuerdo a la ancestría que posee cada segmento. La Figura 18 representa un esquema para tres individuos con diferentes proporciones de ancestría Europea, Nativa y Africana. En la Figura 18A se encuentra un individuo con una contribución similar de ancestría tanto Europea como Nativa (obtenido a partir de su información de Ancestría Local). Este individuo posee un tipo “mosaico”, dado que para cada cromosoma existen numerosos segmentos de distinto tamaño y de diferentes ancestrías (Europea, Nativa, Africana o Desconocido). Cabe señalar, que cuando se mira la disposición de los segmentos cromosómicos con ancestría conocida para la mayoría de los individuos en la Patagonia, se observa que estos presentan en su mayoría este tipo de patrón. En cambio, los individuos presentes en las Figura 18B y Figura 18C corresponden a sujetos que tienen una ancestría mayoritariamente Nativa y Europea respectivamente. En ambos casos, se observa que existe una predominancia de sus respectivas ancestrías, con segmentos cromosómicos mucho más extensos que el patrón “mosaico”.

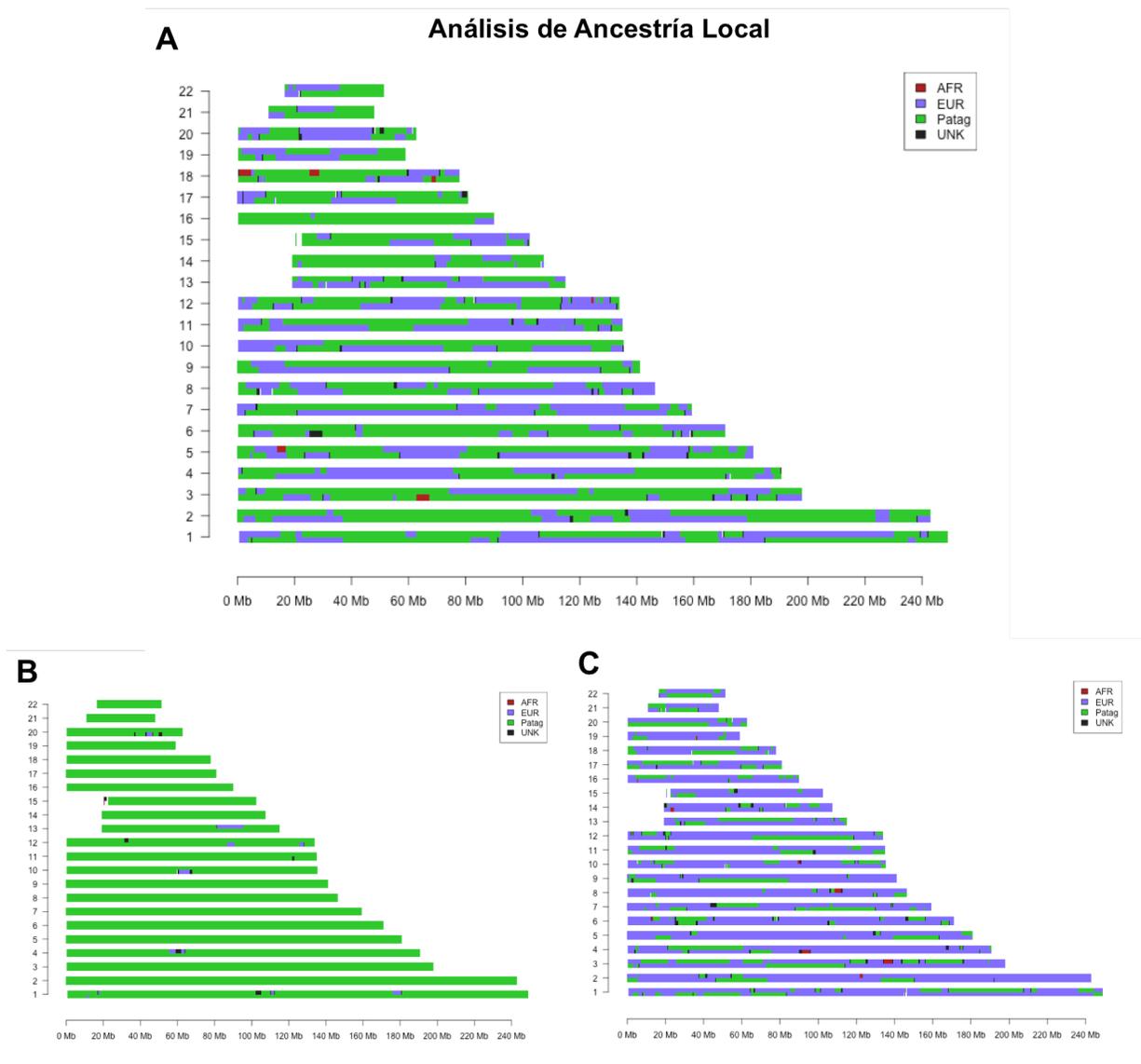


Figura 18. Mapeo Genómico de la Ancestría Local en tres individuos de la Patagonia. En A, se muestra un individuo mestizo con aproximadamente la mitad de su genoma de origen Nativo y la otra mitad Europeo. En B se observa un individuo con una ancestría nativa del 99%. En C vemos un individuo con una ancestría Europea aproximadamente del 80%. En el eje X se observa el largo del cromosoma, mientras que en el eje Y se encuentra el número de los cromosomas autosómicos para cada individuo.

Para comprobar la correcta asignación de las ancestrías locales, se compararon con los resultados de Ancestría Global mediante un test de correlación para estos individuos. En la Figura 19 se encuentra el análisis para todos los individuos de la Patagonia. Las correlaciones se encuentran separadas por ancestría Amerindia, Europea y Africana. Con las dos primeras ancestrías se puede notar que

el coeficiente de correlación entre la inferencia global con la local es igual o superior al 99%, revelando una estrecha concordancia entre estos dos resultados. En cambio para la ancestría Africana vemos que el coeficiente de correlación se encuentra en un 53%, muy por debajo de la ancestría Nativa y Europea.

Dado que la ancestría global aportó evidencia de un componente Aymara importante en las poblaciones de la Patagonia, se realizó un análisis de ancestría local, utilizando para esto tres grupos de referencia; un grupo Aymara (Puno), del Centro Sur de Chile (Pehuenche y Huilliche) y Europeo (IBS). En la **Figura 20** se muestra el test de correlación para comparar la precisión sobre las estimaciones de ancestría para este análisis. En términos generales vemos que los coeficientes de correlación son menores que en el análisis anterior, por lo que se puede desprender que hay una precisión más baja al momento de inferir ancestría local. Esto se ejemplifica para el componente Aymara el cual presenta un coeficiente de correlación del 73% con un aparente sobre-estimación para la ancestría local, mientras que el coeficiente de correlación del componente del Centro Sur de Chile desciende a un 94% con una subestimación por parte de la ancestría local. No obstante, para la comparación del componente Europeo vemos que el coeficiente de correlación corresponde a un 99%, valor muy similar al análisis anterior.

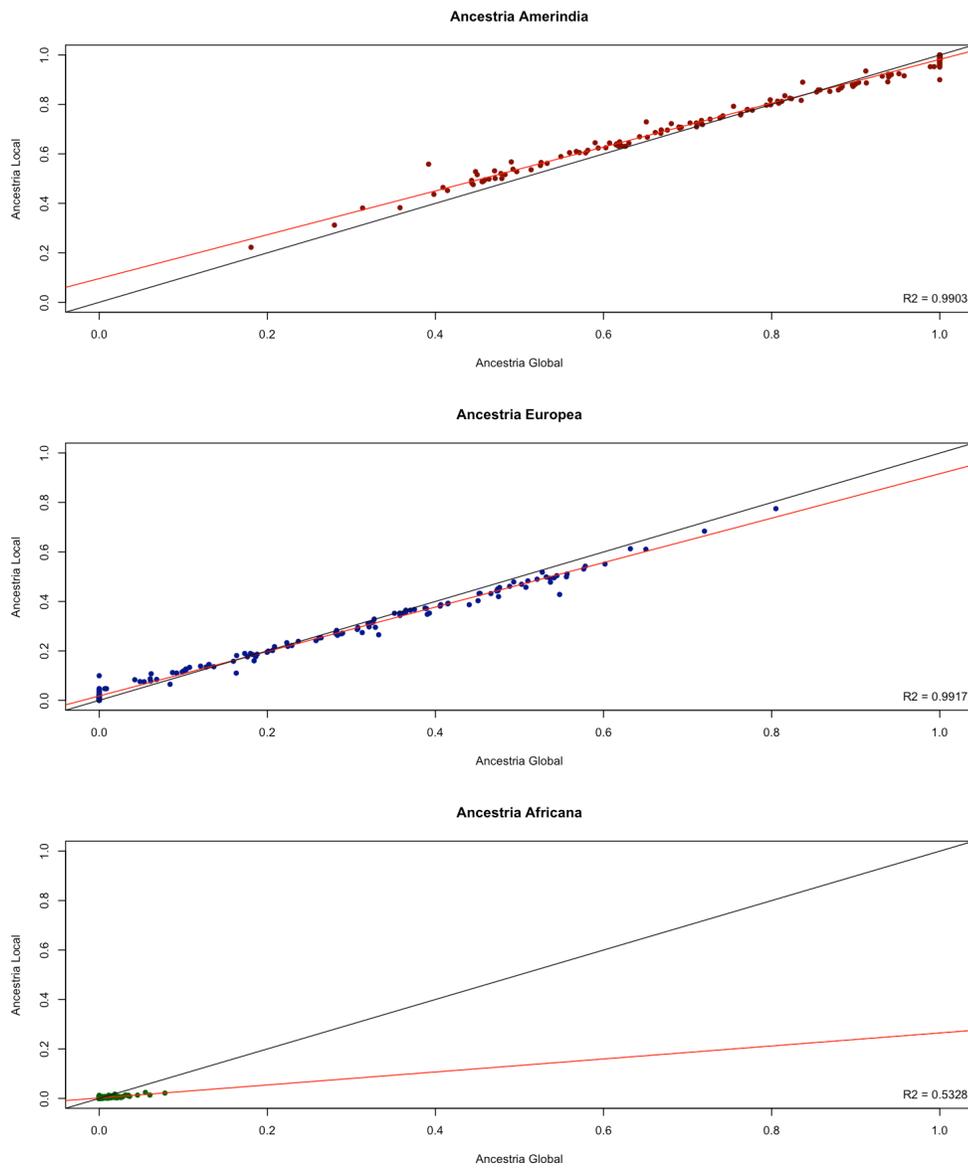


Figura 19. Test de Correlación para la inferencia de Ancestría Global realizado por ADMIXTURE y RFMix en las poblaciones de la Patagonia. En la figura superior se muestra el test de correlación para la Ancestría Amerindia, en medio la ancestría Europea y en la figura inferior la Ancestría Africana.

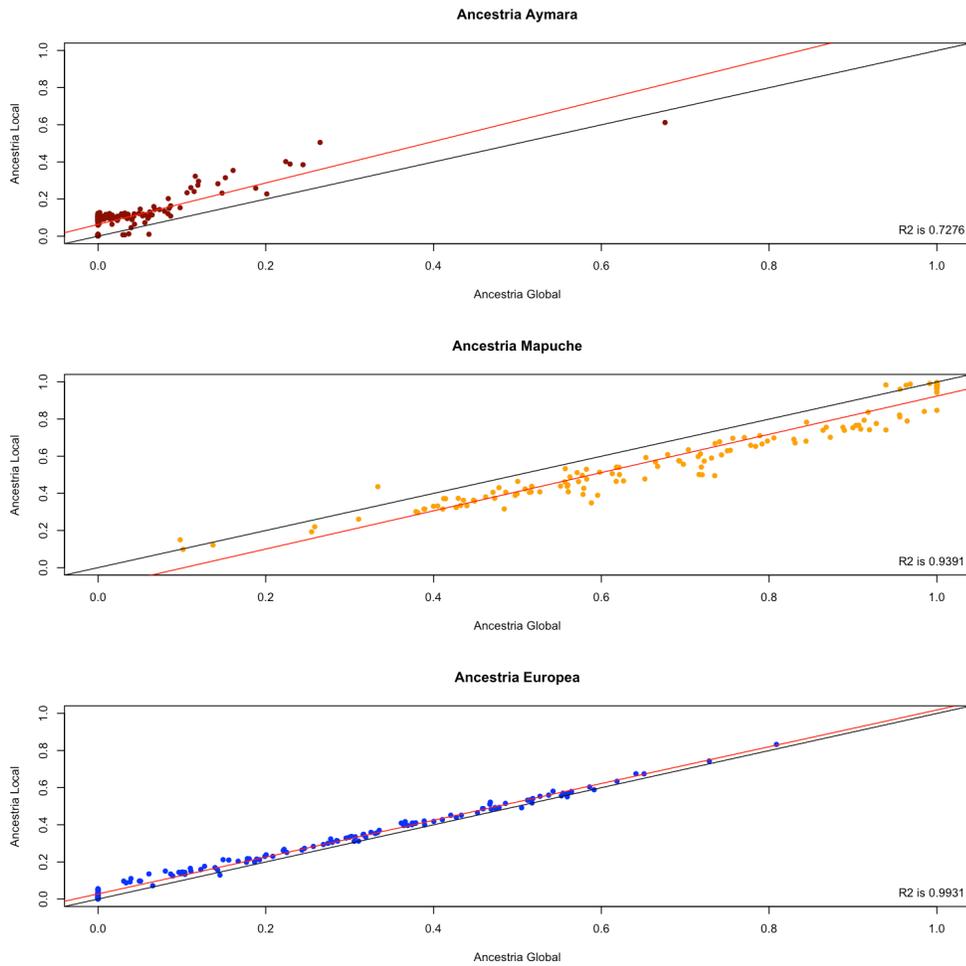


Figura 20. Test de Correlación para la inferencia de Ancestría Global realizado por ADMIXTURE y la ancestría local realizado por RFMix en las poblaciones de la Patagonia. En la figura superior se muestra el test de correlación para la Ancestría Mapuche, en la figura del medio la ancestría Aymara y en la figura inferior la Ancestría Africana.

5.3.4. Modelamiento de los procesos de mestizaje en la Patagonia

Se analizaron las longitudes de cada segmento cromosómico con ancestría conocida (“tracts”) con la finalidad de determinar diversos parámetros de las poblaciones de la Patagonia: tiempo inicial de mezcla, proporciones iniciales para cada ancestría, y sus respectivos cambios debido a eventos de migración. Para esto se pusieron a prueba dos modelos de migración, y se evaluó mediante likelihood el escenario que mas se ajusta a nuestros datos (Tabla V). Por un lado se evaluó un

modelo que comprende una mezcla inicial entre Europeos y Nativos Americanos, con un posterior pulso (migración) Africano (**EUR,NAT+AFR**), mientras que el otro modelo añade un segundo pulso de origen Europeo (**EUR,NAT+AFR+EUR**). Como se aprecia en los valores de likelihood de la **Tabla IV**, el modelo que mas se ajustó para los Chono y Yámana fue el de **EUR,NAT+AFR**, mientras que para las poblaciones Tehuelches y Kawésqar, el modelo con un segundo pulso Europeo (**EUR,NAT,AFR+AFR+EUR**) fue el que mejor se ajustó con los datos observados.

Tabla V. Cuadro con los valores de Likelihood para cada modelo de migración. G corresponde al tiempo inicial de mezcla en generaciones.

	Migrations Models			
	EUR,NAT+AFR		EUR,NAT+AFR+EUR	
	log Likelihood	G*	log Likelihood	G*
Chono	-244	10	-245	10
Tehuelches	-223	9	-212	11
Kawesqar	-140	9	-137	11
Yámana	-370	9	-373	11

5.3.4.1. Componentes Europeo, Nativo y Africano

Como podemos ver a partir de la Tabla V, los tiempos de mezcla inicial van entre un rango de 9 a 12 generaciones, los cuales corresponden aproximadamente hace 225 y 300 años atrás (considerando 25 años por generación). En la **Figura 21** podemos apreciar los tiempos de mestizaje y la distribución de los tracts para cada población, de acuerdo al modelo de migración específico para cada uno de estos. Cabe recalcar que para ningún escenario es posible constatar una contribución importante del pulso Africano posterior a la mezcla inicial, manteniéndose proporciones inferiores al 0,5%. Esto induce que las proporciones de los principales componentes ancestrales (Amerindio y Europeo) se mantengan constantes y sin significativos cambios a lo largo del tiempo. Esta afirmación no incluye a las poblaciones Tehuelches y Kawésqar, las cuales presentan un modelo de pulso

Europeo tardío y que a diferencia del pulso Africano, si se refleja en un cambio en las proporciones ancestrales.

La población de Laitec ascendentes de Chono, ubicada al sur de la Isla de Chiloé, presenta un tiempo de mezcla de 10 generaciones atrás con unas proporciones iniciales del 80% para la ancestría Nativa Americana, convirtiéndose en la más alta en las poblaciones de la Patagonia Chilena. Para el caso de los Tehuelches en la Patagonia Argentina podemos inferir una contribución inicial del 86% para el componente Amerindio, frente a un 14% del Europeo hace 11 generaciones. Sin embargo, el segundo pulso Europeo el cual ocurrió hace 7 generaciones aproximadamente (175 años aprox.) contribuyó con un aumento del 8% para este componente. Algo muy similar ocurre con los individuos Kawéskar, los cuales presentan inicialmente proporciones del 76% para el componente Amerindio y un 24% para el componente Europeo los cuales aumenta en un 3 y 16% hace 6 y 5 generaciones atrás respectivamente mediante un segundo pulso. Por último, la población Yámana presenta una mezcla inicial del 70% para la ancestría Nativa Americana y un 30% para la Europea hace 9 generaciones atrás aproximadamente.

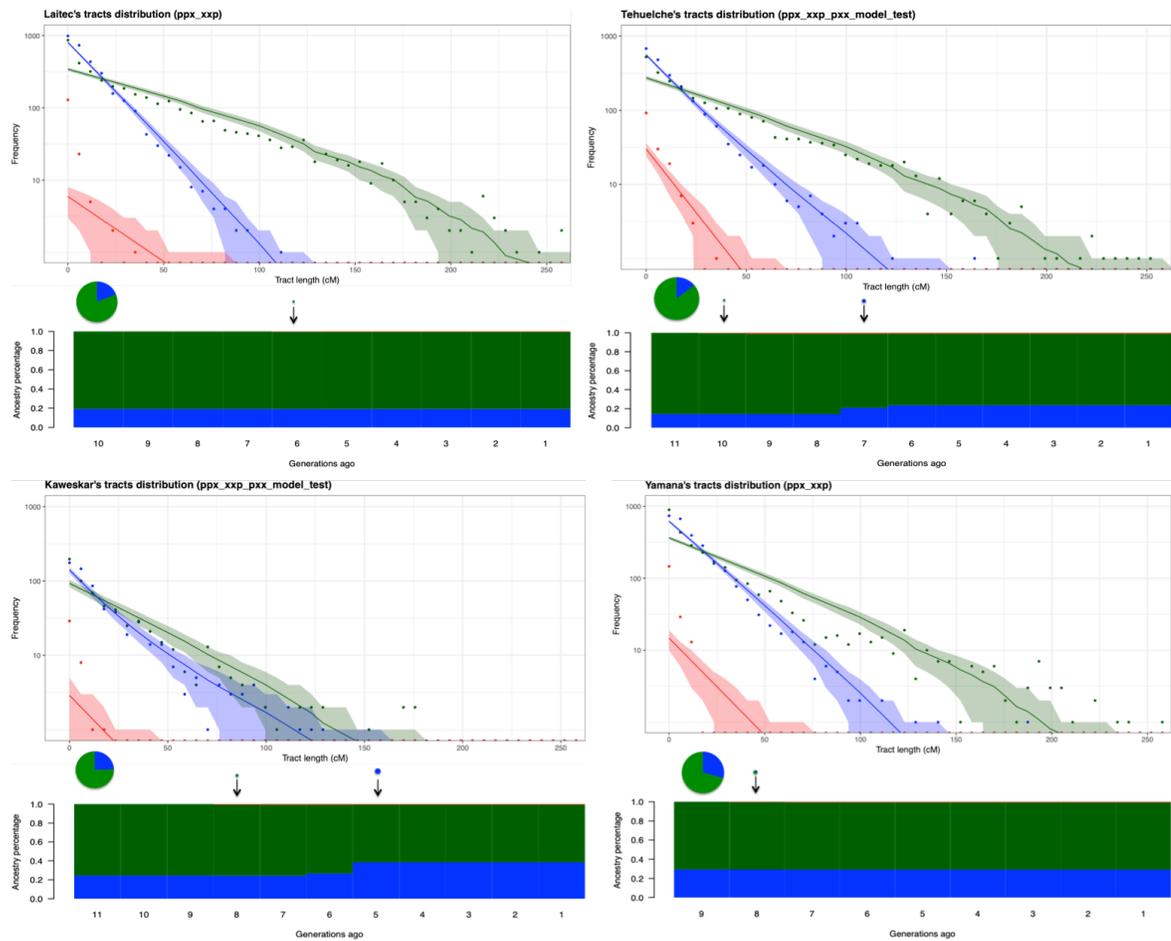


Figura 21. Análisis de TRACTs y tiempos de mestizaje en las poblaciones de la Patagonia. Se muestran los modelos de TRACTs que mejor se ajustan para los datos observados en las poblaciones Chono, Tehuelches, Kawéskar y Yámana. El decaimiento de la ancestría (arriba) compara la distribución de las longitudes de tracts observados y esperados basado en el modelo de migración propuesto. El área sombreada corresponde al 68,7% del intervalo de confianza. Los puntos indican los valores observados. El modelo de migración (parte inferior del gráfico) muestra el cambio en las proporciones de ancestría en el tiempo. De izquierda a derecha se observa la escala en generaciones, los círculos en la parte superior indican la ocurrencia de los eventos de mezcla, siendo el tamaño del círculo la magnitud de los migrantes que llegaron de cada población ancestral. Para las poblaciones Tehuelches y Kawéskar se mostraron modelos con un segundo pulso Europeo (ppx_xxp_pxx). Los colores verde, azul y rojo representan a las ancestrías Mapuche, Europeo y Africano respectivamente.

5.3.4.2. Componentes Aymara, Mapuche y Europeos

Se evaluaron los componentes ancestrales en las poblaciones de la Patagonia previo al contacto Europeo (Figura 22). Para esto se descompuso el componente

Amerindio en las ancestrías Aymara y Mapuches (Pehuenche-Huilliche) de acuerdo a lo visto en los análisis de ancestría global, dentro del cual se evidenció un importante aporte de estas ancestrías en las poblaciones de la Patagonia, especialmente en las de la Patagonia Austral (Kawéskar y Yámana). Dado a que el análisis de ancestría local no permite mas de tres grupos de referencia, se removió al grupo Africano dada su baja contribución, y se reemplazó por el grupo de referencia Aymara (Puno). A diferencia del análisis anterior, no se compararon dos modelos de migración debido a que el enfoque está orientado a comparar las frecuencias entre los componentes amerindios antes del pulso Europeo. Del análisis se desprende una importante participación del componente Aymara en los grupos Kawéskar con un 59% hace 11 generaciones atrás, seguido inmediatamente por un pulso Europeo de aproximadamente un 30%. El resto de las localidades presenta una composición entre el componente Aymara y Mapuche de aproximadamente un 25%, remontándose hace aproximadamente 12, 14 y 16 generaciones para los Chono, Tehuelches y Yámana respectivamente.

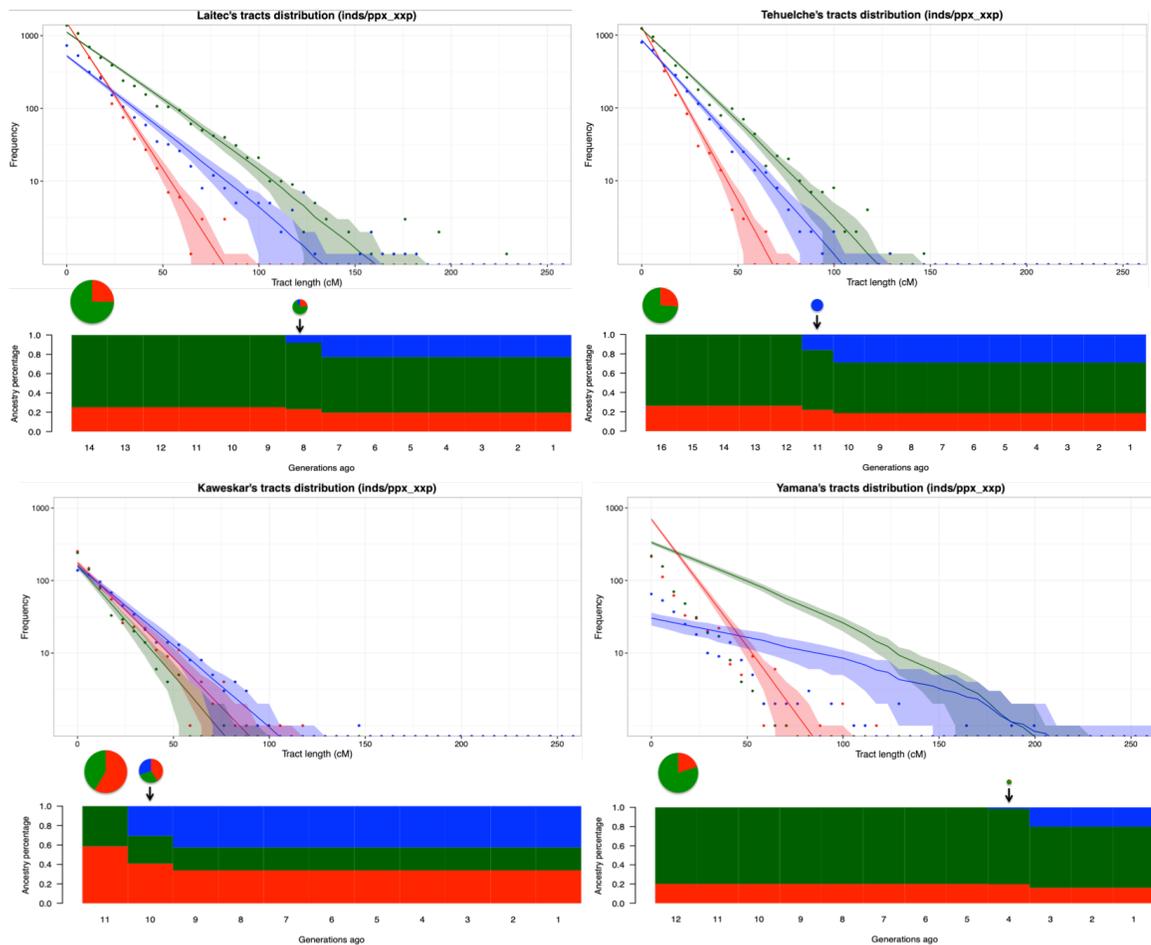


Figura 22. Análisis de TRACTs y tiempos de mestizaje en las poblaciones de la Patagonia para las ancestrías Aymara, Mapuche y Europa. Se muestran los modelos de TRACTs que mejor se ajustan para los datos observados en las poblaciones Chono, Tehuelche, Kawéskar y Yámana. El decaimiento de la ancestría (arriba) compara la distribución de las longitudes de tracts observados y esperados basado en el modelo de migración propuesto. El área sombreada corresponde al 68,7% del intervalo de confianza. Los puntos indican los valores observados. El modelo de migración (parte inferior del gráfico) muestra el cambio en las proporciones de ancestría en el tiempo. De izquierda a derecha se observa la escala en generaciones, los círculos en la parte superior indican la ocurrencia de los eventos de mezcla, siendo el tamaño del círculo la magnitud de los migrantes que llegaron de cada población ancestral. Los colores verde, azul y rojo representan a las ancestrías Mapuche, Aymara y Africano respectivamente.

5.4. Enmascaramiento del Genoma no Nativo

Dado que los resultados muestran una clara evidencia de contribución Europea en los individuos de la Patagonia, se procedió a enmascarar todas aquellas regiones

del genoma de origen No Nativo. Esto posee una alta relevancia al momento de analizar señales de selección sobre el genoma de individuos de la Patagonia, puesto que de haber regiones bajo presión selectiva, estos tienen que encontrarse en variantes de origen Nativo. Este paso se hace especialmente necesario para los análisis de Population Branch Statistic (PBS) e integrated Haplotype Score (iHS), exceptuando el Mapeo Genómico de Ancestría. El proceso general se encuentra descrito en la **Figura 23** y se realiza separadamente para cada individuo. Este método parte con un archivo faseado en formato *VCF* (**Figura 23A**), posteriormente se procedió a separar el set de cromosomas maternos y paternos (**Figura 23B**). Asociado a esto, se utilizó la información entregada por la ancestría local para cada individuo, donde como podemos ver, esta entrega la ancestría para una región del cromosoma en específico (rectángulos verdes). Esto se usó para obtener dentro de que rango se ubicada cada posición del SNP y la ancestría que posee. En la **Figura 23C** vemos para haplotipo el resultado de la intersección, para el haplotipo A vemos que los SNPs están asociados a una ancestría Amerindia (Patagonia), mientras que para el otro haplotipo los SNPs se encuentran asociados al componente Europeo. Después de obtener la ancestría para cada SNP, se procedió a enmascarar mediante datos perdidos (".") todos los SNPs que tuvieran ancestría Europea, Africana o desconocida (RFMix no siempre puede determinar la ancestría para todos los SNPs). Esto se realizó por separado para cada set de cromosomas materno y paterno para luego ser fusionado finalmente, obteniendo un archivo en formato *VCF* faseado y solo con aquellos SNPs que tengan una ancestría Amerindia (**Figura 23D**).

El ejercicio implementado en esta sección se aplicó para todos los individuos de la Patagonia, y como resultado final podemos ver que existe una gran cantidad de datos perdidos a lo largo del genoma para cada individuo los cuales se tomarán en cuenta para los análisis posteriores que requieran de datos enmascarados.

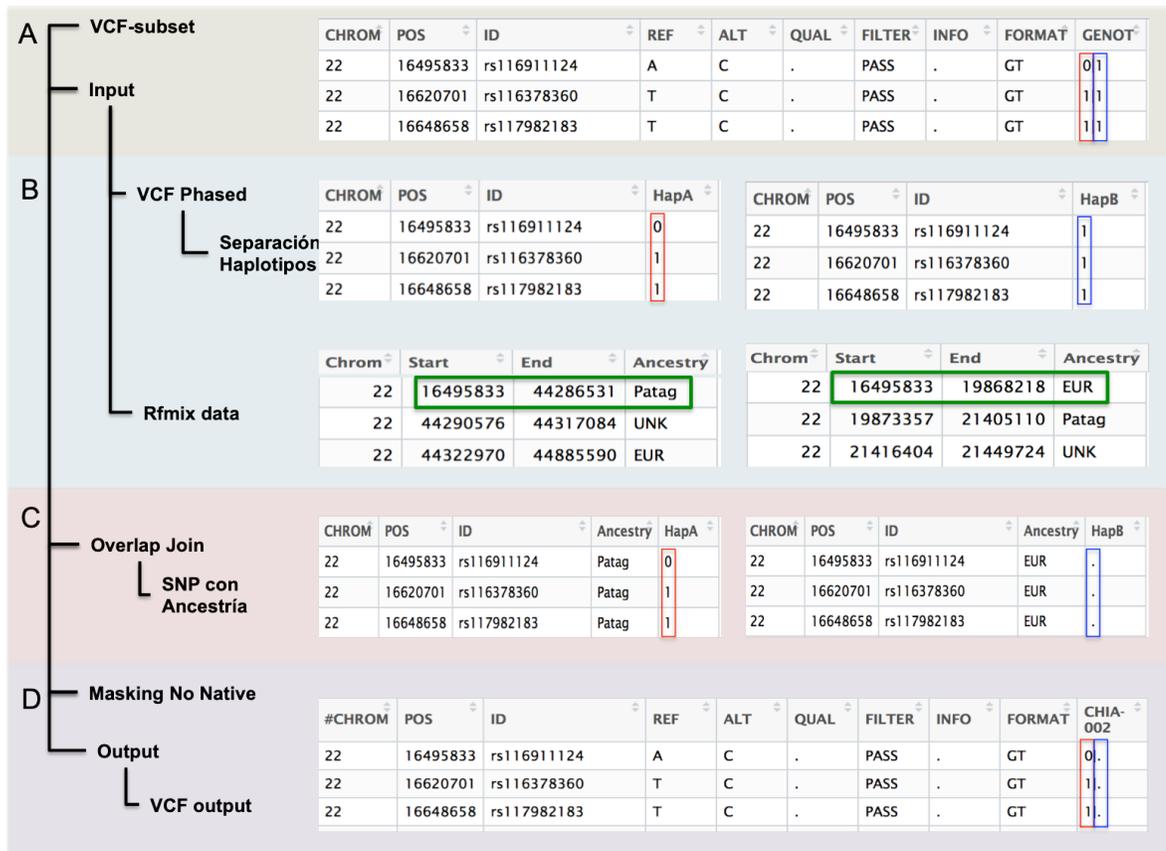


Figura 23. Esquema del proceso de enmascaramiento de regiones No Nativas del genoma. En la figura solo se muestra la información para un individuo y únicamente para el cromosoma 22. Los rectángulos rojos y azules señalan el set de SNPs para un juego cromosómico distinto (eg. Materno y paterno). Los rectángulos verdes indican la ancestría para un rango o región cromosómica específica.

5.5. Análisis de Selección

5.5.1. Señales de Selección pre-contacto Europeo

5.5.1.1. Population Branch Statistic (PBS)

Se realizó el análisis de PBS para dos escenarios posibles; A) uno con las poblaciones de la Patagonia (Chono, Tehuelches, Kawéskar y Yámana), el cual se contrastó con los Pehuenche y Huilliche, mientras que el otro escenario B) corresponde únicamente a las poblaciones cazadoras recolectoras marítimas (Chono, Kawéskar y Yámana), utilizando esta vez a los individuos Tehuelche como grupo de contraste. Para ambos escenarios se utilizó como grupo externo a los Europeos (IBS), además se removieron aquellos SNPs con una ancestría de origen No Nativo. En la Tabla VI se muestra un resumen con el número de SNPs y ventanas utilizadas, además de los marcadores por sobre el 1% para ambos escenarios. Por otro lado, en la Figura 24 vemos la intersección de los genes que se encuentran bajo selección para los dos posibles escenarios.

Tabla VI. Número de SNPs, ventanas y Genes analizadas para los escenarios A y B junto con los marcadores sobre el 1% de la distribución de los datos.

	SNPs				Ventanas (200 kb)				
	Nº SNPs	Nº Genes	Nº SNPs top1%	Nº Genes top1%	Nº Ventanas 200kb	Nº Genes	Nº Ventanas > 20 SNPs	Prop SNPs > 1% en Ventanas	Nº Genes
Escenario A (IBS;PEHHUI;CR) ¹	404430	21865	4046	2729	13317	20126	9080	94	227
Escenario B (IBS;TEH;CRM) ²	387327	21745	3875	2720	13313	20123	8732	94	236

1. **Escenario A:** Europa (IBS), Pehuenche-Huilliche (PEHHUI) y Cazadores Recolectores de la Patagonia (CR).

2. **Escenario B:** Europa (IBS), Tehuelches (TEH) y Cazadores Recolectores Marítimos de la Patagonia (CRM).

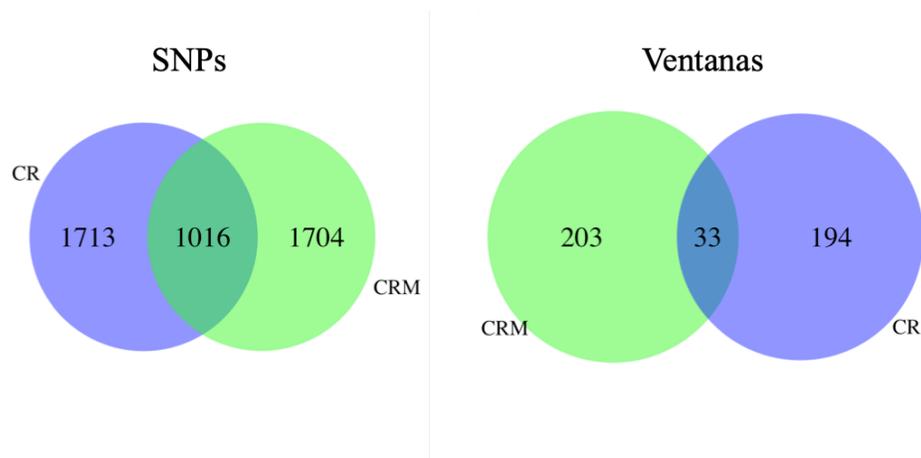


Figura 24. Intersección de genes encontrados bajo selección (top 1%) para los análisis realizados por SNPs (izquierda) y por ventanas de 200Kb (derecha). CR (cazadores recolectores) corresponde al escenario A, compuesto por los grupos Chono, Tehuelche, Kawéskar y Yámana, mientras que CRM (cazadores recolectores marítimos) corresponde a los grupos Chono, Kawéskar y Yámana.

La Figura 25 muestra bajo el primer escenario, un total de 4.046 SNPs con valores de PBS por sobre el 1% de la distribución de los datos. Los valores más elevados se pueden ver en los cromosomas 16, 15, 10 y 2. En la Tabla VII podemos ver el top 10 de aquellos marcadores por sobre el 1% de los datos, de estos solo cinco se encuentran en genes descritos, tres en loci sin un gen asociados y dos sin información. Dentro de los genes que se encuentran descritos en la base de datos³ podemos destacar a **ASXL2** el cual codifica un miembro de reguladores epigenéticos que permiten la unión de enzimas modificadores de histonas. Este gen se encuentra asociado a cáncer de varios tejidos (mamas, páncreas, ovario, próstata y sangre), sin embargo este gen también juega un rol importante en la adipogénesis, el cual esta relacionado con cualquier proceso que active o aumente la frecuencia, tasa o grado de diferenciación de adipocitos⁴. El resto de los genes se encuentra relacionado con supresión de tumores, factores de expresión, metabolización del ácido hialurónico y regulación en la actina del citoesqueleto.

³ Los genes para cada SNP fueron obtenidos de <https://www.ncbi.nlm.nih.gov/snp/>

⁴ El proceso biológico es obtuvo a partir de <http://amigo.geneontology.org/amigo/term/GO:0045600>

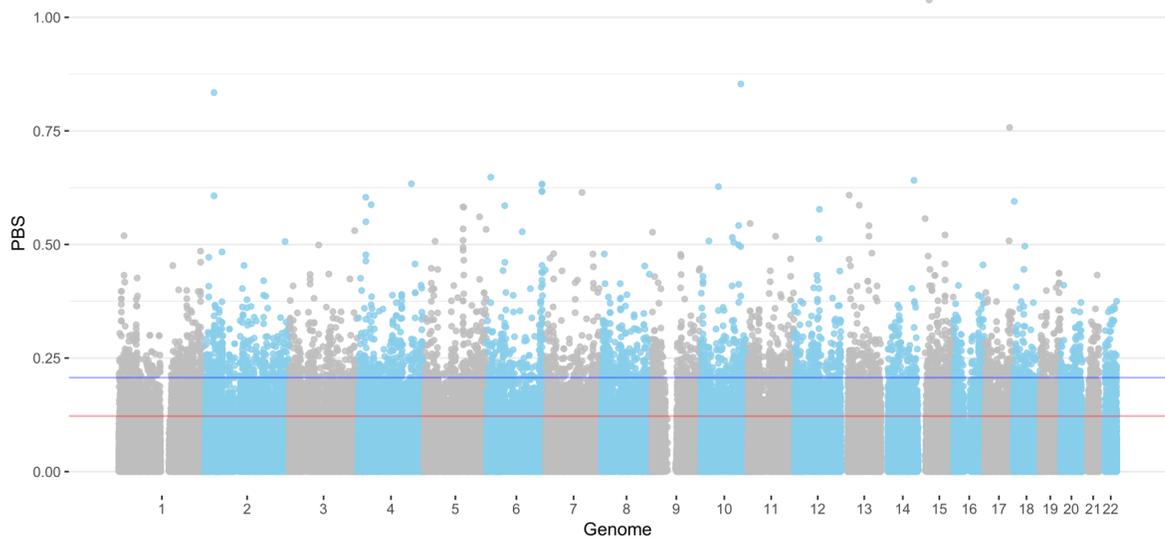


Figura 25. Manhattan Plot para los valores de PBS sobre las poblaciones de la Patagonia para el escenario A (IBS; Pehuenche/Huilliche; Cazadores Recolectores). Las líneas horizontales roja y azul representan el 5 y 1% del percentil respectivamente sobre la distribución de los datos a lo largo del genoma.

Tabla VII. Top 10 de los valores de PBS sobre el 1% de los datos observados para el escenario A por SNPs.

Ranking	Cromosoma	Posición	rsID	PBS	Genes
1	16	75822042	rs2866065	1.0499290	LOC105371348
2	15	32099353	rs2221329	1.0381440	OTUD7A
3	10	115326193	rs3850690	0.8533093	HABP2
4	2	26074547	rs74261335	0.8342096	ASXL2
5	17	71294754	rs904387	0.7574286	CDC42EP4
6	6	11613708	rs542807	0.6481946	Sin Información
7	14	95463806	rs7160237	0.6411432	Sin Información
8	4	154478608	rs12506179	0.6335318	TMEM131L
9	6	159475468	rs2073137	0.6327482	LOC105378083
10	6	159481653	rs12663082	0.6327482	LOC105378083, LOC112267968

Se registraron un total de 13316 ventanas de 200 kb a lo largo del genoma para el escenario A, de las cuales 9080 poseen más de 20 SNPs **Tabla VI**. De estas últimas filtradas, se evaluó la proporción de ventanas con marcadores sobre el 1% de la distribución de los datos a nivel genómico. El umbral azul representa el puntaje de corte del 1% sobre la distribución de las proporciones de las ventanas, indicando

un total de 94 que sobrepasan ese valor (Figura 26). A partir de esto podemos observar que los peaks mas altos corresponden a los cromosomas 11, 6 y 4. En la **Tabla VIII** se encuentra el top 10 de ventanas de 200 Kb con una proporción de SNPs con valores de PBS mayor al 1%. Los genes que se encontraron dentro del top 10 se buscaron en la base de datos ⁵, donde destacan principalmente la presencia de dos genes asociados con el metabolismo de lípidos (**ZNF202**, **LRP2**) y uno asociado con la metabolización de la molécula del colesterol (**GRAMD1B**). Por otro lado, se identificaron dos genes que codifican para proteínas transmembrana en la mitocondria y asociados con el transporte de moléculas y ATP (**ABCB10** y **ABCB11**). Por último, el gen **NUP133** cobra especial interés dado que codifica para una nucleoporina involucrada con varios procesos biológicos, entre estos controla la regulación a la respuesta celular al calor y regulación de procesos glicolíticos.

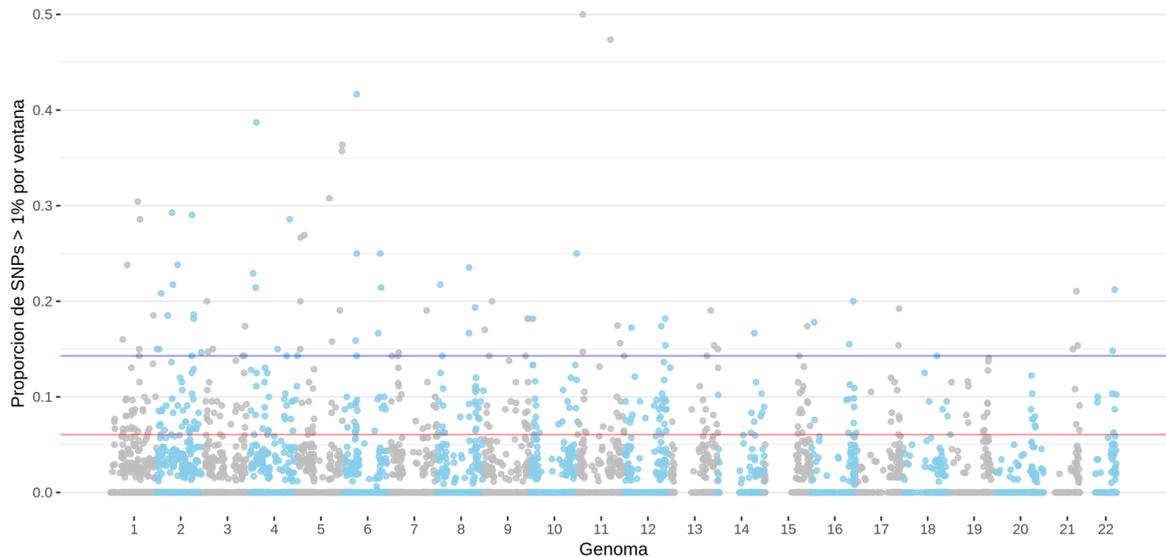


Figura 26. Proporción de SNPs con valores de PBS sobre el 1% en ventanas de 200kb sobre las poblaciones de la Patagonia para el escenario A (IBS; Pehuenche/Huilliche; Cazadores Recolectores). Las líneas horizontales roja y azul representan el 5 y 1% del percentil respectivamente sobre la distribución de los datos a lo largo del genoma.

⁵ Los genes dentro de cada rango dentro del cromosoma fueron obtenidos de <https://www.ncbi.nlm.nih.gov/gene>

Tabla VIII. Top 10 de los valores de PBS sobre el 1% de los datos observados para el escenario A mediante ventanas de 200Kb.

Ranking	Cromosoma	Inicio	Término	Prop SNP > 1%	Genes
1	11	123600000	123800000	0.5	SCN3B, ZNF202 , GRAMD1B , OR6X1, RNU1-21P, LOC105369544, LOC105369543
2	11	32400000	32600000	0.4736845	WT1, LOC107982234, EIF3M, WT1-AS, LOC106707172, HNRNPA3P9, LOC107984351
3	6	159200000	159400000	0.4166666	FND1, LINC02529, LOC105378084
4	4	126400000	126600000	0.3870967	RBM48P1, LOC105377411
5	5	89800000	90000000	0.3636363	LOC102724637
6	5	88000000	88200000	0.3571428	TMEM161B, LOC105379070
7	5	27800000	28000000	0.3076923	Sin Información
8	1	229400000	229600000	0.3043478	ACTA1, NUP133 , ABCB10 , TAF5L, HMG2P19, RNU4-21P, RN7SKP276, RNA5SP78, LOC101927478
9	2	169000000	169200000	0.2926829	ABCB11 , LRP2 , DHRS9, UBE2V1P6
10	2	40800000	41000000	0.2903225	LOC105374497

Para el escenario B, se analizaron un total de 387326 SNPs, de los cuales 3875 poseen un valor de PBS que se encuentra por sobre el 1% de los datos (Tabla VI). En la Figura 27 se aprecia el Manhattan plot con la distribución de los valores de PBS para cada SNP a lo largo del genoma. Se puede apreciar que a diferencia del escenario A, existen otros peaks de marcadores principalmente en los cromosomas 19, 8, 1 y 3. Esta información se encuentra detallada en la Tabla IX con el top 10 de los SNPs con mayor puntaje de PBS, de los cuales vale la pena destacar al gene **GFRAL** el cual codifica para un receptor de la familia GDNF y que participa en la reducción de la ingesta de nutrientes en respuesta a un exceso en la dieta. Otro gen de relevancia para nuestros datos corresponde al gen **WNT9B**, el cual modela la respuesta celular a la falta de nutrientes y sales minerales, así como también hacia el ácido retinoico. Este metabolito (Vitamina A) y regulador de transcripción genética ayuda a la reducción en la masa grasa y mejora la sensibilidad de la insulina (Okla et al., 2017).

El resto de los genes descritos se encuentran relacionados con respuesta a estímulo de luz, percepción visual, unión de iones de zinc, etc.

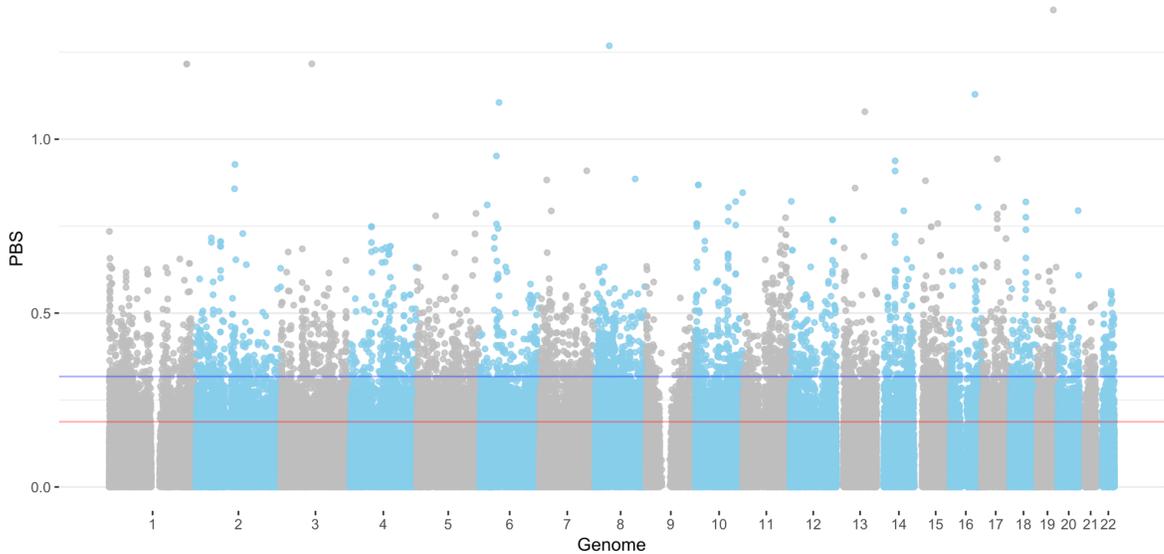


Figura 27. Manhattan Plot para los valores de PBS sobre las poblaciones de la Patagonia para el escenario B (IBS; Tehuelche; Cazadores Recolectores Marítimos). Las líneas horizontales roja y azul representan el 5 y 1% del percentil respectivamente sobre la distribución de los datos a lo largo del genoma.

Tabla IX. Top 10 de los valores de PBS sobre el 1% de los datos observados para el escenario B mediante el análisis de SNPs.

Ranking	Cromosoma	Posición	rsID	PBS	Genes
1	19	46480407	rs4239537	1,37143	Sin Información
2	8	40685227	rs16890007	1,268572	ZMAT4
3	3	88625185	rs77196386	1,216636	Sin Información
4	1	222598404	rs35452783	1,216008299	Sin Información
5	1	222620067	rs17163130	1,216008299	Sin Información
6	16	71332135	rs16971012	1,12881	LOC102723786
7	6	55194771	rs17756843	1,105559	GFRAL
8	13	80874400	rs1952666	1,078900159	Sin Información
9	6	47769943	rs1410830	0,9515249	OPN5
10	17	44921929	rs2083797	0,9432496	WNT9B, LRRC37A2

Cuando se analizaron las ventanas de 200 kb para el escenario B pudimos ver que de un total de 13313 ventanas, 8732 poseían sobre 20 SNPs. En la Figura 28 vemos la proporción de ventanas (>20 SNPs) con marcadores sobre el 1% de la

distribución de los datos a nivel genómico. La línea representa el puntaje de corte del 1% sobre la distribución de las proporciones de las ventanas, indicando un total de 94 que sobrepasan este valor. En la Tabla X se muestra el top 10 de las ventanas con mayor proporción de SNPs con valores de PBS >1%. Los peaks mas altos corresponden a los cromosomas 2, 17 y 11. De estos, resulta interesante la repetición del cromosoma 11 con la comparación del escenario A, el cual arrojó la misma región y con los mismos genes, de manera que nuevamente aparecen **ZNF202** y **GRAMD1B** relacionados con el metabolismo de lípidos y la metabolización de la molécula de colesterol.

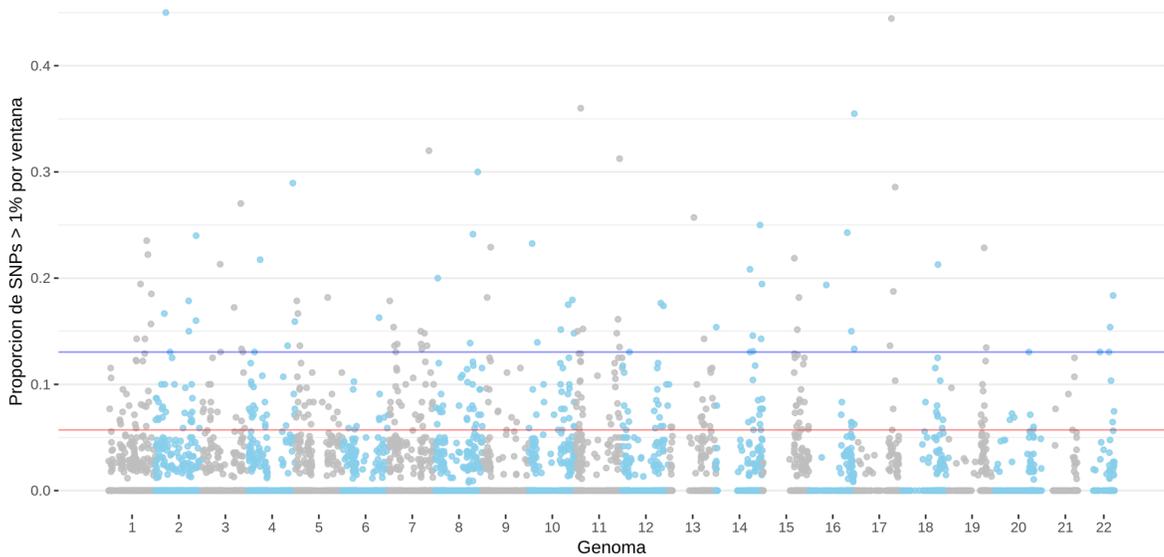


Figura 28. Proporción de SNPs con valores de PBS sobre el 1% en ventanas de 200kb sobre las poblaciones de la Patagonia para el escenario B (IBS; Tehuelche; Cazadores Recolectores Marítimos). Las líneas horizontales roja y azul representan el 5 y 1% del percentil respectivamente sobre la distribución de los datos a lo largo del genoma.

Tabla X. Top 10 de los valores de PBS sobre el 1% de los datos observados para el escenario B mediante el análisis de ventanas de 200Kb

Ranking	Cromosoma	Inicio	Término	Prop SNP > 1%	Genes
1	2	149400000	149600000	0.45	MMADHC,LYPD6,LOC101929231, RPL17P13
2	17	44800000	45000000	0.4444444	GFAP, EFTUD2, GJC1, CCDC103, KIF18B, HIGD1B, C1QL1, FAM187A, MIR6783, RN7SL405P, LOC112268183, LOC107987243, LOC105371793
3	11	123600000	123800000	0.36	SCN3B, ZNF202 , GRAMD1B , OR6X1, RNU1-21P, LOC105369544, LOC105369543
4	16	89800000	90000000	0.3548387	MC1R, TUBB3, FANCA, TCF25, SPIRE2, AFG3L1P, DEF8, CENPBD1, SNORA119, LOC112268179, LOC107984817, LOC105371419
5	7	67200000	67400000	0.32	TYW1, PMS2P4, STAG3L4, SPDYE21P, LINC01372, LOC105375337
6	11	86000000	86200000	0.3125	PICALM, FNTAP1, LINC02695, RNU6-560P
7	8	77600000	77800000	0.3	Sin Información
8	4	86800000	87000000	0.2894736	PTPN13, AFF1, SLC10A6, C4orf36, AFF-AS1, RPL6P13, TECRP1, LOC260422
9	17	63200000	63400000	0.2857142	TANC2, LOC105371857, LOC105371856, LOC100420064
10	3	61000000	61200000	0.2702702	FHIT

5.5.1.2. integrated Haplotype Score (iHS)

Se evaluaron las señales de selección basado en desequilibrio de ligamiento mediante iHS para cada SNP y por ventanas de 200 kb. Este análisis se llevo a cabo para los grupos de cazadores recolectores con las mismas agrupaciones efectuadas anteriormente; a) grupos cazadores recolectores (Chono, Tehuelche, Kawéskar y Yámana), y b) cazadores recolectores marítimos (Chono, Kawéskar y Yámana). Dado que este análisis toma en cuenta aquellos SNPs en desequilibrio de ligamiento, es de vital importancia que estos marcadores contiguos tengan el mismo origen ancestral –a diferencia de PBS-, además iHS no permite datos perdidos. Por esta razón, se procedió a maximizar la cantidad de haplotipos de origen Amerindio tomando en

cuenta los tamaños de la muestra (Figura 29). Para esto se aplicó la tasa de pérdida (*--mind, plink*) en diferentes porcentajes con el fin de remover aquellos individuos con menor componente Amerindio, luego para cada porcentaje se evaluó el número de loci sin datos perdidos en el nuevo subset de muestra.

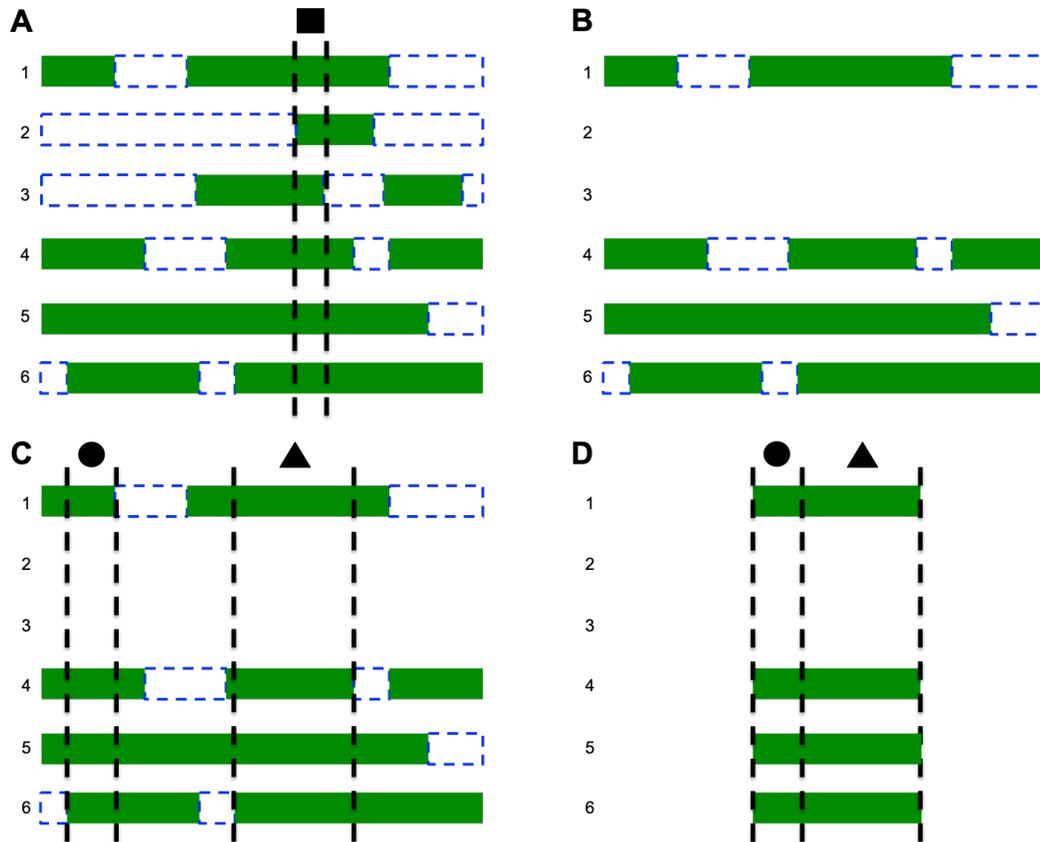


Figura 29. Representación del proceso de filtro de individuos y SNPs de origen Amerindio para los análisis de iHS. En A vemos seis haplotipos, donde los segmentos de color verde son de origen Amerindio, mientras que los segmentos con línea azul punteada corresponden a regiones de origen Europeo, y por lo tanto, se relleno estas regiones como “datos perdidos” en el proceso de Enmascaramiento. El cuadrado negro junto con las líneas negras verticales indican una región de origen Amerindio común a todos los haplotipos. En B se mantuvieron solo aquellos haplotipos con un componente Amerindio por sobre un umbral determinado. En C se identificaron aquellas regiones con marcadores Amerindios presentes en todo el nuevo subset filtrado (representados por un círculo y triángulo). En D se removieron todos aquellos SNPs que posean al menos un haplotipo sin información de genotipos. Finalmente se obtiene un nuevo subset de individuos con una menor cantidad de marcadores sin datos perdidos y de origen Amerindio.

Se realizaron diferentes filtros para los grupos cazadores recolectores de la Patagonia, estos valores se encuentran en la Figura 30. Antes de realizar los filtros

por componente Amerindio vemos que solo hay tres SNPs que poseen información para los 74 individuos que conforman el grupo de cazadores recolectores, el resto de los marcadores posee al menos en un individuo genotipos perdidos. Cuando removimos aquellos individuos con un componente Amerindio menor al 50%, el número de SNPs asciende a 144, de esta manera, a medida que vamos aumentando el umbral de componente Amerindio el número de SNPs aumenta en detrimento del número de individuos, el cual disminuye. Evaluamos la homogeneidad en la distribución de los SNPs a lo largo del genoma para dos rangos de componente Amerindio: 77 y 78% (Figura 31). Para esto, se generaron ventanas de 200Kb y se contaron cuantos marcadores habían en cada una. Como podemos ver, la frecuencia de las ventanas se encuentra distribuida hacia la izquierda con valores que van desde 0 a 100 SNPs por ventana, la mediana para los dos rangos corresponde a 36 marcadores. Tomando esto en cuenta, se eligieron en total 14 individuos, seis posibles Chono (Laitec), cinco Tehuelches, dos Yámana y un Kawéskar con al menos un 78% de componente Amerindio, quedándonos para este análisis con un total de 124.347 SNPs.

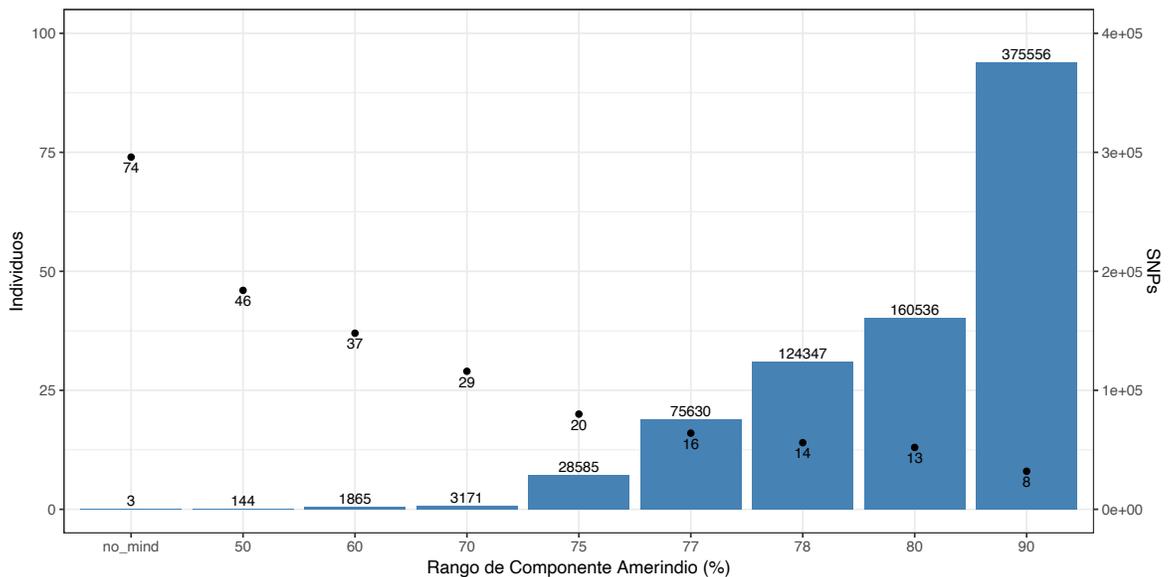


Figura 30. Número de individuos y SNPs para cada rango de componente Amerindio en grupos cazadores recolectores de la Patagonia. En el eje X se muestra los porcentajes de componente Amerindio utilizados, el S.A representa la ausencia de filtros por componente Amerindio (*sin mind*). El eje Y izquierdo muestra el numero de individuos (puntos), mientras que el eje Y derecho revela el número de SNPs (columnas).

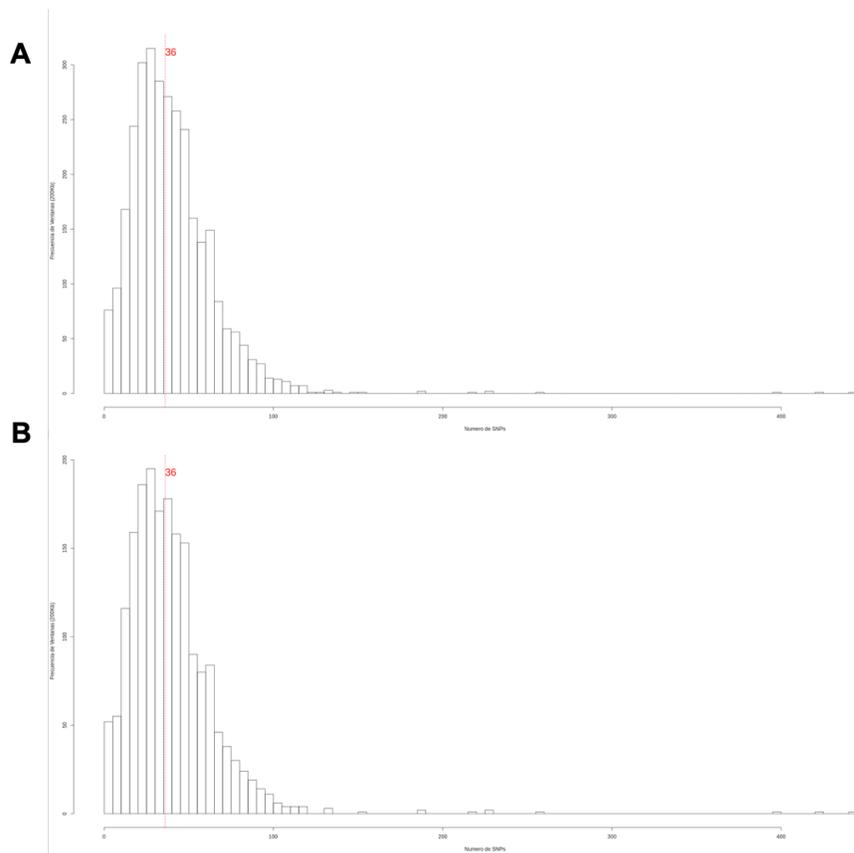


Figura 31. Frecuencia de SNPs por ventana de 200 Kb para diferentes rangos de Ancestría Amerindia. En el eje horizontal vemos el número de SNPs contenidos en cada ventana, mientras que en el eje vertical la frecuencia de ventanas. En A se observa la distribución de los marcadores a lo largo del genoma utilizando un rango de al menos un 77% de ancestría amerindia y en la figura inferior (B) se incluyeron aquellos individuos con al menos un 78% de ancestría amerindia. La línea vertical roja punteada indica la mediana para la distribución de los datos.

Una vez determinado el nuevo subset de individuos se realizó el análisis de iHS a nivel de SNPs y ventanas de 200 kb, el resumen de los filtros que se realizaron y el numero total de genes encontrados se encuentran en la Tabla XI. En la Figura 32 se muestra el análisis de iHS para los SNPs, donde podemos ver que existen varios SNPs con un valor de $|iHS| > 2.5$ (1274 SNPs), el cual coincide aproximadamente con el 1% de la distribución de los datos. Se puede ver que los valores mas extremos

corresponden a los cromosomas 5, 18, 15 y 16. Esto se puede ver mas claramente en la Tabla XII, donde destaca el gen **SLC39A10**, el cual pertenece a una subfamilia de proteínas con características estructurales de transportadores de Zinc. Este elemento es esencial para múltiples procesos, incluido al metabolismo de lípidos⁶. Además, cobra especial interés el gen **ADAM17** el cual codifica para una familia de proteínas de membrana e involucradas en muchos procesos biológicos, de los cuales se destacan la respuesta celular frente a estímulos de lipoproteínas de alta densidad y regulación negativa de termogénesis inducida por frío⁷. El resto de los genes esta asociado entre otras funciones al desarrollo de la fibra del músculo cardíaco, vasculogénesis, reguladores de transcripción, morfogénesis dendrítica, desarrollo del sistema nervioso, receptores de insulina, entre otros.

Tabla XI. Número de SNPs, Ventanas y Genes analizados para iHS en poblaciones cazadores recolectores de la Patagonia.

	SNPs				Ventanas				
	Nº SNPs	Nº Genes	Nº SNPs > 2.5	Nº Genes > 2.5	Nº Ventanas 200kb	Nº Genes > 20 SNPs	Prop SNPs (>2.5) > 1% en Ventanas	Nº Genes	
CR¹	64265	5097	1274	540	2730	4072	1442	15	55

CR: Cazadores Recolectores de la Patagonia (Chono, Tehuelches, Kawéskar y Yámana).

⁶ Extraído a partir de <https://www.ncbi.nlm.nih.gov/gene/57181>

⁷ Extraído a partir de <https://www.ncbi.nlm.nih.gov/gene/6868>

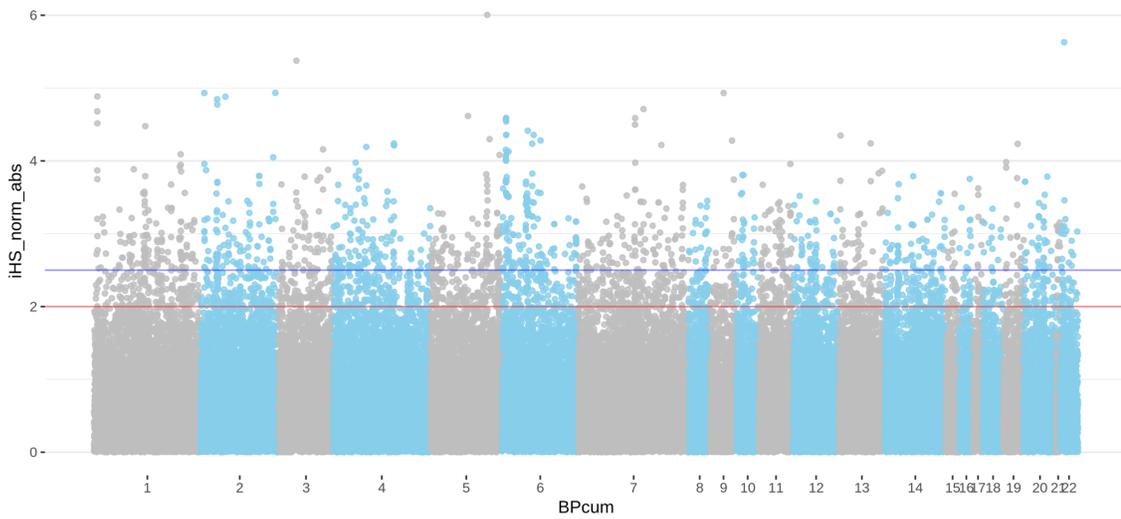


Figura 32. Valores de iHS por SNPs en poblaciones cazadores recolectores de la Patagonia. Los valores fueron normalizados y se muestra para todo el genoma. Las líneas horizontales roja y azul representan un puntaje de corte de $|iHS| > 2$ e $|iHS| > 2.5$ respectivamente.

Tabla XII. Top 10 de los valores de iHS sobre el 1% de los datos observados mediante el análisis de SNPs.

Ranking	Cromosoma	rsID	Posición	iHS norm	Genes
1	13	rs767932	101417627	6,35188	NALCN-AS1
2	5	rs156092	139518989	6,00407	Sin Información
3	22	rs4822653	26180538	5,62927	MYO18B
4	3	rs17069353	63755029	5,37573	LOC100130345
5	2	rs76324176	196500724	4,93424	SLC39A10
6	2	rs10495562	9631340	4,93264	IAH11, ADAM17
7	9	rs72699101	3555163	4,93249	Sin Información
8	1	rs17385519	8527520	4,8854	RERE
9	2	rs2888637	40127182	4,88168	LOC112268436
10	2	rs848516	36764402	4,84573	CRIM1

En el análisis de la Figura 33 se puede observar el Top 10 de las ventanas sobre el 1% de la distribución de los datos, los cuales se encuentran descritos en su totalidad en la Tabla XIII. De estos destacan cinco ventanas cercanas bajo selección ubicadas en el cromosoma 6, específicamente abarcando una región que va desde los 30-30,2Mb y desde 44.6Mb hasta los 45.4Mb. Algunos de los genes que se

encuentran en esta región están asociados con el Complejo Mayor de Histocompatibilidad (MHC).

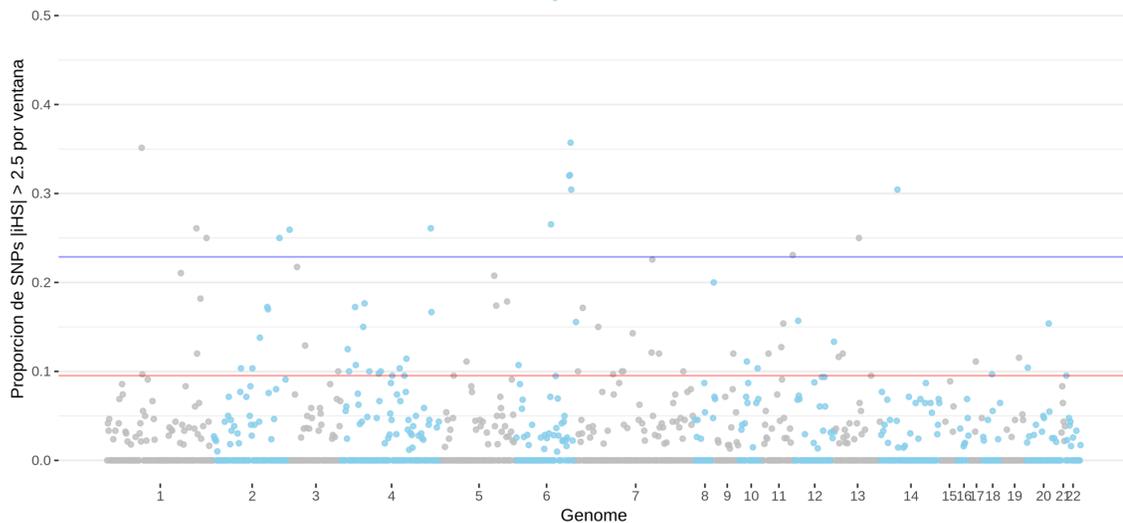


Figura 33. Proporción de SNPs con valores de $|iHS| > 2.5$ dentro de cada ventana de 200Kb en poblaciones cazadoras recolectoras de la Patagonia. Los valores fueron normalizados y se muestra para todo el genoma. Las líneas horizontales roja y azul representan el top 5 y 1% de los datos respectivamente.

Tabla XIII. Top 10 de los valores de iHS sobre el 1% de los datos observados mediante análisis de ventanas de 200Kb.

Ranking	Cromosoma	Inicio	Término	Prop SNP $ iHS > 2.5$	Genes
1	6	30000000	30200000	0,52	ZNRD1, TRIM26, TRIM31, PPP1R11, RNF39, ZNRD1ASP, TRIM15, TRIM10, TRIM40, HLA-J, ETF1P1, HCG4P3, TRIM-AS1, PAIP1P1
2	6	45000000	45200000	0,357142857	SUPT3H, MIR586, RBM22P4, LOC100420052
3	1	181600000	181800000	0,351351351	CACNA1E, RNA5SP70
4	6	44800000	45000000	0,320754717	SUPT3H, NUDT19P4, LOC107986600, LOC101929770
5	6	44600000	44800000	0,32	LOC101929770
6	6	45200000	45400000	0,304347826	RUNX2, SUPT3H, LOC109611593
7	14	32000000	32200000	0,304347826	ARHGAP5, ARHGAP5-AS1, LOC105370440
8	6	153000000	153200000	0,265306122	RGS17, MTRF1L, RNA5SP224
9	1	76200000	76400000	0,260869565	ST6GALNAC3, LOC100418965
10	4	86600000	86800000	0,260869565	PTPN13, LOC107986189

5.5.2. Señales de Selección post-contacto Europeo

5.5.2.1. Detección de Loci bajo Selección mediante Ancestría

Este tipo análisis nos permite visualizar la información de Ancestría Local desde otra perspectiva, dando una visión acerca de las frecuencias de ancestrías a lo largo del genoma sobre un grupo de individuos. En una primera instancia, se evaluó el mapeo genómico de ancestría en dos escenarios: a) cazadores recolectores de la Patagonia (Chono, Tehuelche, Kawéskar y Yámana), y b) cazadores recolectores marítimos de la Patagonia (Chono, Kawéskar y Yámana). En la Tabla XIV se aprecian los promedios, y sus respectivas desviaciones y errores estándar para las ancestrías Amerindia, Europea y Africana a lo largo del genoma. En términos generales vemos que para ambas comparaciones existe una mayor frecuencia de la ancestría Amerindia, con valores muy similares (75,6 y 76,2% respectivamente). La ancestría Europea se encuentra representada en un 23% aproximadamente en ambos grupos, mientras que el componente Africano esta casi ausente con un valor inferior al 0,5%. Por otro lado, en la Tabla XV vemos un resumen con el número de SNPs y genes analizados para cada escenario, así como también el proceso de filtro de aquellos marcadores por sobre el 1% de la distribución de los datos.

Tabla XIV. Resumen de los promedios de ancestría a nivel genómico para los grupos cazadores recolectores de la Patagonia y sus respectivas medidas de dispersión.

		Escenario A (LAI,TEH,KAW,YAM)	Escenario B (LAI,KAW,YAM)
Ancestría Amerindia	Media	0.7568805	0.7615024
	Desviación Estándar	0.0516117	0.0570262
	Error Estándar	0.0000703	0.0000777
Ancestría Europea	Media	0.2381943	0.2345993
	Desviación Estándar	0.0525292	0.0573877
	Error Estándar	0.0000716	0.0000782
Ancestría Africana	Media	0.0049250	0.0038982
	Desviación Estándar	0.0073183	0.0080750
	Error Estándar	0.0000997	0.0000110

En **Escenario A** (Chono, Tehuelche, Kawéskar y Yámana), En **Escenario B** (Chono, Kawéskar y Yámana)

Tabla XV. Resumen con el número de SNPs y genes para el análisis de Detección de Loci bajo Selección mediante Ancestrías.

	Nº SNPs	Nº Genes	Nº SNPs > top 1%	Genes > top 1% Nativos	Genes > top 1% Europeos
Escenario A	538040	22625	5518	200	331
Escenario B	538040	22676	7378	217	348

En **Escenario A** (Chono, Tehuelche, Kawéskar y Yámana), En **Escenario B** (Chono, Kawéskar y Yámana)

Para ambos escenarios se mostraron las fluctuaciones en porcentaje de las ancestrías Nativas y Europeas a lo largo del genoma. El componente Africano no se analizó dado que la representación en estas poblaciones es baja (0.4% aproximadamente). Con los grupos cazadores recolectores de la Patagonia (Figura 34, gráfico superior) vemos que la mayoría de las variantes que están por sobre el 1% de la distribución de los datos se encuentran concentradas en el cromosoma 6, conformando la extensión mas larga en el genoma, y en menor grado al cromosoma 21. Se registraron un total de 200 genes asociados a estos cromosomas, donde 194 corresponden al cromosoma 6, estando la mayoría asociados al complejo mayor de histocompatibilidad. Dentro de esta familia, los genes **MICA** y **MICB** cobran cierto interés, puesto que entre tantos procesos, permite regular la respuesta al calor y al ácido retinoico⁸. Por otro lado, también resulta interesante destacar el gen **LTA**, el cual corresponde a una citoquina producida por linfocitos. Esta proteína se encuentra involucrada en varios procesos metabólicos, de los cuales vale mencionar distintas respuestas frente a hipoxia y nutrientes⁹. En lo que respecta al metabolismo de lípidos, el gen **APOM** posee particular interés debido a que esta involucrado en varias funciones relacionadas con procesos metabólicos de lipoproteínas, respuesta a glucosa y al metabolismo del colesterol¹⁰. Para el cromosoma 21 encontramos un set de genes que principalmente están asociados con la respuesta inmune,

⁸ Extraído a partir de <https://www.ncbi.nlm.nih.gov/gene/100507436>

⁹ Extraído a partir de <https://www.ncbi.nlm.nih.gov/gene/4049>

¹⁰ Extraído a partir de <https://www.ncbi.nlm.nih.gov/gene/55937>

transporte transepitelial, organización mitocondrial, desarrollo del corazón, entre otros.

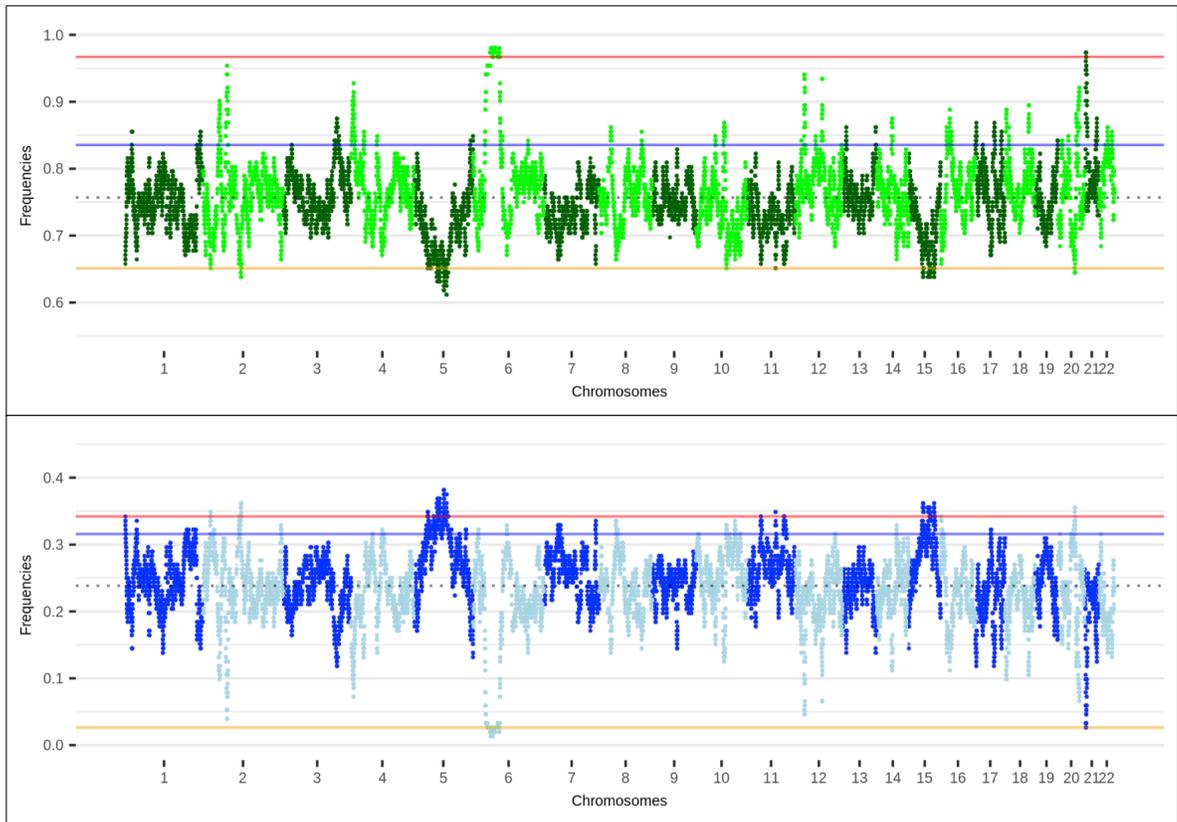


Figura 34. Detección de loci bajo selección a lo largo del Genoma basado en ancestrías para los grupos cazadores recolectores de la Patagonia (Chono, Tehuelche, Kawéskar y Yámana). Las líneas horizontales azul y roja representan el top 5 y 1% de los datos respectivamente. La línea horizontal anaranjada representa el top 1% inferior de la distribución de los datos. La figura superior muestra la distribución de la ancestría Amerindia a lo largo del Genoma, mientras que la figura inferior revela la distribución de la ancestría Europea a lo largo del Genoma.

El análisis para los grupos cazadores recolectores marítimos de la Patagonia (Figura 35, gráfico superior) nos muestra los mismos peaks por sobre el 1% en comparación con el análisis anterior (cromosoma 6 y 21), sin embargo aparecen otros nuevos loci en los cromosomas. En total se registraron 217 genes, de los cuales cinco pertenecen al cromosoma 2. De estos, están relacionados con respuesta celular

al estrés oxidativo, fosforilación de proteínas, reparación de ADN, generación de gametos¹¹.

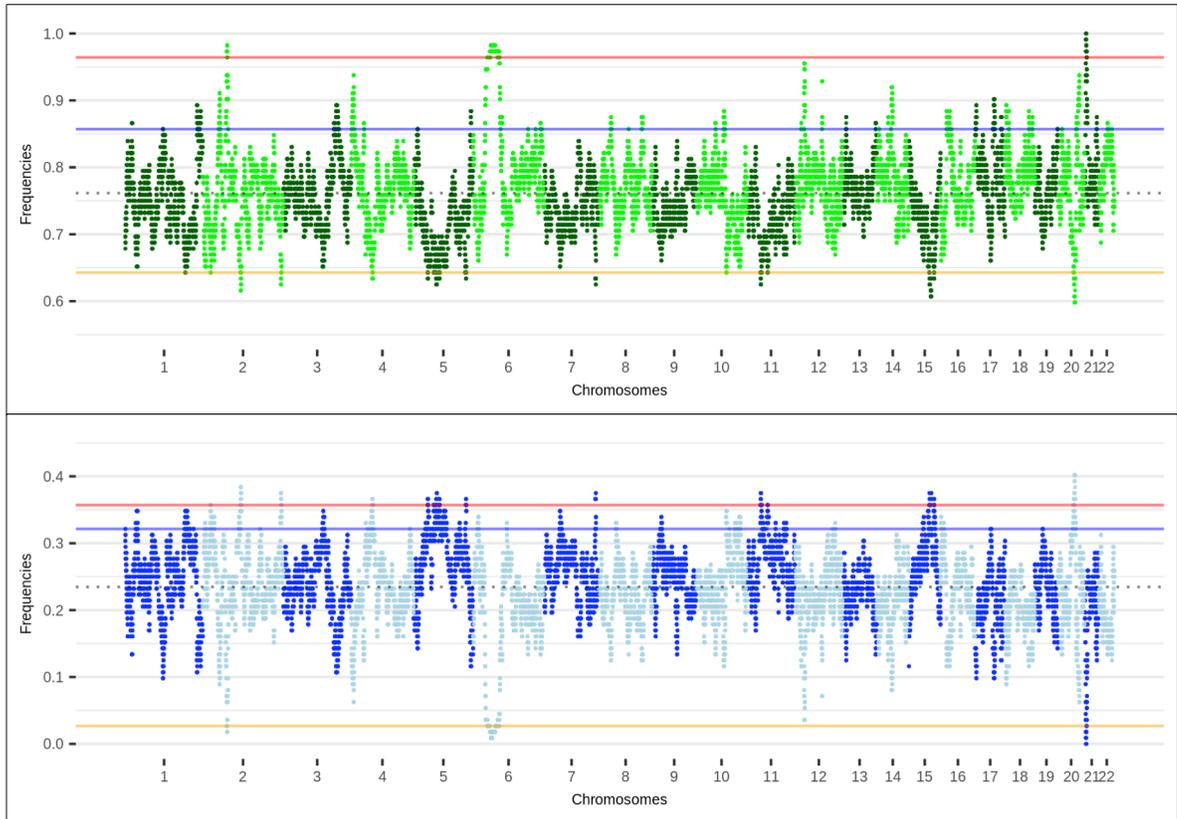


Figura 35. Detección de loci bajo selección a lo largo del Genoma basado en ancestrías para los grupos cazadores recolectores marítimos de la Patagonia (Chono, Kawéskar y Yámana). Las líneas horizontales azul y roja representan el top 5 y 1% de los datos respectivamente. La línea horizontal anaranjada representa el top 1% inferior de la distribución de los datos. La figura superior muestra la distribución de la ancestría Amerindia a lo largo del Genoma, mientras que la figura inferior revela la distribución de la ancestría Europea a lo largo del Genoma.

5.6. Enriquecimiento de Procesos Biológicos

Se identificaron mediante Annovar (K. Wang et al., 2010) aquellos genes que se encontraron por sobre el 1% de los datos para los tres análisis: PBS, iHS y Detección de loci bajo Selección mediante Ancestría. La caracterización de los

¹¹ Extraído a partir de <https://www.ncbi.nlm.nih.gov/gene/>

procesos biológicos enriquecidos se llevaron a cabo para todos los grupos de la Patagonia, esto quiere decir Cazadores Recolectores y Cazadores Recolectores Marítimos de la Patagonia. El enriquecimiento de los procesos biológicos se determinaron por la representación de genes asociado con un valor de significancia (p-value <0.05) y corregido por FDR (<0.05) a través de GO ontology (Mi, Muruganujan, Casagrande, & Thomas, 2013). Además, se evaluaron las rutas metabólicas, enzimáticas o biológicas sobre las cuales participan estos procesos mediante test de enriquecimiento a través del sitio www.webgestalt.org. En los resultados, se describieron únicamente aquellas rutas con una importancia biológica congruente con las hipótesis planteadas en este estudio, independiente de su nivel de significancia, dado que únicamente queremos explorar los datos de la Patagonia.

5.6.1. Population Branch Statistic (PBS)

5.6.1.1. Cazadores Recolectores de la Patagonia

Se realizó el análisis de enriquecimiento para los grupos cazadores recolectores de la Patagonia con un total de 2729 genes que se encontraron por sobre el 1% de los datos realizados a través de PBS. En la Tabla XVI se registraron un total de 46 procesos biológicos significativos (P-value < 0.05), de los cuales la mayoría está relacionado con procesos asociados al sistema nervioso, desarrollo de órganos y procesos celulares. De estos, cabe destacar la *respuesta celular al ácido retinoico* (GO:0071300), *regulación del transporte de iones de potasio* (GO:0043266) y *desarrollo del oído interno* (GO:0048839).

Tabla XVI. Enriquecimiento de procesos biológicos para los grupos cazadores recolectores de la Patagonia a partir de los genes por sobre el 1% identificados a partir de PBS.

GO	GO biological process complete	O	E	ER	P-value	FDR
GO:0051932	1. synaptic transmission, GABAergic	9	1.33	6.75	1.09E-04	1.52E-02
GO:0007214	2. gamma-aminobutyric acid signaling pathway	13	2.79	4.66	5.57E-05	9.04E-03
GO:0003300	3. cardiac muscle hypertrophy	13	3.39	3.83	2.53E-04	2.88E-02
GO:0086001	4. cardiac muscle cell action potential	17	5.58	3.05	2.75E-04	2.99E-02
GO:0071300	5. cellular response to retinoic acid	22	8.12	2.71	1.90E-04	2.35E-02
GO:0043266	6. regulation of potassium ion transport	26	10.18	2.55	1.38E-04	1.81E-02
GO:0060078	7. regulation of postsynaptic membrane potential	28	11.03	2.54	7.21E-05	1.08E-02
GO:0048839	8. inner ear development	42	21.94	1.91	3.07E-04	3.29E-02
GO:0001822	9. kidney development	54	30.55	1.77	2.70E-04	3.00E-02
GO:0003013	10. circulatory system process	69	42.91	1.61	4.68E-04	4.64E-02
GO:0072359	11. circulatory system development	133	94.67	1.40	3.72E-04	3.81E-02

GO= código de proceso biológico. O= número de genes observados en cada categoría GO. E= número de genes esperado para cada categoría GO. ER= Enriquecimiento de cada proceso biológico. P-value= significancia para cada proceso (<0.05). FDR= valores significativos de p-valor con FDR <0.05 de Benjamin & Hochberg (1995).

Se evaluó la asociación entre las regiones bajo selección y procesos biológicos y rutas metabólicas o de señalización celular mediante el test de sobrerrepresentación. En la Figura 36 se aprecia el top 10 de aquellas rutas con un enriquecimiento significativo ($p\text{-value} < 0.05$). De estas podemos destacar la secreción y síntesis de cortisol en las células fasciculadas de la zona adrenal (riñón). Este glucocorticoide afecta varias funciones fisiológicas, entre ellas el metabolismo de glucosa y lípidos, homeostasis metabólica y adaptación al estrés.

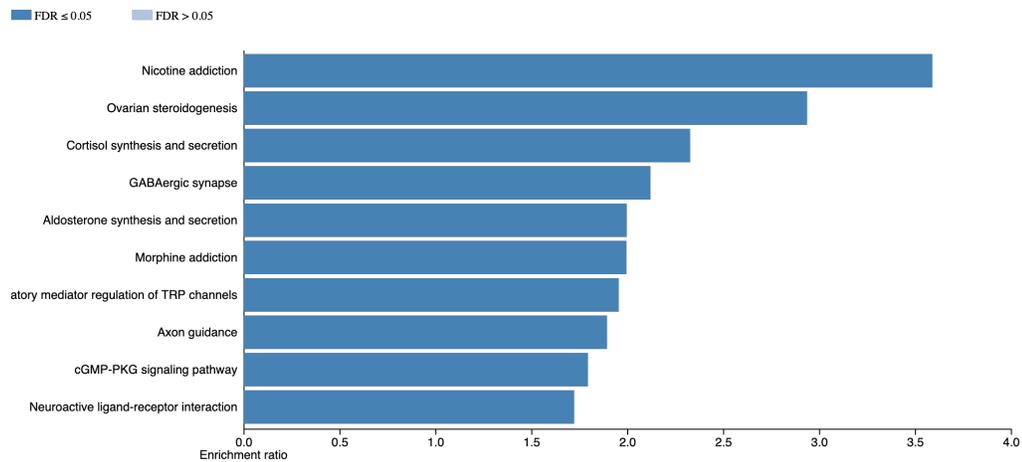


Figura 36. Esquema con la representación de rutas biológicas enriquecidas para los análisis de PBS sobre los SNPs en los individuos de la Patagonia. Cada barra horizontal representa una ruta biológica y el ancho corresponde a la proporción de enriquecimiento. El NES (Normalized Enrichment Score) indica la distribución de las categorías a lo largo de una lista de genes rankeados (valor de PBS). Valores extremos de NES indican un set de genes con un alto ranking y pertenecientes a una misma categoría, representando una regulación positiva (en azul). Los colores de tonalidades fuertes representan categorías con corrección a posterior (FDR < 0.05).

Se detectó un total de 93 ventanas por sobre el 1% de la distribución de los datos para los grupos cazadores recolectores de la Patagonia. De estas ventanas, destacan algunos procesos biológicos importantes para este estudio, los cuales están asociados con el procesamiento de insulina, lípidos y diferenciación de las células del oído interno (p-value < 0.05) (Tabla XVII).

Tabla XVII. Enriquecimiento de procesos biológicos en ventanas de 200 kb para los grupos cazadores recolectores de la Patagonia, a partir de los genes por sobre el 1% identificados a partir de PBS.

GO	GO biological process complete	O	E	ER	P-value	FDR*	Genes
GO:0030070	1. insulin processing	1	.03	29.25	4.43E-02	1.23E00	ERO1B
GO:0030259	2. lipid glycosylation	2	.14	14.63	1.05E-02	1.05E00	GBGT1;ABO
GO:0060113	3. inner ear receptor cell differentiation	3	.62	4.88	2.64E-02	1.18E00	KIF3A;FAT4;HES5
GO:0009247	4. glycolipid biosynthetic process	3	.72	4.18	3.84E-02	1.23E00	GBGT1;LARGE1;PIGP

FDR*=valores no significativos

GO= código de proceso biológico. O= número de genes observados en cada categoría GO. E= número de genes esperado para cada categoría GO. ER= Enriquecimiento de cada proceso biológico. P-value= significancia para cada proceso (<0.05). FDR= valores significativos de p-valor con FDR <0.05 de Benjamin & Hochberg (1995).

5.6.1.2. Cazadores Recolectores Marítimos de la Patagonia

En lo que respecta a los grupos cazadores recolectores marítimos de la Patagonia, pudimos detectar un total de 2720 genes por sobre el 1% de los datos. De estos se detectaron 26 procesos biológicos significativos (P-value < 0.05), de los cuales cabe mencionar algunos procesos que se repiten con análisis previos como la *regulación en el transporte de potasio* (GO:0006813) (Tabla XVIII), desarrollo del sistema nervioso, desarrollo de órganos, entre otros. Para el análisis de rutas metabólicas se puede apreciar que se vuelve a repetir la secreción y síntesis del cortisol, junto con otras como la secreción de insulina.

Tabla XVIII. Enriquecimiento de procesos biológicos para los grupos cazadores recolectores marítimos de la Patagonia a partir de los genes por sobre el 1% identificados a partir de PBS.

GO	GO biological process complete	O	E	ER	P-value	FDR
GO:0007214	1. gamma-aminobutyric acid signaling pathway	12	2.76	4.34	1.76E-04	3.56E-02
GO:0007416	2. synapse assembly	25	9.85	2.54	2.03E-04	3.95E-02
GO:0021953	3. central nervous system neuron differentiation	46	19.70	2.34	2.56E-06	1.11E-03
GO:0048738	4. cardiac muscle tissue development	39	17.78	2.19	4.14E-05	1.17E-02
GO:0007411	5. axon guidance	60	27.63	2.17	6.91E-07	3.71E-04
GO:0098742	6. cell-cell adhesion via plasma-membrane adhesion molecules	58	27.51	2.11	1.78E-06	8.41E-04
GO:0006813	7. potassium ion transport	37	17.66	2.10	1.63E-04	3.34E-02
GO:0007268	8. chemical synaptic transmission	76	44.20	1.72	3.50E-05	1.07E-02
GO:0030900	9. forebrain development	71	42.40	1.67	1.29E-04	2.68E-02
GO:0042391	10. regulation of membrane potential	76	46.00	1.65	1.22E-04	2.56E-02
GO:0098660	11. inorganic ion transmembrane transport	102	66.18	1.54	1.04E-04	2.28E-02
GO:0009887	12. animal organ morphogenesis	156	102.34	1.52	1.77E-06	8.59E-04
GO:0050767	13. regulation of neurogenesis	132	88.77	1.49	3.96E-05	1.14E-02
GO:0035295	14. tube development	138	92.97	1.48	2.84E-05	9.02E-03
GO:0051241	15. negative regulation of multicellular organismal process	182	127.92	1.42	1.24E-05	4.47E-03
GO:0048646	16. anatomical structure formation involved in morphogenesis	137	96.93	1.41	2.32E-04	4.25E-02
GO:0071495	17. cellular response to endogenous stimulus	170	123.60	1.38	1.11E-04	2.40E-02
GO:0051094	18. positive regulation of developmental process	192	145.22	1.32	2.72E-04	4.92E-02
GO:0045893	19. positive regulation of transcription, DNA-templated	211	161.43	1.31	2.29E-04	4.24E-02
GO:0032879	20. regulation of localization	357	289.48	1.23	8.47E-05	2.03E-02
GO:0010646	21. regulation of cell communication	461	385.57	1.20	8.47E-05	2.00E-02
GO:0023051	22. regulation of signaling	465	390.13	1.19	1.03E-04	2.28E-02
GO:0048523	23. negative regulation of cellular process	572	483.82	1.18	2.08E-05	6.88E-03

GO:0050911	24. detection of chemical stimulus involved in sensory perception of smell	18	43.60	.41	3.64E-05	1.07E-02
GO:0031424	25. keratinization	6	21.86	.27	2.06E-04	3.95E-02
GO:0019080	26. viral gene expression	1	12.25	.08	1.78E-04	3.54E-02

GO= código de proceso biológico. O= número de genes observados en cada categoría GO. E= número de genes esperado para cada categoría GO. ER= Enriquecimiento de cada proceso biológico. P-value= significancia para cada proceso (<0.05). FDR= valores significativos de p-valor con FDR <0.05 de Benjamin & Hochberg (1995).

En lo que respecta al análisis de PBS sobre ventanas de 200 kb podemos desprender un total de 93 ventanas por sobre el 1% de los datos, los que arrojaron 447 procesos biológicos (P-value < 0.05). En la Tabla XIX se puede apreciar que existen diferentes procesos asociados a metabolismo de lípidos, diferenciación celular ligada a células pardas, transporte de electrones en la mitocondria, morfogénesis del otolito y regulación positiva y negativa frente a la termogénesis inducida por frío. En la mayoría de estos participan el gen **ADIPOQ** (cromosoma 3). En lo que respecta a la identificación de rutas metabólicas destaca la termogénesis, la cual se produce en los adipocitos de la grasa parda. Además, vemos que existen otras rutas asociadas a fenotipos como diabetes mellitus tipo 2 e hígado graso no-alcohólico.

Tabla XIX. Enriquecimiento de procesos biológicos en ventanas de 200 kb para los grupos cazadores recolectores marítimos de la Patagonia, a partir de los genes por sobre el 1% identificados a partir de PBS.

GO	GO biological process complete	O	E	ER	P-value	FDR*
GO:1902958	1. positive regulation of mitochondrial electron transport, NADH to ubiquinone	1	.02	44.30	3.31E-02	2.10E00
GO:0032474	2. otolith morphogenesis	1	.03	29.53	4.39E-02	2.29E00
GO:0050873	3. brown fat cell differentiation	2	.33	6.11	4.66E-02	2.14E00
GO:0120163	4. negative regulation of cold-induced thermogenesis	3	.52	5.78	1.73E-02	3.36E00
GO:0006635	5. fatty acid beta-oxidation	3	.64	4.66	2.94E-02	2.50E00
GO:0070542	6. response to fatty acid	4	.95	4.22	1.70E-02	3.37E00
GO:0120162	7. positive regulation of cold-induced thermogenesis	4	1.08	3.69	2.58E-02	2.38E00
GO:0045598	8. regulation of fat cell differentiation	4	1.33	3.00	4.79E-02	2.17E00

FDR*=valores no significativos

GO= código de proceso biológico. O= número de genes observados en cada categoría GO. E= número de genes esperado para cada categoría GO. ER= Enriquecimiento de cada proceso biológico. P-value= significancia para cada proceso (<0.05). FDR= valores significativos de p-valor con FDR <0.05 de Benjamin & Hochberg (1995).

5.6.2. integrated Haplotype Score (iHS)

El análisis de enriquecimiento para iHS se realizó sobre 540 genes que se encuentran por sobre el 1% de la distribución de los datos. Se encontraron un total de 77 procesos biológicos (P-value < 0.05) de los cuales se encontraron procesos asociados a la *diferenciación y regulación de adipocitos, procesos asociados a la regulación de polisacáridos y glicoproteínas, respuesta celular al ácido retinoico* y por último, procesos relacionados con el *estímulo a la temperatura y termogénesis adaptativa* Tabla XX, siendo particularmente el gen THADA el que se encuentra participando en este proceso y descrito previamente como una región bajo adaptación por parte de diversas poblaciones de Siberia (Cardona et al., 2014).

Tabla XX. Enriquecimiento de procesos biológicos para los grupos cazadores recolectores de la Patagonia a partir de los genes con valores de |iHS| > 2.5 por SNPs.

GO	GO biological process complete	O	E	ER	P-value	FDR*
GO:0019373	1. epoxygenase P450 pathway	4	.21	19.42	9.65E-04	1.78E00
GO:1990845	2. adaptive thermogenesis	2	.10	19.42	2.45E-02	3.43E00
GO:0050872	3. white fat cell differentiation	2	.10	19.42	2.45E-02	3.27E00
GO:0010560	4. positive regulation of glycoprotein biosynthetic process	2	.10	19.42	2.45E-02	2.95E00
GO:0050951	5. sensory perception of temperature stimulus	3	.51	5.83	3.17E-02	3.00E00
GO:0016048	6. detection of temperature stimulus	3	.51	5.83	3.17E-02	2.93E00
GO:0033692	7. cellular polysaccharide biosynthetic process	4	.72	5.55	1.45E-02	3.28E00
GO:0034605	8. cellular response to heat	4	.82	4.85	2.02E-02	3.29E00
GO:0071300	9. cellular response to retinoic acid	4	1.13	3.53	4.46E-02	3.07E00
GO:0045600	10. positive regulation of fat cell differentiation	5	1.54	3.24	3.31E-02	2.95E00
GO:1900034	12. regulation of cellular response to heat	5	1.54	3.24	3.31E-02	2.93E00

FDR*=valores no significativos

GO= código de proceso biológico. O= número de genes observados en cada categoría GO. E= número de genes esperado para cada categoría GO. ER= Enriquecimiento de cada proceso biológico. P-value= significancia para cada proceso (<0.05). FDR= valores significativos de p-valor con FDR <0.05 de Benjamin & Hochberg (1995).

Para el caso de las rutas biológicas (p-value < 0.05) podemos destacar de algunas rutas relacionadas con el metabolismo y degradación de ácidos grasos, así como también con el metabolismo de B Alanina. En todos estas rutas actúa principalmente el gen **ACADM**, el cual esta relacionado con la enzima acyl-Coa deshidrogenasa.

Para el análisis con las ventanas de 200 kb para estas poblaciones se registraron solo 12 genes distribuidos sobre el top 1% de los datos. Se detectaron 124 procesos biológicos enriquecidos (P-value < 0.05), de los cuales vale mencionar los que están relacionados con el proceso metabólico de la carnitina, una molécula encargada del transporte de lípidos hacia la mitocondrial (Hoppel, 2003), transporte de sacarosa, metabolismo de lípidos y regulación de la glucosa. Para los procesos relacionados con el metabolismo de lípidos y azúcares nos aparece el gen **ACADM** como principal protagonista.

Tabla XXI. Enriquecimiento de procesos biológicos en ventanas de 200kb para los grupos cazadores recolectores de la Patagonia, a partir de los genes por sobre el 1% identificados a partir de iHS.

GO	GO biological process complete	O	E	ER	P-value	FDR*	Genes
GO:0010255	1. glucose mediated signaling pathway	1	.01	73.12	2.68E-02	1.00E00	AGER
GO:0019254	2. carnitine metabolic process, CoA-linked	1	.01	73.12	2.68E-02	1.00E00	ACADM
GO:0032094	3. response to food	1	.01	73.12	2.68E-02	1.00E00	GHR
GO:0015770	4. sucrose transport	1	.01	73.12	2.68E-02	1.00E00	SLC45A1
GO:0010225	5. response to UV-C	1	.03	36.56	3.99E-02	1.00E00	BCL3
GO:0033539	6. fatty acid beta-oxidation using acyl-CoA dehydrogenase	1	.03	36.56	3.99E-02	1.00E00	ACADM
GO:0045329	7. carnitine biosynthetic process	1	.03	36.56	3.99E-02	1.00E00	ACADM
GO:0051793	8. medium-chain fatty acid catabolic process	1	.03	36.56	3.99E-02	1.00E00	ACADM

FDR*=valores no significativos

GO= código de proceso biológico. O= número de genes observados en cada categoría GO. E= número de genes esperado para cada categoría GO. ER= Enriquecimiento de cada proceso biológico. P-value= significancia para cada proceso (<0.05). FDR= valores significativos de p-valor con FDR <0.05 de Benjamin & Hochberg (1995).

En lo que respecta a las rutas metabólicas, podemos ver que dos procesos que están asociados con el metabolismo de ácidos grasos y la degradación de estos (p-value < 0.05) manteniendo el mismo patrón; el gen **ACADM** se encuentra presente en la mayoría de estas rutas metabólicas (**Figura 37**).

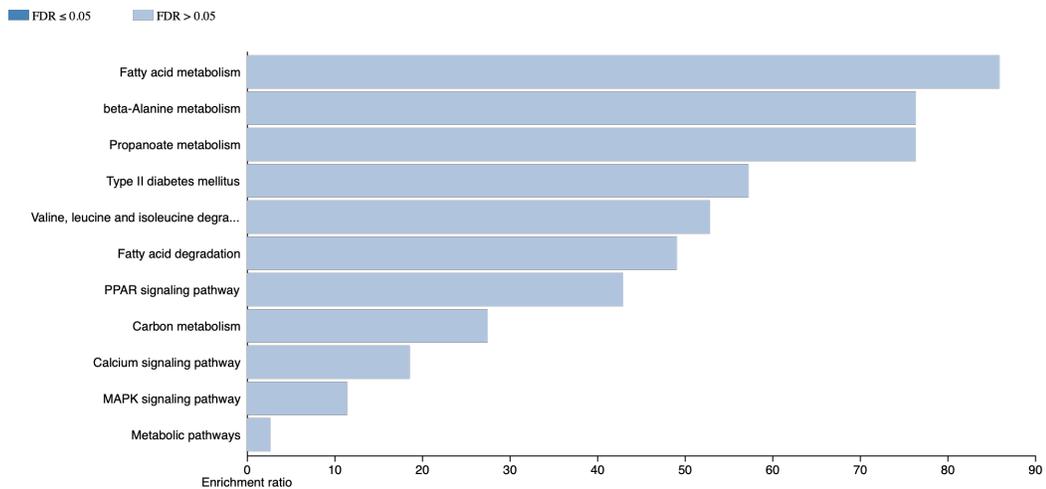


Figura 37. Esquema con la representación de rutas biológicas enriquecidas para los análisis de |iHS| sobre las ventanas de 200 Kb en los individuos de la Patagonia. Cada barra horizontal representa una ruta biológica y el ancho corresponde a la proporción de enriquecimiento. El NES (Normalized Enrichment Score) indica la distribución de las categorías a lo largo de una lista de genes rankeados (valor de PBS). Valores extremos de NES indican un set de genes con un alto ranking y pertenecientes a una misma categoría, representando una regulación positiva (en azul). Los colores de tonalidades suaves representan categorías con corrección a posterior no significativa (FDR < 0.05).

5.6.3. Detección de Loci bajo Selección mediante Ancestría

Para este análisis se identificaron un total de 200 genes por sobre el 1% de los datos, de los cuales se encontraron 17 procesos biológicos sobre-representados significativos (P-value < 0.05) a nivel genómico (Tabla XXII). Para este grupo podemos ver que la mayoría de los procesos biológicos está relacionado con el Complejo Mayor de Histocompatibilidad (MHC) asociado a la respuesta inmune, de hecho la mayoría de los genes son de la familia HLA. No obstante, aparecen los genes **MICA** y **MICB**, asociados a diferentes procesos biológicos como la *activación de células T gamma-delta*, *regulación positiva de las células natural killer*, *inmunidad mediada por linfocitos* y *procesos virales*. También se encuentra asociado el gen **LTA**, el cual participa en la *regulación positiva en la respuesta inflamatoria frente a estímulos de antígenos*.

Tabla XXII. Enriquecimiento de procesos biológicos para los grupos cazadores recolectores de la Patagonia a partir de los genes por sobre el 1% identificados en aquellos Loci bajo selección mediante Ancestría.

GO	GO biological process complete	O	E	ER	P-value	FDR
GO:0002503	1. peptide antigen assembly with MHC class II protein complex	4	.04	> 100	4.99E-07	2.69E-04
GO:1990668	2. vesicle fusion with endoplasmic reticulum-Golgi intermediate compartment (ERGIC) membrane	3	.05	63.64	4.33E-05	1.23E-02
GO:0002476	3. antigen processing and presentation of endogenous peptide antigen via MHC class Ib	4	.08	53.04	3.42E-06	1.49E-03
GO:0046629	4. gamma-delta T cell activation	3	.06	53.04	6.45E-05	1.71E-02
GO:0002486	5. antigen processing and presentation of endogenous peptide antigen via MHC class I via ER pathway, TAP-independent	3	.07	45.46	9.15E-05	2.30E-02
GO:0002480	6. antigen processing and presentation of exogenous peptide antigen via MHC class I, TAP-independent	3	.08	35.36	1.65E-04	3.54E-02
GO:0002863	7. positive regulation of inflammatory response to antigenic stimulus	3	.09	31.82	2.14E-04	4.28E-02
GO:0060333	8. interferon-gamma-mediated signaling pathway	12	.62	19.29	7.47E-12	2.33E-08
GO:0019886	9. antigen processing and presentation of exogenous peptide antigen via MHC class II	13	.86	15.15	1.49E-11	3.88E-08
GO:0045954	10. positive regulation of natural killer cell mediated cytotoxicity	4	.27	14.63	2.43E-04	4.63E-02
GO:0002479	11. antigen processing and presentation of exogenous peptide antigen via MHC class I, TAP-dependent	8	.61	13.06	3.91E-07	2.18E-04
GO:0042102	12. positive regulation of T cell proliferation	7	.82	8.53	2.82E-05	9.18E-03
GO:0050852	13. T cell receptor signaling pathway	12	1.44	8.32	4.83E-08	3.77E-05
GO:0035967	14. cellular response to topologically incorrect protein	8	1.19	6.73	3.75E-05	1.10E-02
GO:0002449	15. lymphocyte mediated immunity	7	1.12	6.24	1.81E-04	3.78E-02
GO:0001819	16. positive regulation of cytokine production	14	3.84	3.65	4.20E-05	1.21E-02
GO:0016032	17. viral process	18	5.86	3.07	3.10E-05	9.88E-03

GO= código de proceso biológico. O= número de genes observados en cada categoría GO. E= número de genes esperado para cada categoría GO. ER= Enriquecimiento de cada proceso biológico. P-value= significancia para cada proceso (<0.05). FDR= valores significativos de p-valor con FDR <0.05 de Benjamin & Hochberg (1995).

Cuando se evaluaron las principales rutas metabólicas, biológicas y/o enzimáticas se detectaron en su mayoría fenotipos relacionados con patologías inmunes, virales, bacterianas, artritis reumatoide, lupus eritematoso y diabetes mellitus tipo I (p-value < 0.05) (Figura 38).

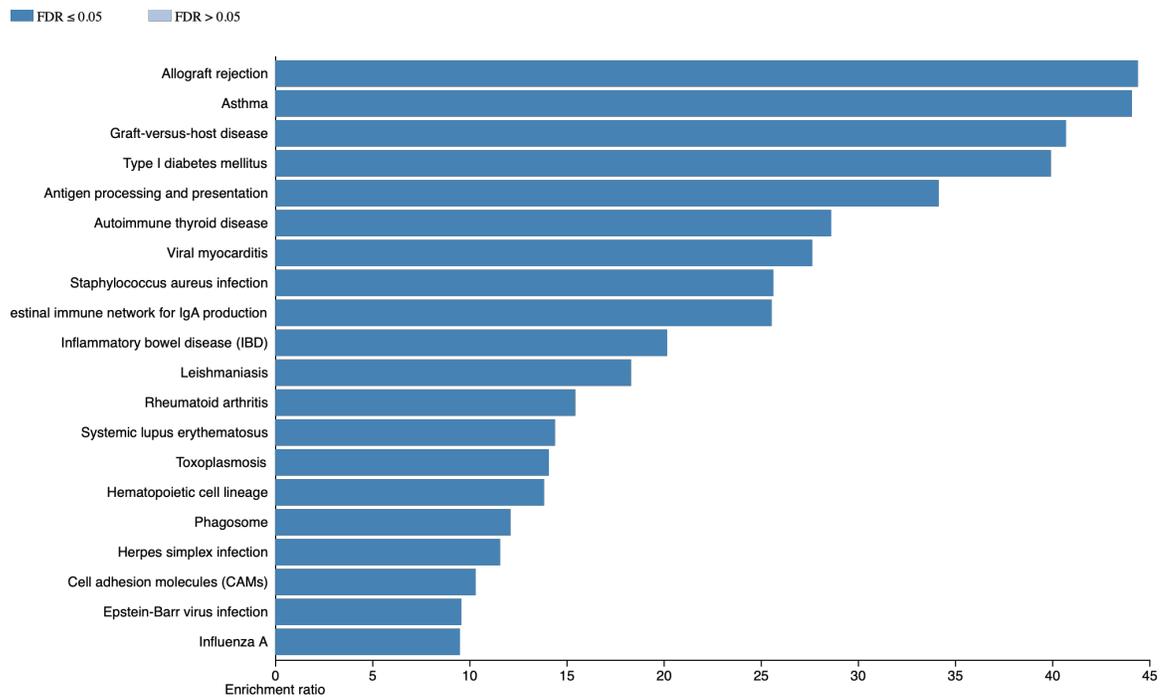


Figura 38. Esquema con la representación de rutas biológicas enriquecidas en grupos cazadores recolectores de la Patagonia. Cada barra horizontal representa una ruta biológica y el ancho corresponde a la proporción de enriquecimiento. El NES (Normalized Enrichment Score) indica la distribución de las categorías a lo largo de una lista de genes rankeados (valor de PBS). Valores extremos de NES indican un set de genes con un alto ranking y pertenecientes a una misma categoría, representando una regulación positiva (en azul). Los colores de tonalidades fuertes representan categorías con corrección a posterior (FDR < 0.05).

Se evaluaron los procesos biológicos para la ancestría Europea para un total de 331 genes sobre el 1% de los datos, de los cuales se encontraron mas de 200 funciones sobre-representadas (p-value < 0.05). Sin embargo, algunas de las funciones mas interesantes que pudimos encontrar corresponden al procesamiento de la vitamina E, respuesta celular a la ausencia de luz, resistencia contra bacterias y virus, procesos catabólicos para NAD y NADP y metabolismo de ácidos grasos (Tabla XXIII). En cuanto a las rutas metabólicas se encontraron principalmente asociadas a la degradación de glicanos, glicosaminoglicanos y biosíntesis de glicofosfolípidos (p-value < 0.05).

Tabla XXIII. Enriquecimiento de procesos biológicos para los grupos cazadores recolectores de la Patagonia a partir de los genes por sobre el 1% identificados en aquellos Loci bajo selección mediante Ancestría Europea.

GO	GO biological process complete	O	E	ER	P-value	FDR*	Genes
GO:0080058	protein deglutathionylation	1	.01	74.85	2.62E-02	1.00E00	GLRX
GO:0010189	vitamin E biosynthetic process	1	.01	74.85	2.62E-02	1.00E00	PLTP
GO:0045225	negative regulation of CD4 biosynthetic process	1	.01	74.85	2.62E-02	1.00E00	ACOT8
GO:0071485	cellular response to absence of light	1	.01	74.85	2.62E-02	1.00E00	PAPD4
GO:0070963	positive regulation of neutrophil mediated killing of gram-negative bacterium	1	.01	74.85	2.62E-02	1.00E00	F2RL1
GO:1990603	dark adaptation	1	.01	74.85	2.62E-02	1.00E00	PAPD4
GO:2000208	positive regulation of ribosomal small subunit export from nucleus	1	.01	74.85	2.62E-02	1.00E00	RIOK2
GO:0040040	thermosensory behavior	1	.01	74.85	2.62E-02	1.00E00	SLC12A5
GO:0051460	negative regulation of corticotropin secretion	1	.01	74.85	2.62E-02	1.00E00	CRHBP
GO:0019677	NAD catabolic process	1	.01	74.85	2.62E-02	1.00E00	NUDT12
GO:0002528	regulation of vascular permeability involved in acute inflammatory response	2	.05	37.42	2.51E-03	1.00E00	C2CD4B; C2CD4A
GO:0006742	NADP catabolic process	1	.03	37.42	3.90E-02	1.00E00	NUDT12
GO:0060139	positive regulation of apoptotic process by virus	1	.03	37.42	3.90E-02	1.00E00	BCL2L11
GO:1900169	regulation of glucocorticoid mediated signaling pathway	1	.03	37.42	3.90E-02	1.00E00	AKAP13
GO:0035865	cellular response to potassium ion	2	.13	14.97	1.05E-02	1.00E00	CACNA1H ;CRHBP
GO:0033540	fatty acid beta-oxidation using acyl-CoA oxidase	2	.19	10.69	1.84E-02	1.00E00	ACOT8;A COXL

GO= código de proceso biológico. O= número de genes observados en cada categoría GO. E= número de genes esperado para cada categoría GO. ER= Enriquecimiento de cada proceso biológico. P-value= significancia para cada proceso (<0.05). FDR= valores significativos de p-valor con FDR <0.05 de Benjamin & Hochberg (1995).

Se registraron un total de 217 genes por sobre el 1% para los grupos cazadores recolectores marítimos de la Patagonia, con 26 procesos biológicos sobre-representados (P-value < 0.05) (Tabla XXIV). De estos, la mayoría se repiten con la comparación anterior, situándose dentro del proceso de respuesta inmune, no obstante vemos que aparece un proceso biológico relacionado con la *respuesta al calor*, donde los principales genes recaen en **MICA** y **MICB** tal como se habían descrito previamente.

Tabla XXIV. Enriquecimiento de procesos biológicos para los grupos cazadores recolectores marítimos de la Patagonia a partir de los genes por sobre el 1% identificados en aquellos Loci bajo selección a través de Ancestría.

GO	GO biological process complete	O	E	ER	P-value	FDR
GO:0002486	1. antigen processing and presentation of endogenous peptide antigen via MHC class I via ER pathway, TAP-independent	7	.07	> 100	1.80E-11	1.22E-08
GO:0002503	2. peptide antigen assembly with MHC class II protein complex	4	.04	> 100	5.03E-07	1.21E-04
GO:0002428	3. antigen processing and presentation of peptide antigen via MHC class Ib	8	.09	94.09	1.12E-12	1.10E-09
GO:0002480	4. antigen processing and presentation of exogenous peptide antigen via MHC class I, TAP-independent	7	.09	82.33	5.89E-11	3.29E-08
GO:0042270	5. protection from natural killer cell mediated cytotoxicity	4	.06	70.57	1.49E-06	3.06E-04
GO:1990668	6. vesicle fusion with endoplasmic reticulum-Golgi intermediate compartment (ERGIC) membrane	3	.05	63.51	4.35E-05	6.48E-03
GO:0046629	7. gamma-delta T cell activation	3	.06	52.92	6.49E-05	9.31E-03
GO:0002729	8. positive regulation of natural killer cell cytokine production	3	.06	52.92	6.49E-05	9.22E-03
GO:0002863	9. positive regulation of inflammatory response to antigenic stimulus	3	.09	31.75	2.15E-04	2.67E-02
GO:0001916	10. positive regulation of T cell mediated cytotoxicity	7	.23	30.87	1.20E-08	4.59E-06
GO:0060333	11. interferon-gamma-mediated signaling pathway	18	.62	28.87	6.55E-20	2.05E-16
GO:2001198	12. regulation of dendritic cell differentiation	3	.11	26.46	3.37E-04	3.99E-02
GO:0002479	13. antigen processing and presentation of exogenous peptide antigen via MHC class I, TAP-dependent	12	.61	19.54	6.54E-12	5.68E-09
GO:0045954	14. positive regulation of natural killer cell mediated cytotoxicity	5	.27	18.25	1.51E-05	2.48E-03
GO:0019886	15. antigen processing and presentation of exogenous peptide antigen via MHC class II	13	.86	15.12	1.53E-11	1.20E-08
GO:0046596	16. regulation of viral entry into host cell	4	.28	14.11	2.76E-04	3.37E-02
GO:0060337	17. type I interferon signaling pathway	8	.60	13.23	3.57E-07	9.00E-05
GO:0042102	18. positive regulation of T cell proliferation	8	.82	9.73	3.02E-06	5.76E-04
GO:0032729	19. positive regulation of interferon-gamma production	5	.56	8.97	3.30E-04	3.94E-02
GO:0050852	20. T cell receptor signaling pathway	12	1.45	8.30	4.95E-08	1.52E-05
GO:0002449	21. lymphocyte mediated immunity	8	1.13	7.06	2.74E-05	4.37E-03
GO:0035967	22. cellular response to topologically incorrect protein	8	1.20	6.67	4.02E-05	6.04E-03
GO:0009408	23. response to heat	6	.91	6.62	3.94E-04	4.53E-02
GO:0002460	24. adaptive immune response based on somatic recombination of immune receptors built from immunoglobulin superfamily domains	7	1.26	5.57	3.51E-04	4.06E-02
GO:0043901	25. negative regulation of multi-organism process	8	1.59	5.04	2.55E-04	3.14E-02

GO:0016032	26. viral process	19	5.92	3.21	1.02E-05	1.75E-03
-------------------	-------------------	----	------	------	----------	----------

GO= código de proceso biológico. O= número de genes observados en cada categoría GO. E= número de genes esperado para cada categoría GO. ER= Enriquecimiento de cada proceso biológico. P-value= significancia para cada proceso (<0.05). FDR= valores significativos de p-valor con FDR <0.05 de Benjamin & Hochberg (1995).

Para la identificación de rutas metabólicas se pudo constatar que las mismas categorías que fueron detectadas para el grupo anterior, encontrando patologías asociadas a enfermedades inmunes, afecciones de carga viral y bacteriana, así como también afecciones como asma y diabetes mellitus tipo I (Figura 39).

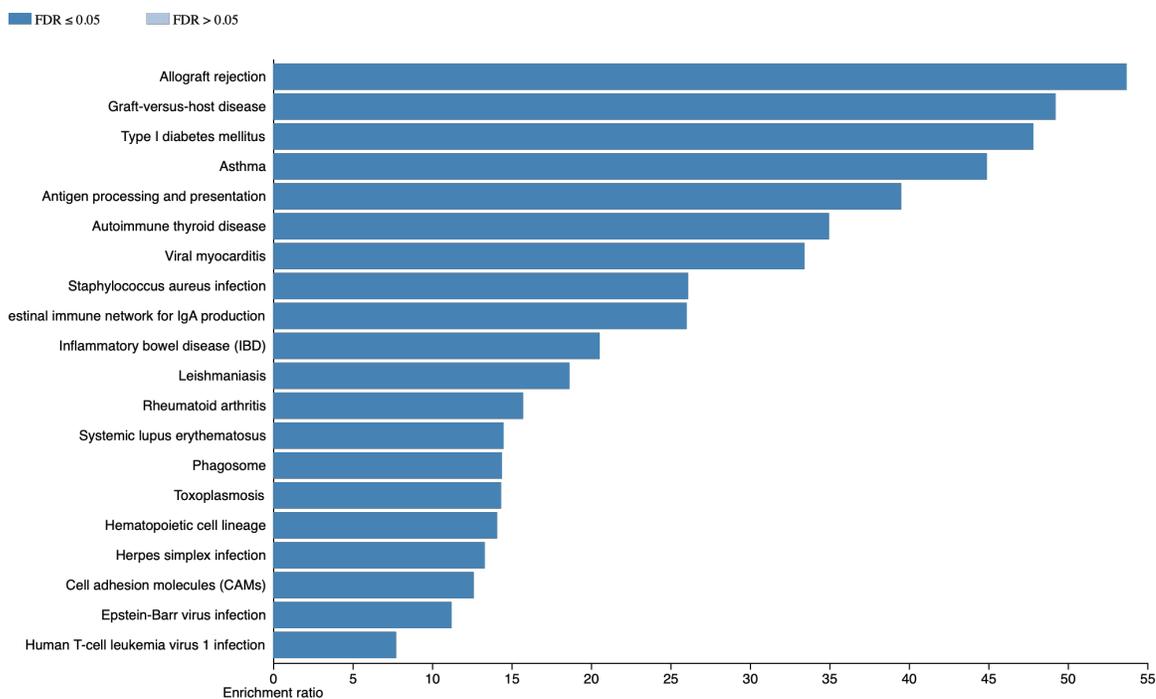


Figura 39. Esquema con la representación de rutas biológicas enriquecidas en grupos cazadores recolectores marítimos de la Patagonia. Cada barra horizontal representa una ruta biológica y el ancho corresponde a la proporción de enriquecimiento. El NES (Normalized Enrichment Score) indica la distribución de las categorías a lo largo de una lista de genes rankeados (valor de PBS). Valores extremos de NES indican un set de genes con un alto ranking y pertenecientes a una misma categoría, representando una regulación positiva (en azul). Los colores de tonalidades fuertes representan categorías con corrección a posterior (FDR < 0.05).

Se evaluaron los procesos biológicos para la ancestría Europea para un total de 348 genes sobre el 1% de los datos, de los cuales se encontraron dos funciones sobre-representadas. Estas están relacionadas con receptores de trombina y regulación negativa de la actividad endopeptidasa (Tabla XXV). En cuanto a las rutas

metabólicas se encontraron asociadas al sistema angiotensina-renina, biosíntesis de glicofosfolípidos, metabolismo de selenocomponentes, entre otros (p -value < 0.05).

Tabla XXV. Enriquecimiento de procesos biológicos para los grupos cazadores recolectores marítimos de la Patagonia a partir de los genes por sobre el 1% identificados en aquellos Loci bajo selección a través de Ancestría Europea.

GO	GO biological process complete	O	E	ER	P-value	FDR
GO:0070493	thrombin-activated receptor signaling pathway	4	.13	29.90	2.90E-05	4.12E-02
GO:0010466	negative regulation of peptidase activity	18	3.40	5.29	2.59E-08	4.06E-04
GO:0010951	negative regulation of endopeptidase activity	16	3.28	4.87	4.44E-07	2.32E-03
GO:0045861	negative regulation of proteolysis	18	4.64	3.88	1.90E-06	4.25E-03
GO:0052547	regulation of peptidase activity	22	5.92	3.72	2.66E-07	2.08E-03
GO:0052548	regulation of endopeptidase activity	20	5.56	3.60	1.54E-06	4.82E-03
GO:0051346	negative regulation of hydrolase activity	21	6.11	3.44	1.67E-06	4.36E-03
GO:0030162	regulation of proteolysis	25	9.69	2.58	2.20E-05	3.43E-02
GO:0051248	negative regulation of protein metabolic process	36	14.68	2.45	1.12E-06	4.38E-03
GO:0032269	negative regulation of cellular protein metabolic process	33	13.76	2.40	4.93E-06	9.63E-03
GO:0051246	regulation of protein metabolic process	64	37.11	1.72	1.09E-05	1.89E-02

GO= código de proceso biológico. O= número de genes observados en cada categoría GO. E= número de genes esperado para cada categoría GO. ER= Enriquecimiento de cada proceso biológico. P-value= significancia para cada proceso (<0.05). FDR= valores significativos de p-valor con FDR <0.05 de Benjamin & Hochberg (1995).

5.7. Genes Candidatos a Selección Positiva en la Patagonia

Se describieron en la Tabla XXVI un set de genes que se encontraron dentro del top 10 de las señales mas potentes que arrojaron los análisis de PBS e iHS, tanto para SNPs como por ventanas de 200 Kb. Estos genes fueron señalados de especial interés para responder las hipótesis relacionadas con frío y dieta dada sus características asociadas con el metabolismo de lípidos (ácidos grasos de cadena larga y colesterol), transporte de electrones en la mitocondria, respuestas asociadas a termogénesis y la regulación celular frente a la ingesta de nutrientes.

Tabla XXVI. Descripción de los genes candidatos para frío y dieta en la Patagonia con los análisis realizados con PBS e iHS (SNPs y ventenas de 200Kb).

Crom	rsID	Inicio	Termino	Gen	Marcador		Grupo		Análisis
					SNPs	Ventana (200Kb)	CR	CRM	
2	rs74261335			ASXL2	X		X		PBS
6	rs17756843			GFRAL	X			X	PBS
17	rs2083797			WNT9B	X			X	PBS
11		123600000	123800000	ZNF202, GRAMD1B		X	X	X	PBS
1		229400000	229600000	NUP133, ABCB10		X	X		PBS
2		169000000	169200000	ABCB11, LRP11		X	X		PBS
2	rs76324176			SLC39A10	X		X		iHS
2	rs10495562			ADAM17	X		X		iHS

Nos dispusimos a comparar estos genes bajo selección entre los grupos de la Patagonia (cazadores recolectores) y los cazadores recolectores marítimos únicamente para los análisis de PBS –no se realizó para iHS puesto que solo se pudo evaluar para cazadores recolectores- tanto para SNPs como por ventanas (200Kb). Para el caso de los SNPs, evaluamos el Top 10 de los genes bajo selección para los cazadores recolectores (CR) y realizamos un intersección con los genes sobre el 1% para los grupos cazadores recolectores marítimos de la Patagonia (CRM). Respecto a esto, podemos ver que solo hay cuatro genes dispuestos en el 1% de los CRM presentes en el Top 10 de los CR. Notamos específicamente que el gen **ASXL2** se encuentra en la posición 3449, muy por debajo de la potente señal de selección vista por los CR de la Patagonia (Figura 40). Para los análisis realizados a través de ventanas (200Kb) podemos ver que de las 10 ventanas dentro del Top 10 para los CR, solo 3 se encuentran dentro del 1% para los CRM de la Patagonia. De estas destaca la ventana que contiene los genes **ZNF202** y **GRAMD1B** las cuales están dentro del Top 10 para ambos grupos (Figura 41).

Cuando hacemos la comparación al sentido opuesto, es decir, el intersección entre el top 10 de los CRM con el 1% de los CR para los análisis mediados por SNPs vemos 5 señales presentes, del cual destaca el gen **GFRAL** el cual se ubica en el puesto 872 de las señales más potentes en los CR, siendo una de las más potentes en los

CRM. Para los análisis realizados por ventanas no se detectó ninguna intersección entre estas dos agrupaciones (Figura 42).

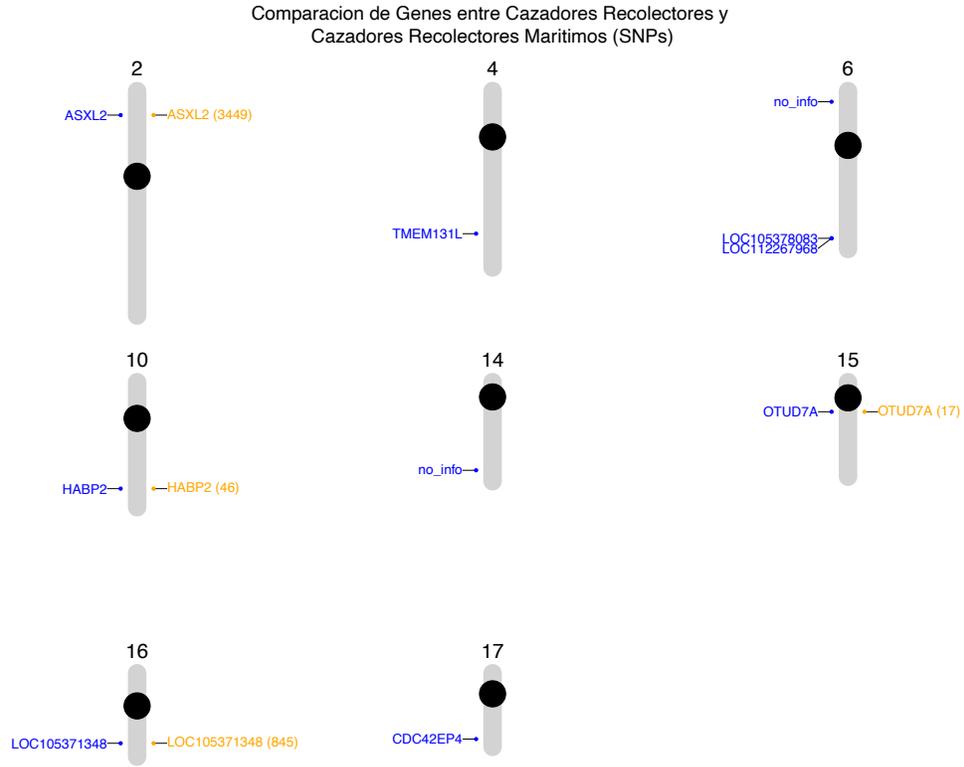


Figura 40. Intersección de genes bajo selección entre el Top 10 de los cazadores recolectores de la Patagonia (azules, lado izquierdo de cada cromosoma) y el 1% de los cazadores recolectores marítimos de la Patagonia (naranjos, lado derecho de cada cromosoma). Se mostraron solo aquellos cromosomas con genes dentro del Top 10. Aquellos con 'no_info' representan SNPs sin asociación con un gen. Los análisis están realizados por SNPs.

Comparacion de Genes entre Cazadores Recolectores y Cazadores Recolectores Maritimos (Ventanas)

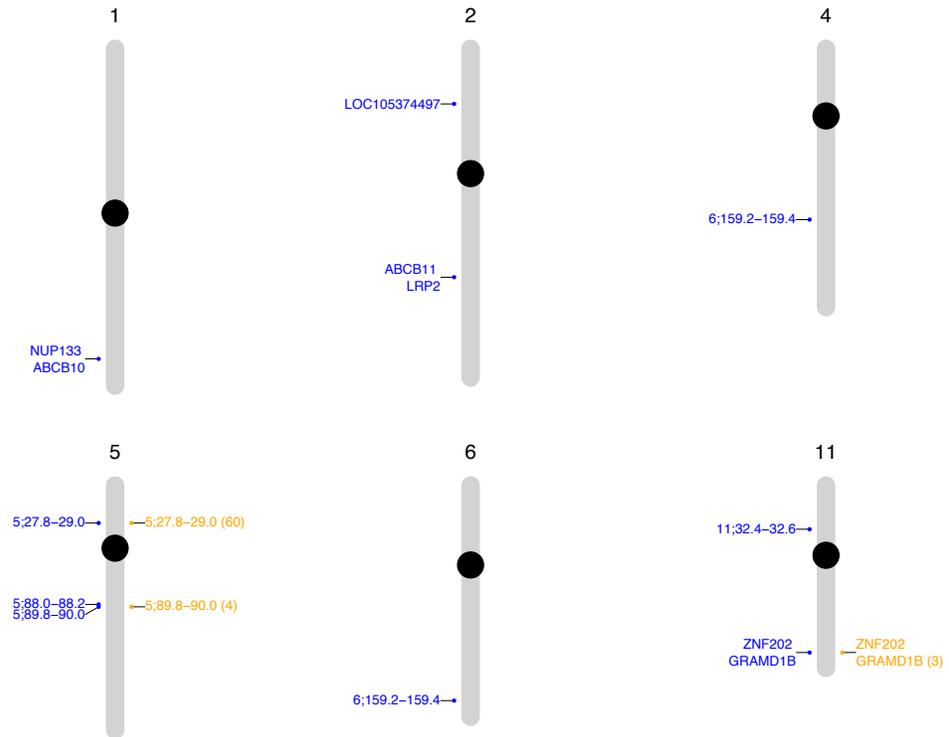


Figura 41. Intersección de genes bajo selección entre el Top 10 de los cazadores recolectores de la Patagonia (azules, lado izquierdo de cada cromosoma) y el 1% de los cazadores recolectores marítimos de la Patagonia (naranjos, lado derecho de cada cromosoma). Se mostraron solo aquellos cromosomas con genes dentro del Top 10. Los análisis están realizados por ventanas de 200Kb y la nomenclatura representa el cromosoma y el tamaño de la ventana. Se describieron únicamente los nombres de aquellos genes en ventanas con especial interés como regiones candidatas a frío y dieta.

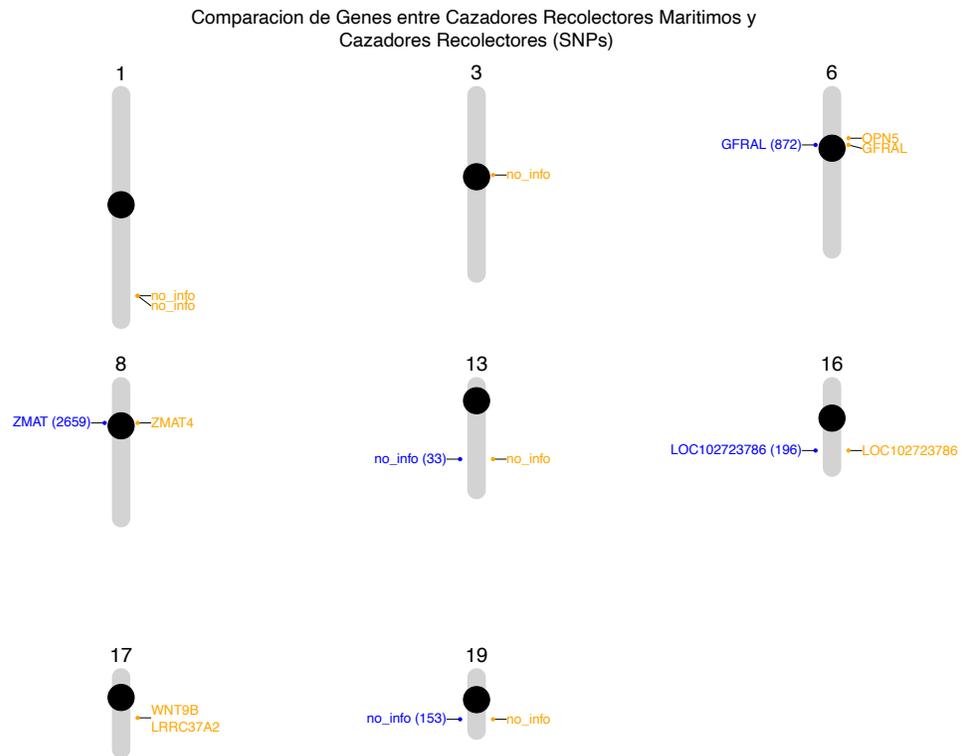


Figura 42. Intersección de genes bajo selección entre el Top 10 de los cazadores recolectores marítimos de la Patagonia (naranja, lado derecho de cada cromosoma) y el 1% de los cazadores recolectores de la Patagonia (azules, lado izquierdo de cada cromosoma). Se mostraron solo aquellos cromosomas con genes dentro del Top 10. Aquellos con 'no_info' representan SNPs sin asociación con un gen. Los análisis están realizados por SNPs.

5.7.1. Comparación de Regiones bajo Selección con Genes asociados a Frío y Dieta

Se evaluó para ambos grupos de la Patagonia las distancias entre las ventanas con distribución sobre el 1% de los datos para los valores de PBS y aquellos genes asociados en estudios previos a frío y dieta. Esto con la finalidad de determinar la existencia de regiones cromosómicas que estuvieran ligadas o asociadas con genes bajo selección en diversas poblaciones, principalmente Siberia. En la Figura 43

vemos la distribución de las ventanas bajo selección solo para los cromosomas donde se encontraron previamente genes asociados a frío y dieta. A partir de esto podemos ver que si bien la mayoría de las ventanas se encuentran fuera del rango de los genes de interés, existen algunas regiones bajo selección asociadas a **THADA** (banda roja) en el cromosoma 2 y a **BDKRB2** (banda negra) en el cromosoma 14.

Luego de esta comparación, se procedió a evaluar específicamente estas regiones asociadas a **THADA** y **BDKRB2** pero esta vez asociando los valores de PBS para cada SNPs. Para el caso del gen **THADA** ubicado en el cromosoma 2 (43.2-43.5 Mb), vemos que existe una señal de selección en la región (40.8-41.2 Mb) que dista 2 Mb del gen y que pertenece al grupo compuesto por cazadores recolectores de la Patagonia. Mas aún, aparece un SNP (rs12613957) con un valor de PBS sobre el 1% y que se encuentra junto a **THADA** (Figura 44). Este gen se encuentra bajo una señal de selección en diversas poblaciones de Siberia, y que se encuentra jugando un rol en mecanismos de respuesta al frío (Cardona et al., 2014). Por otro lado, el gen **BDKRB2** que se encuentra asociado a la vasodilatación y vasoconstricción, ambos mecanismos relacionados con la respuesta al buceo en poblaciones nativa nómades de Indonesia (M. Ilardo & Nielsen, 2018). Se puede apreciar que para nuestros datos existe una ventana bajo selección ubicada en el cromosoma 14 (95.6-95.8 Mb) ubicada a solo 0.4 Mb de distancia de este gen, además es posible detectar un set de SNPs bajo selección que no habían sido identificado a través de las ventanas y que se encuentran mucho mas cerca de este gen (Figura 45).

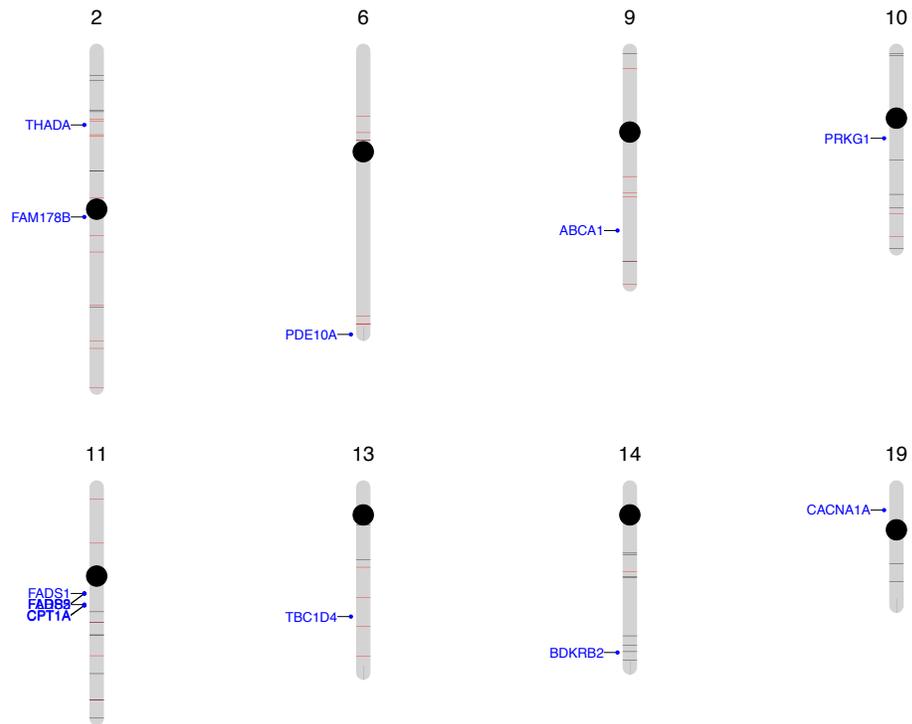


Figura 43. Comparación entre los genes asociados previamente a frío y dieta con las ventanas sobre el 1% de los valores de PBS. En azul se encuentran los genes asociados a frío y dieta descritos previamente. Las bandas horizontales dentro de cada cromosoma corresponden a ventanas de 200Kb con valores por sobre el 1% de los datos. Las bandas en color rojo corresponden a las señales para los cazadores recolectores de la Patagonia, mientras que las bandas negras corresponden a los grupos cazadores recolectores marítimos de la Patagonia.

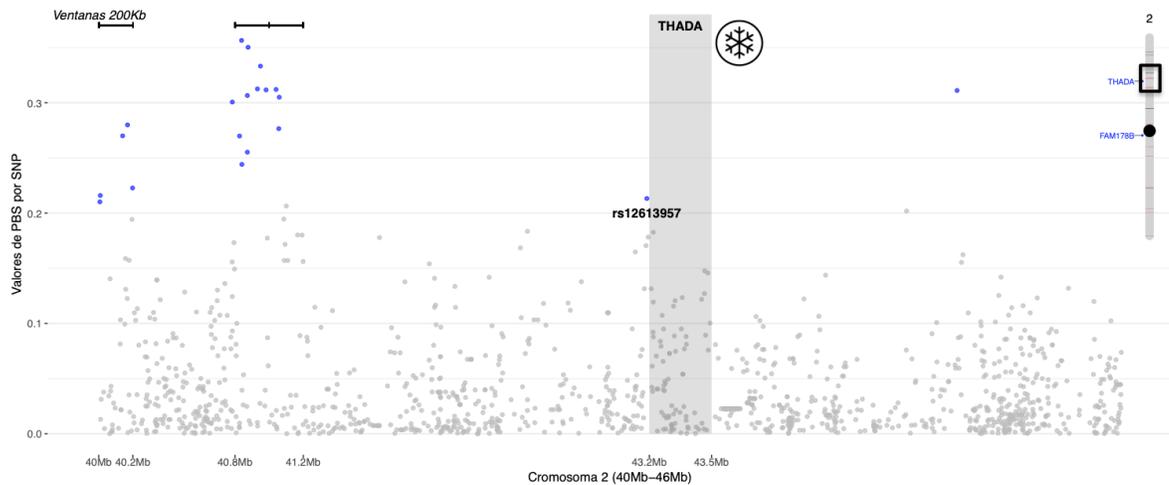


Figura 44. Distribución de los valores de PBS por SNPs entre las regiones 40-46Mb del cromosoma 2. La distribución de los valores por sobre el 1% de los datos esta dado por el color azul. El recuadro gris señala la región del gen THADA. Las líneas horizontales en el margen superior izquierdo muestra las ventanas de 200Kb que aparecen cercanas al gen de interés. Al costado derecho se muestra el cromosoma 2 con el recuadro en negro donde se realizó el zoom dentro del análisis. El símbolo al lado del gen THADA representa que esta asociado al frío.

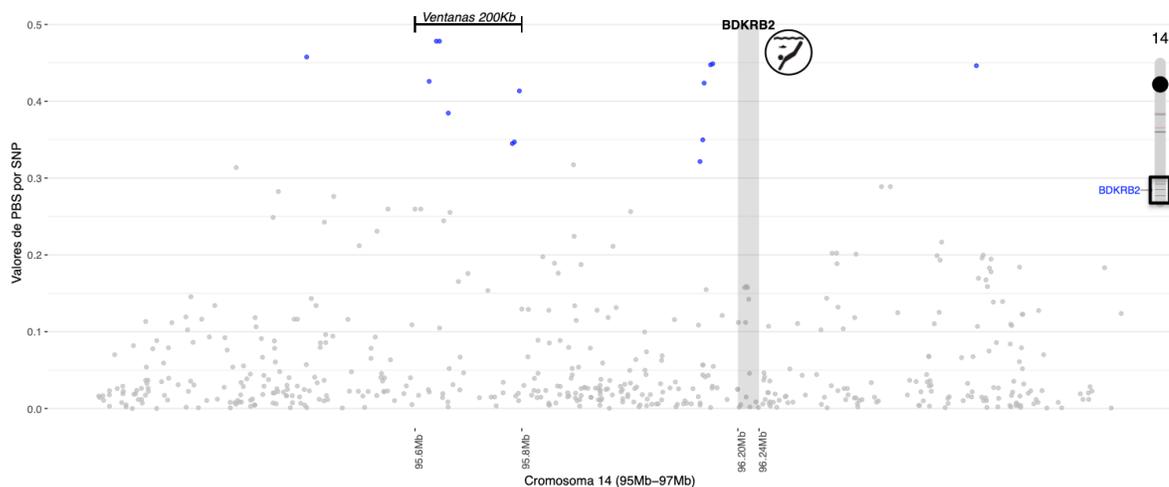


Figura 45. Distribución de los valores de PBS por SNPs entre las regiones 95-97Mb del cromosoma 14. La distribución de los valores por sobre el 1% de los datos esta dado por el color azul. El recuadro gris señala la región del gen BDKRB2. Las líneas horizontales en el margen superior muestra las ventana de 200Kb que aparece cercana al gen de interés. Al costado derecho se muestra el cromosoma 14 con el recuadro en negro donde se amplificó el análisis. El símbolo al costado del gen BDKRB2 indica que esta relacionado a buceo.

Puesto que este análisis se llevó a cabo únicamente mediante ventanas, se procedió a comparar si las señales bajo selección que se encontraron para los SNPs

se encuentran cerca de algunos de los genes descritos previamente para frío, dieta y buceo. Esto se realizó para los valores de PBS y separado para los escenarios CR (cazadores recolectores) y CRM (cazadores recolectores marítimos). El primer escenario lo podemos ver en la Figura 46 con aquellos genes que resultaron más interesantes de evaluar, destacando esencialmente los genes **FADS1** y **PRKG1** asociados a dieta marina y a frío respectivamente en poblaciones de Siberia (Cardona et al., 2014; Fumagalli et al., 2015). Estos SNPs parecen conformar una región cromosómica estrechamente ligada a estos genes, especialmente para **PRKG1**. También podemos ver un clúster de SNPs bajo selección próximos al gen **CPT1A**, relacionado con desarrollar procesos biológicos para el frío y dieta marina en diversas poblaciones de Siberia (Cardona et al., 2014; Clemente et al., 2014). Por último, es posible ver un set de tres SNPs bajo selección dentro del gen **CACNA1E**, asociado a la presión del flujo sanguíneo en poblaciones con estilos de vida relacionados a prácticas buceo (Ilardo et al., 2018).

Para los cazadores recolectores marítimos de la Patagonia (Figura 47) podemos ver que la región que comprende el cromosoma 11, especialmente el clúster conformado por los genes **FADS** también poseen una región compuesta por varios SNPs bajo selección muy próxima, mas aún, dos marcadores se encuentran dentro del gen **FADS3** asociado a dieta marina (Fumagalli et al., 2015). Además, podemos ver cinco SNPs que se encuentran relacionados con los genes **PDE10A** y un SNP dentro de **CACNA1A**, ambos involucrados en regulación del tamaño del timo y en la presión arterial en poblaciones con prácticas de buceo (Ilardo et al., 2018).

Interesantemente, vemos que para el cromosoma 11 existe una señal de selección en una región comprendida específicamente entre los 62.6-65.4Mb para los Cazadores Recolectores Marítimos y de 63.1-64.7Mb para el grupo Cazadores Recolectores de la Patagonia. Analizando los genes de esta región podemos encontrar varios asociados con el metabolismo de lípidos y de glicógeno. Esto incluye especialmente a los genes **PLAAT3**, **PLAAT4** y **PLAAT5**, entre otros.

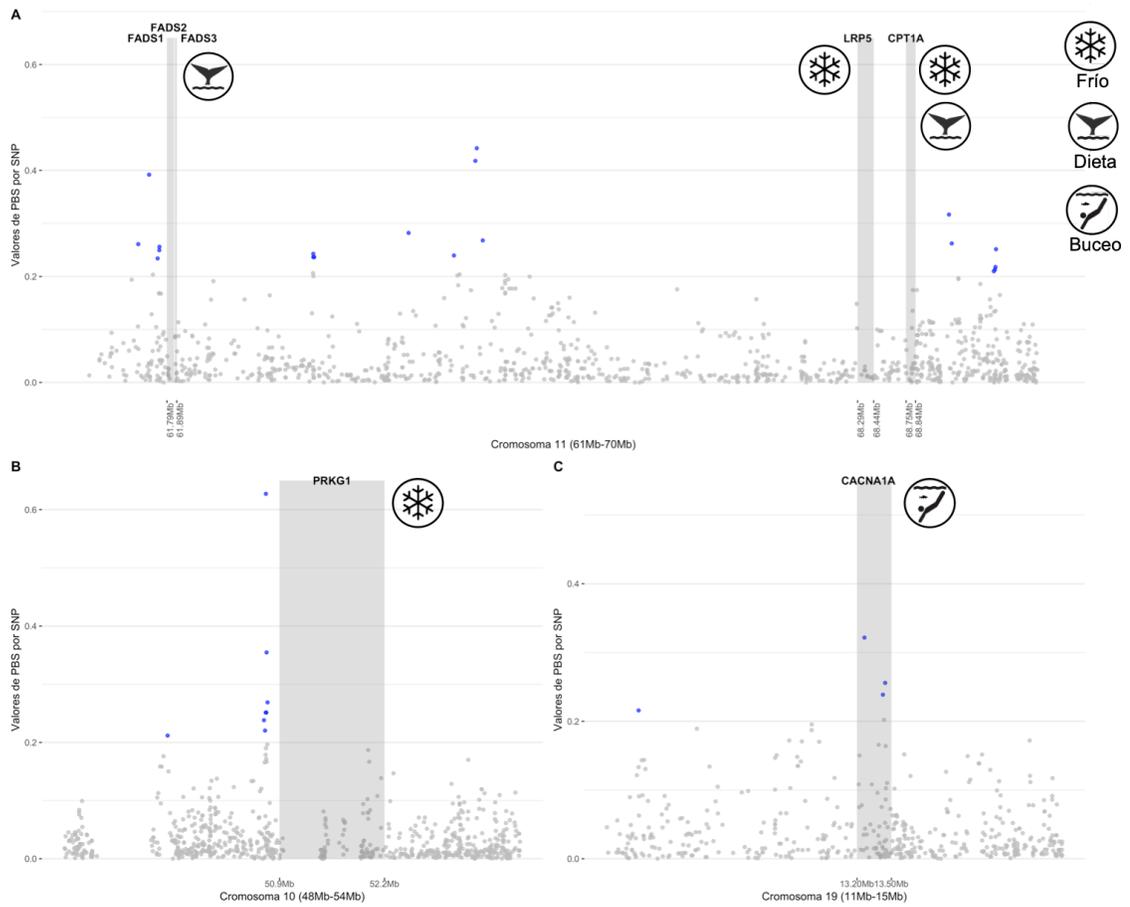


Figura 46. Comparación de valores de PBS por SNPs entre los cazadores recolectores de la Patagonia en regiones cromosómicas de genes descritos previamente en literatura para otras poblaciones nativas. Aquellos valores por sobre la distribución del 1% se encuentra en azul. El rectángulo gris corresponde al gen descrito a comparar. Cada símbolo al costado del gen corresponde a si este esta asociado a frío, dieta y/o buceo.

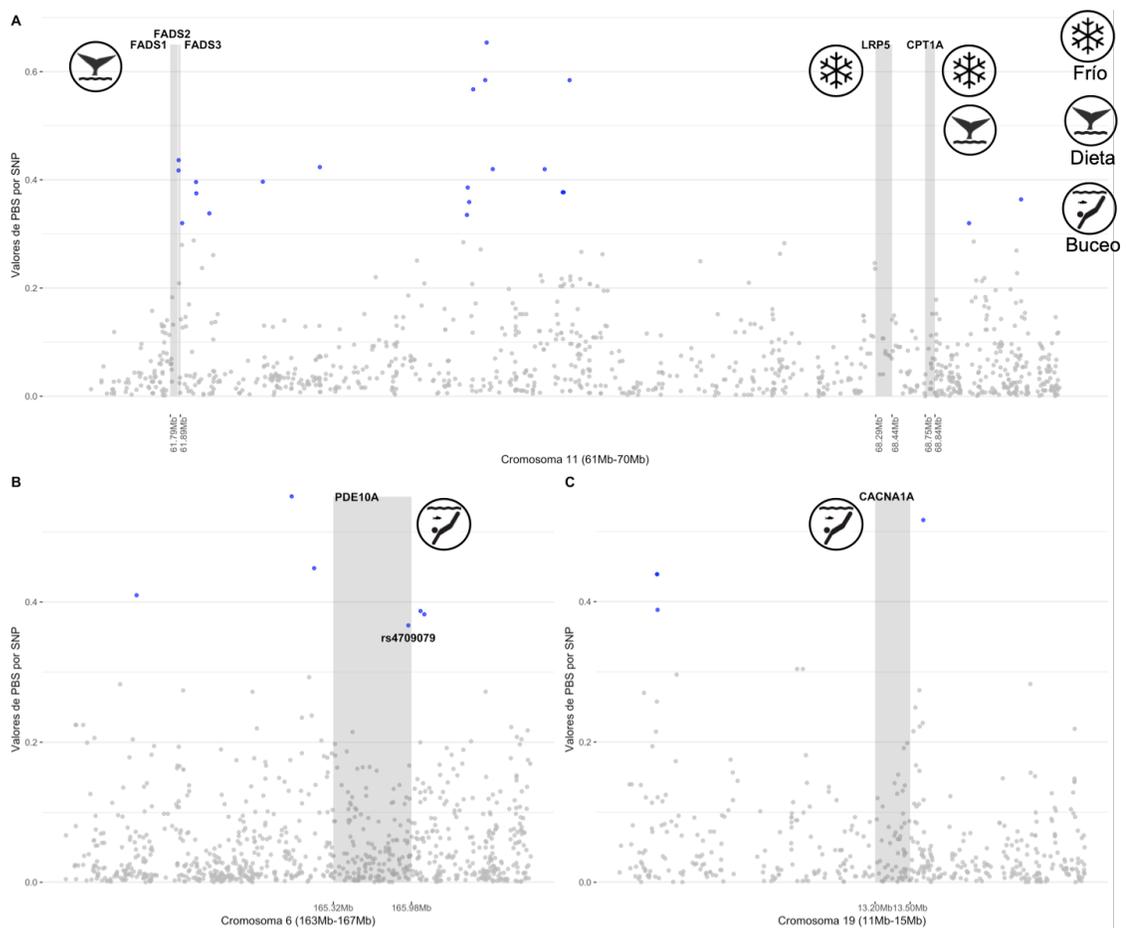


Figura 47. Comparación de valores de PBS por SNPs entre los cazadores recolectores marítimos de la Patagonia en regiones cromosómicas de genes descritos previamente en literatura para otras poblaciones nativas. Aquellos valores por sobre la distribución del 1% se encuentra en azul. El rectángulo gris corresponde al gen descrito a comparar. Cada símbolo al costado del gen corresponde a si este esta asociado a frío, dieta y/o buceo.

En lo que respecta a los análisis de iHS, podemos percatarnos que no hay ventanas bajo selección cercanos a los genes descritos para frío, dieta o buceo. De la misma manera, cuando queremos analizar la distribución de los SNPs con valores de $|iHS| > 2.5$ podemos darnos cuenta de que no hay datos para estos genes debido a que se removieron ciertas regiones del genoma durante el proceso de enmascaramiento del genoma con ancestría no nativa. En la Figura 48 se puede ver el resultado del proceso de enmascaramiento para el análisis de iHS en los cromosomas 2, 10 y 11.

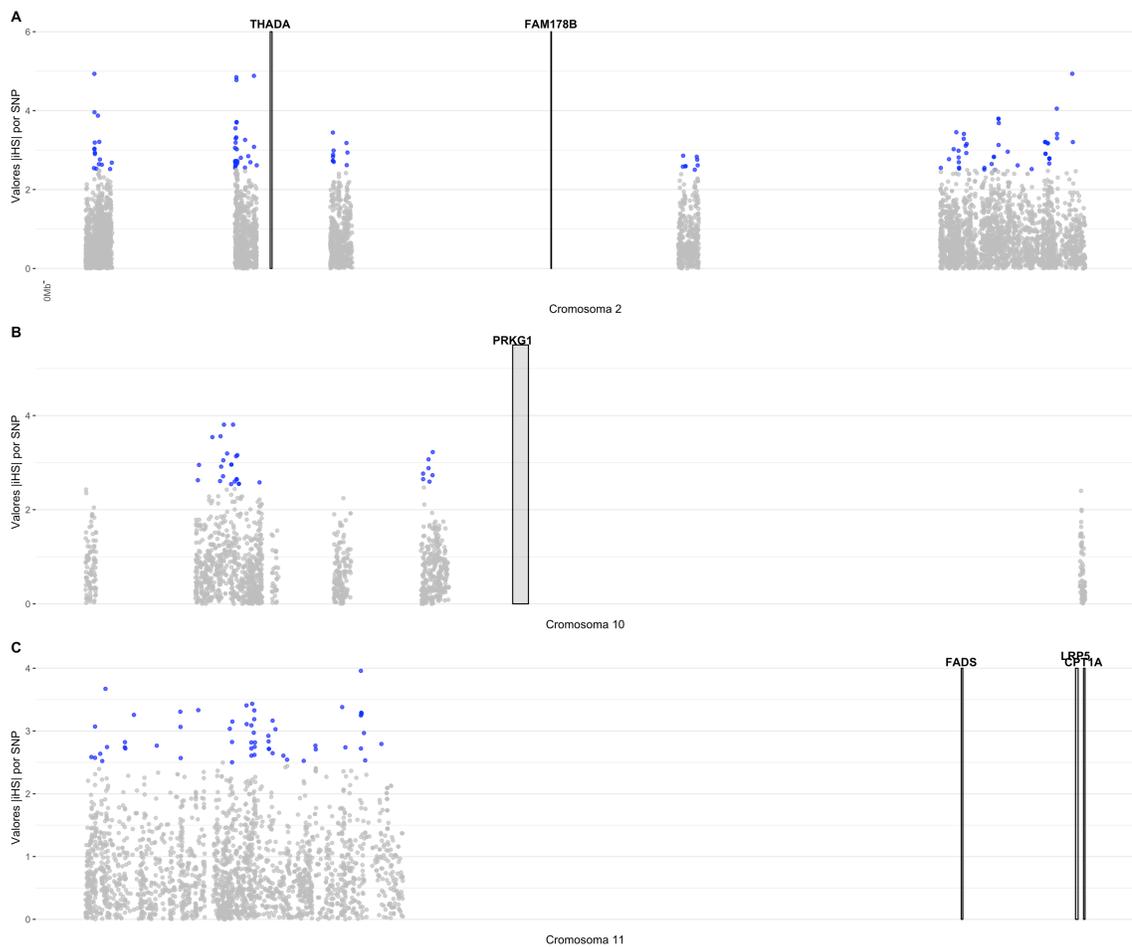


Figura 48. Análisis de iHS para SNPs en grupos Cazadores Recolectores de la Patagonia. Los puntos de color azul corresponden a los valores de $|iHS| > 2.5$. Se muestran únicamente los cromosomas 2, 10 y 11 donde se ubican genes descritos previamente en Siberia bajo selección para frío y dieta. Los recuadros negros verticales corresponden a la ubicación física de estos genes.

5.7.2. Procesos y Rutas Metabólicas en la Patagonia

Dentro de las categorías que más llaman la atención para los grupos cazadores recolectores de la Patagonia, destaca la síntesis y excreción del cortisol, un glucocorticoide relacionado con el metabolismo de glucosa y lípidos, metabolismo homeostático y respuesta al estrés. La distribución de los genes que se encuentran asociados a esta ruta y por sobre el 1% de los valores de PBS (por SNPs) están relacionados principalmente con el cromosoma 1 y 2 (Tabla XXVII). Esta ruta

metabolica también se encuentra en el grupo de los cazadores recolectores de la Patagonia, no obstante sin una corrección a posterior significativa (FDR > 0.05). En la Figura 49 vemos la descripción de la ruta metabolica del cortisol, donde en los recuadros rojos aparecen los genes que fueron encontrados bajo selección en nuestros datos.

Tabla XXVII. Genes involucrados en la ruta metabolica de síntesis y secreción de cortisol en los individuos de la Patagonia.

Gen	Nombre Gen	Cromosoma	Inicio	Termino
HSD3B2	hydroxy-delta-5-steroid dehydrogenase, 3 beta-and steroid delta-isomerase 2	1	119414931	119423039
HSD3B1	hydroxy-delta-5-steroid dehydrogenase, 3 beta-and steroid delta-isomerase 1	1	119507203	119515058
PBX1	PBX homeobox 1	1	164559360	164886047
ADCY3	adenylate cyclase 3	2	24819169	24920017
POMC	proopiomelanocortin	2	25160853	25168851
KCNK3	potassium two pore domain channel subfamily K member 3	2	26692713	26733420
CREB1	cAMP responsive element binding protein 1	2	207529892	207605989
ADCY2	adenylate cyclase 2	5	7396230	7830081
CREB5	cAMP responsive element binding protein 5	7	28299321	28825894
CREB3L2	cAMP responsive element binding protein 3 like 2	7	137874979	138002101
KCNA4	potassium voltage-gated channel subfamily A member 4	11	30009741	30017030
NR4A1	nuclear receptor subfamily 4 group A member 1	12	52022832	52059507
SCARB1	scavenger receptor class B member 1	12	124777628	124863973
ADCY9	adenylate cyclase 9	16	3953387	4116185
GNA11	G protein subunit alpha 11	19	3094410	3124002
PLCB1	phospholipase C beta 1	20	8132265	8884900
PLCB4	phospholipase C beta 4	20	9068710	9480816

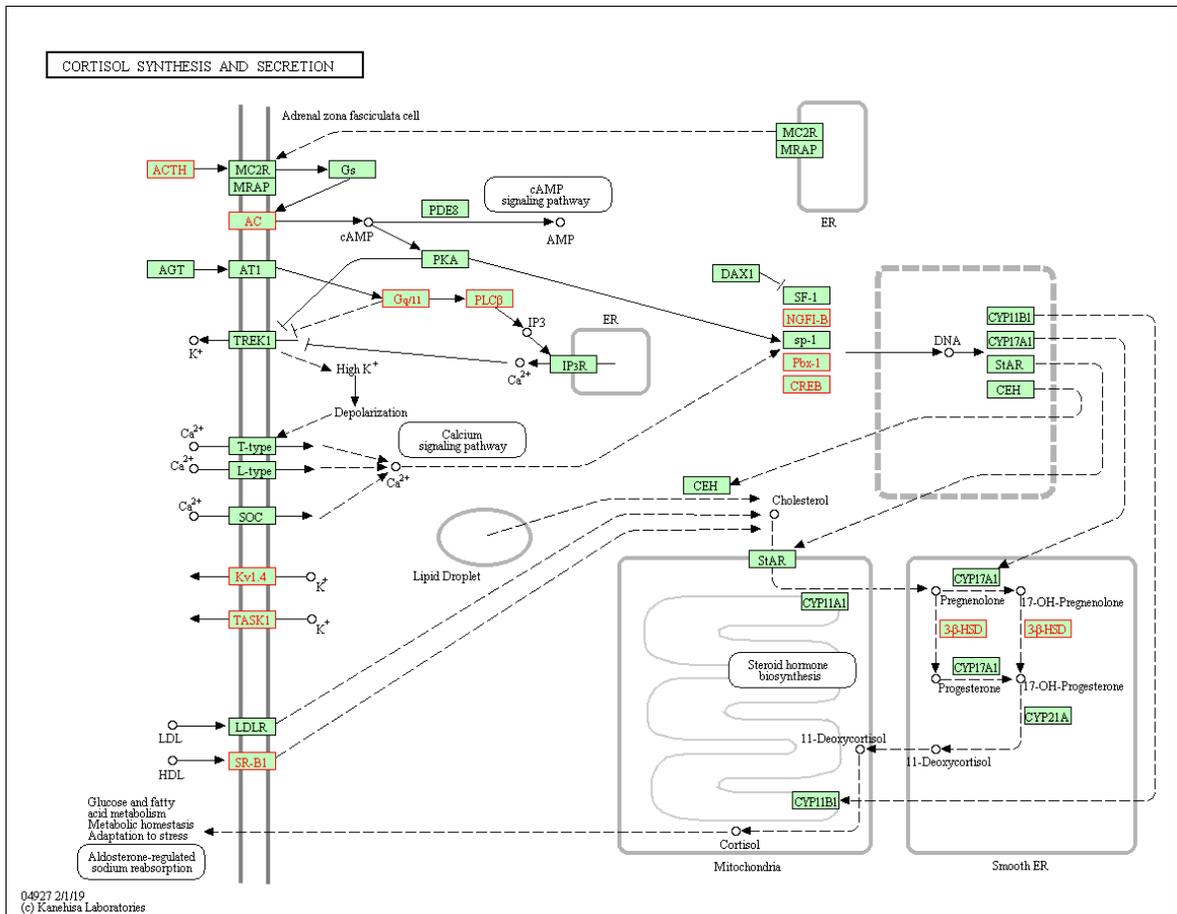


Figura 49. Ruta metabólica descrita para la síntesis y secreción del cortisol (hsa04927) extraído a partir del Mapa de rutas KEGG (www.kegg.jp). Los recuadros en rojo corresponden a los genes sobre-representados en los cazadores recolectores de la Patagonia.

Otra de las rutas que aparecen dentro del análisis de PBS por ventanas de 200 Kb corresponde a la termogénesis, que a pesar de no ser significativa a posteriori (FDR > 0.05), es de vital importancia para nuestros datos, dado que este proceso permite convertir la energía en calor en los adipocitos (Tabla XXVIII). En la Figura 50 se muestra la ruta metabólica relacionada con la termogénesis, donde los recuadros en rojo corresponden a los genes que se encuentran sobre-representados en los cazadores recolectores marítimos de la Patagonia, estos se ubican en los complejos de la cadena mitocondrial, así como también como receptores de membrana celular y dentro del núcleo.

Tabla XXVIII. Genes involucrados en la ruta metabólica de termogénesis en los individuos de la Patagonia.

Gen	Nombre Gen	Cromosoma	Inicio	Término
PRDM16	PR/SET domain 16	1	3068227	3438621
ADCY1	adenylate cyclase 1	7	45574140	45723116
SMARCA2	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 2	9	2015219	2193624
SLC25A29	solute carrier family 25 member 29	14	100277986	100306547
NDUFA3	NADH:ubiquinone oxidoreductase subunit A3	19	54102338	54106974

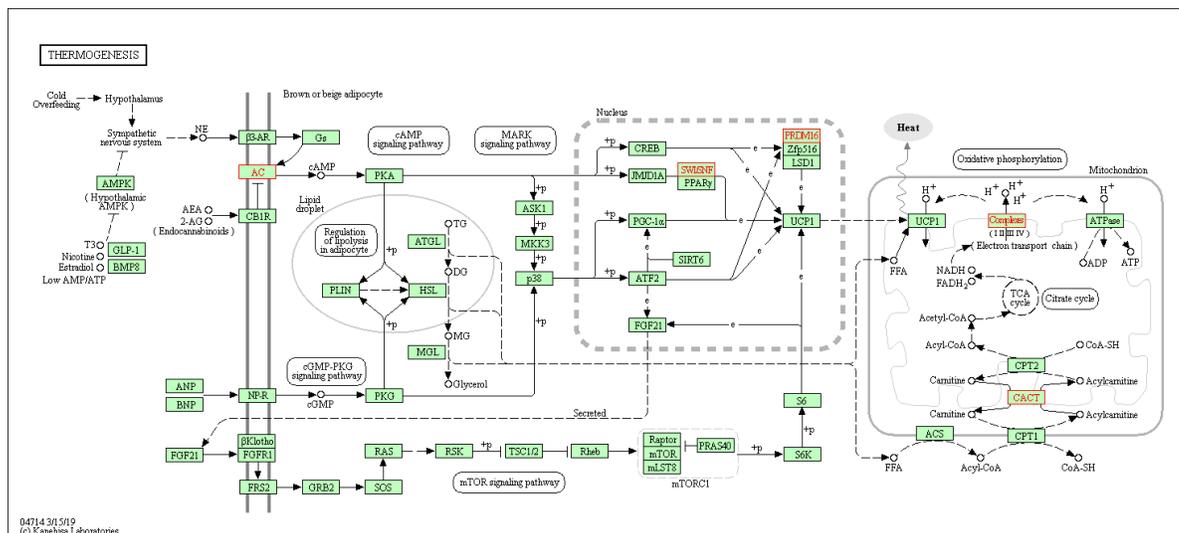


Figura 50. Ruta metabólica descrita para la termogénesis (hsa04714) extraído a partir del Mapa de rutas KEGG (www.kegg.jp). Los recuadros en rojo corresponden a los genes sobre-expresados en los cazadores recolectores marítimos de la Patagonia.

Otro de los genes que aparecen en varios procesos biológicos (FDR > 0.05) directamente asociadas con el metabolismo y degradación de ácidos grasos corresponde a **ACADM**. Este gen codifica para la enzima Acyl-Coenzima A deshidrogenasa y participa en el primer paso de la beta-oxidación mitocondrial de ácidos grasos.

5.8. Análisis sobre Cazadores Recolectores Marítimos en la Patagonia Austral

En esta parte de los resultados nos enfocaremos en una parte específica de la Patagonia, específicamente en Villa Ukika, una localidad ubicada en la Región de Magallanes y la Antártica Chilena. Estos individuos son descendientes de los grupos Kawéskar que habitaban la zona y fueron genotipificados mediante el microarreglo de MEGA (Illumina) para un total de 1.476.122 SNPs distribuidos a lo largo del genoma. Se consideraron también cinco individuos de Villa Ukika descendientes de Yámana, dos individuos Tehuelche, un individuo de Punta Arenas (Kawéskar) y seis de Laitec (Chono), los cuales fueron genotipificados en ambas plataformas. De acá en adelante, por lo tanto exploraremos su diversidad genética y evaluaremos posibles señales de selección de forma más específica en los grupos cazadores recolectores marítimos de la Patagonia Occidental.

5.8.1. Análisis de Estructuración Poblacional mediante Ancestrías

5.8.1.1. Ancestría Global

Se evaluó la ancestría global a través de ADMIXTURE (Alexander & Lange, 2011) sobre las poblaciones de la Patagonia Austral, utilizando para esto las mismas poblaciones de referencia continental (YRI, IBS, EAS, Puno y Pehuenche/Huilliche) para diferentes valores de poblaciones ancestrales (K). En la Figura 51A podemos ver el error de validación cruzada, el cual muestra que los valores más bajos corresponden a un K4,5 y 6. Luego de esta primera proyección para la ancestría global se realizó el coeficiente de parentesco mediante ancestría a través de REAP, el cual arrojó múltiples comparaciones por sobre el 12.5% entre individuos dentro de la localidad de Puerto Edén. Es por esto que se realizaron dos ADMIXTURE con sub-paneles filtrados por REAP, y luego se procedieron a unir en una sola figura (Figura 52). El error de validación cruzada controlado por REAP (Figura 51B) revela que los valores de K siguen siendo más bajos para 4, 5 y 6 grupos ancestrales.

Con un K=4 vemos que los principales grupos continentales corresponden a los Africanos, Asiáticos, Europeos y Amerindios. Las muestras de la Patagonia Austral posee un 93.5% de componente Amerindio, y un 6.5% de componente Europeo, siendo los Kawéskar de Punta Arenas los que mayor poseen este componente (16.4%) (Tabla XXIX). Cuando vemos el K=5 se puede distinguir el quiebre entre un componente Andino (Aymara, Puno) y un componente propio del centro sur de Chile (Pehuenche y Huilliche). Bajo este panorama podemos evidenciar una clara participación del componente centro sur de Chile en las localidades de Puerto Edén y Punta Arenas, mientras que el componente Andino se encuentra en una mayor proporción en los individuos de Puerto Edén, en comparación con los de Punta Arenas. Por último, un K=6 nos permite aproximarnos aparentemente a un nuevo componente Amerindio, esta vez de origen Austral, puesto que aparentemente se encuentra representado únicamente por las localidades de Puerto Edén, Punta Arenas y Villa Ukika. De estas, las localidades con ascendencia Kawéskar (Puerto Edén y Punta Arenas) poseen un 51% de este componente, mientras que los grupos Yámana tienen un 74%. Los dos individuos Tehuelches en este set no poseen este componente y solo un individuo de Laitec (Chono) posee un 11.1% de esta componente Austral (Tabla XXIX).

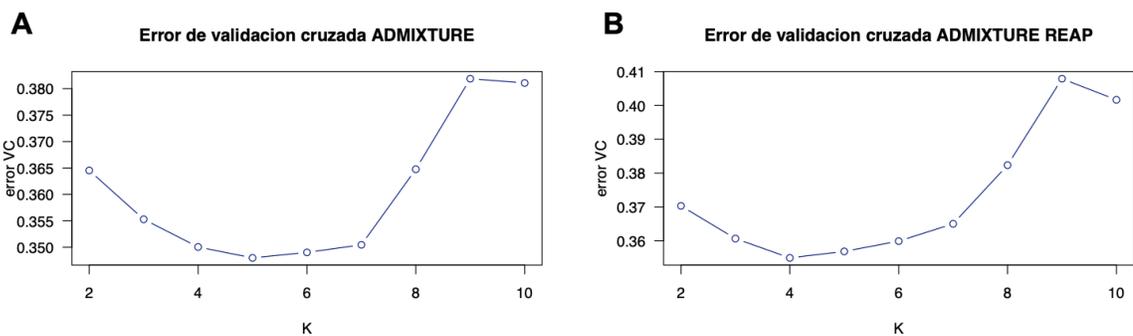


Figura 51. Error de validación cruzada para el análisis de ADMIXTURE (cross-validation, CV). En A se muestra el CV para todos los individuos de la muestra, mientras que en B se muestra el CV posterior al filtrado de individuos con un coeficiente de parentesco mayor a 0.125 mediante REAP.

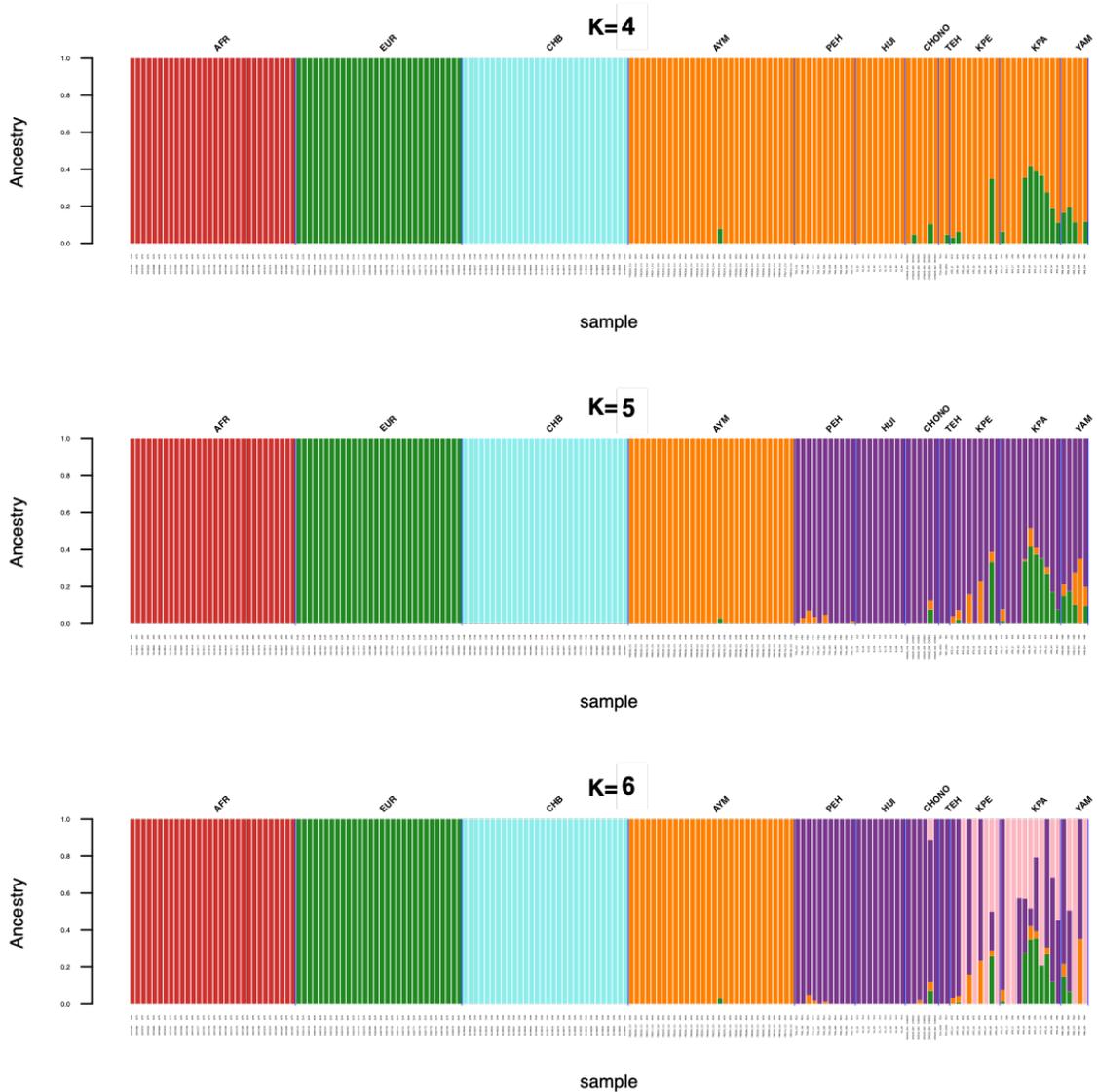


Figura 52. Análisis de Ancestría Global sobre las poblaciones de la Patagonia genotipificados en MEGA. En la figura se observan solo las ancestrías con un K=4, 5 y 6. Los paneles de referencia corresponden a Africa (YRI), Europa (IBS), Este de Asia (CHB), Aymara (Puno), Pehuenche (PEH) y Huilliche (HUI). Este análisis se llevó a cabo mediante la fusión de los dos subset filtrados por REAP.

Tabla XXIX. Comparación entre promedios de ancestrías por población para ADMIXTURE con filtro por parentesco . En paréntesis se encuentran los errores estándares para cada ancestría.

ADMIXTURE, K=6						
	África	Europa	Asia	Aymara	Sur Chile	Patagonia Austral
Chono	0,00001(0)	0,00577(0,00576)	0,00001(0)	0,00595(0,00594)	0,94700(0,01165)	0,04123(0,00588)
Tehuelche	0,00001(0)	0,01302(0,01301)	0,00001(0)	0,06055(0,00306)	0,91494(0,01691)	0,01146(0,00083)
Kawéskar	0,00001(0)	0,12504(0,03491)	0,00001(0)	0,01705(0,00760)	0,31934(0,07276)	0,53851(0,08999)
Yámana	0,00001(0)	0,05474(0,02270)	0,00088(0)	0,19537(0,02237)	0,42592(0,04735)	0,32306(0,04703)
Total	0,00001	0,08591	0,00014	0,04469	0,48572	0,38351
ADMIXTURE REAP, K=6						
	África	Europa	Asia	Aymara	Sur Chile	Patagonia Austral
Chono	0,00001(0)	0,01217(0,01216)	0,00001(0)	0,01121(0,00798)	0,95800(0,03820)	0,01858(0,01857)
Tehuelche	0,00001(0)	0,00400(0,00399)	0,00001(0)	0,03537(0,00162)	0,96058(0,00561)	0,00001(0)
Kawéskar	0,00001(0)	0,10391(0,02982)	0,00001(0)	0,03440(0,01374)	0,35252(0,07456)	0,50913(0,08837)
Yámana	0,00001(0)	0,00001(0)	0,00001(0)	0,07022(0,07021)	0,19331(0,12943)	0,73643(0,19410)
Total	0,00001	0,06543	0,00001	0,03567	0,47534	0,42352

5.8.1.2. Análisis de Componentes Principales

Se realizó el análisis de componentes principales a través de smartPCA (N. Patterson et al., 2006) sobre la muestra de la Patagonia Austral con la finalidad de explorar la diversidad presente en estas localidades. Para esto se incluyeron los mismos grupos de referencia utilizados para la ancestría global: Europeos (n=30), Africanos (n=30), Asiáticos (n=30), Aymara (Puno, n=30) y Pehuenches/Huilliches (n=22). En la Figura 53 se proyectaron los dos primeros componentes principales, el primero de estos permite explicar el 6.9% de la variabilidad de los datos y separa el grupo Africano del resto, mientras que el segundo componente (4.3%) distingue la variabilidad observada principalmente entre los grupos Europeos y Amerindios (Aymara y Pehuenche/Huilliche). Los individuos de la Patagonia Austral se encuentran formando un gradiente entre estos dos grupos ancestrales mencionados. En orden de explorar mas detalladamente la variabilidad en la Patagonia Austral, se procedió a remover del análisis a los grupos Africanos y Asiáticos dada su escaso aporte ancestral en la muestra, y se realizó un nuevo Análisis de Componentes

Principales (**Figura 54**). En esta figura podemos ver que el primer componente (8.7%) permite separar la ancestría Europea de la Amerindia, donde si bien existen algunos individuos inclinados hacia un componente Europeo, la mayoría está plenamente agrupados dentro del componente Amerindio. El segundo componente principal (2.1%) distingue claramente la ancestría Andina, representado por los Aymara y la ancestría Centro Sur de Chile compuesto por los Pehuenches y Huilliches. En este punto se puede apreciar el mayor grado de variabilidad dentro de la Patagonia, donde los individuos principalmente asociados a Chono y los Tehuelches (Patagonia Norte) se encuentran estrechamente asociados con grupos Huilliches, mientras que los grupos asociados a la Patagonia Austral (Kawéskar y Yámana) aparecen conformando una región distinta y mas alejada del Centro Sur de Chile.

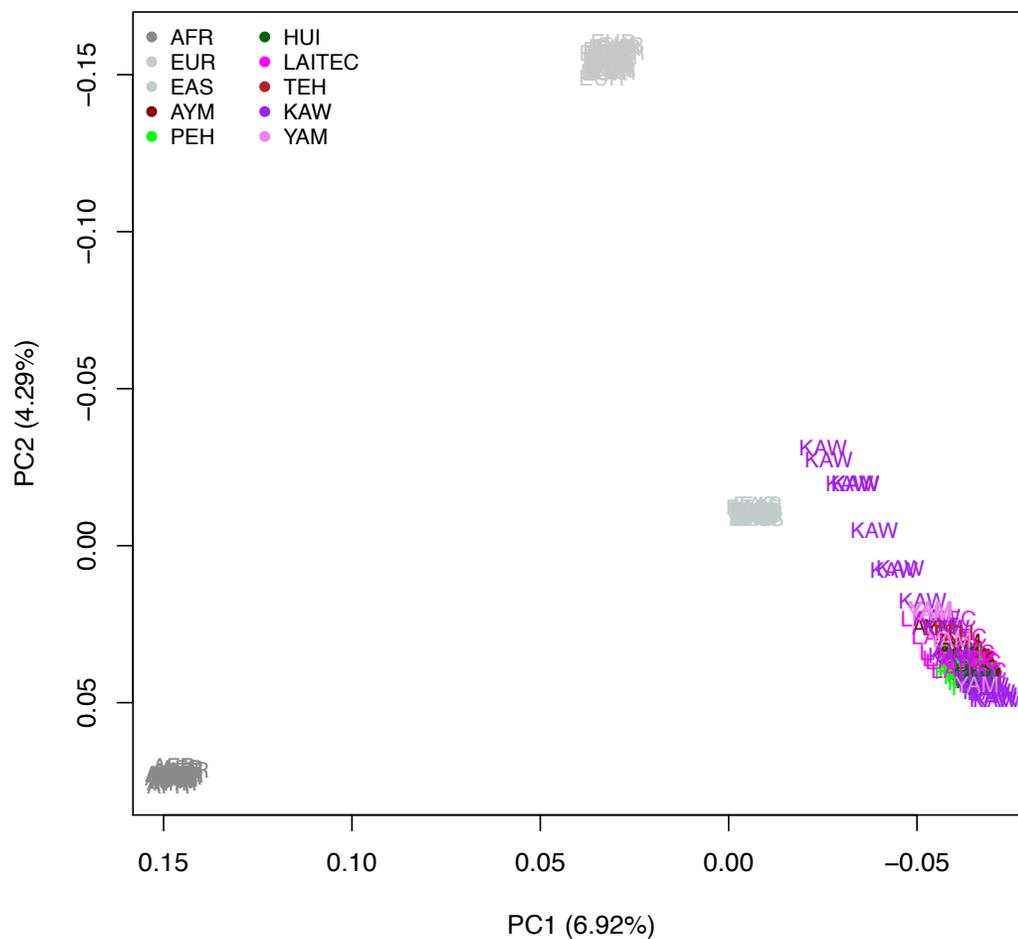


Figura 53. Análisis de Componentes Principales para las muestras de la Patagonia Austral. Se muestran los dos primeros componente, y se utilizaron como referencia a grupos Africanos (AFR), Europeos (EUR), Asiáticos (EAS), Aymara (AYM), Pehuenche (PEH) y Huilliche(HUI). Las localidades analizadas corresponden a Laitec (CHONO) pertenecientes a grupos Chono, Kawéskar (KAW) representado por Puerto Edén y Punta Arenas, Tehuelches (TEH) y Yámanas (YAM).

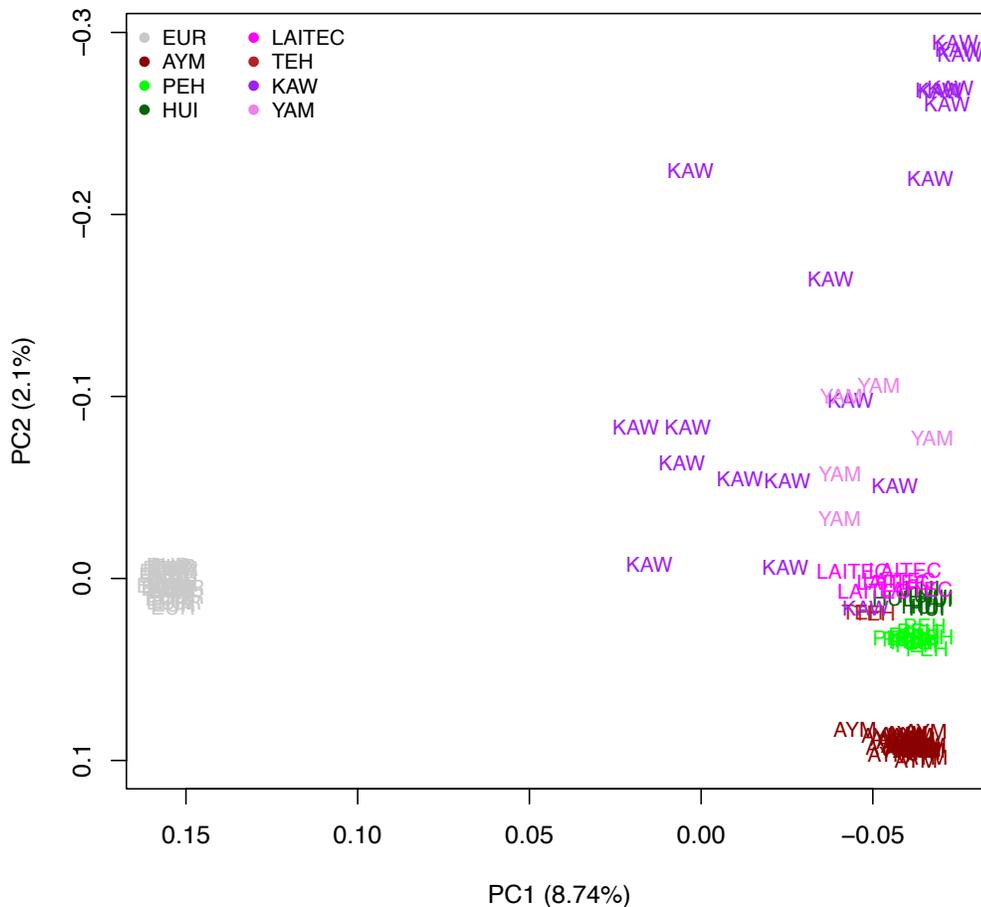


Figura 54. Análisis de Componentes Principales para las muestras de la Patagonia Austral. Se muestra los dos primeros componentes, y se removieron los grupos de referencia de origen Africano y Asiático, manteniendo únicamente a los Europeos (EUR), Aymara (AYM), Pehuenches (PEH) y Huiliches (HUI). Las localidades analizadas corresponden a Laitec (CHONO), Tehuelches (TEH), Kaweskar (KAW) y Yámana (YAM).

5.8.1.3. Análisis de Ancestría Local

Luego de estimar la variabilidad de nuestros datos y determinar los principales grupos ancestrales que están influyendo en las muestras de la Patagonia Austral, se procedió a inferir la Ancestría Local sobre el genoma de estos individuos con la finalidad de poder profundizar en la asignación de segmentos del genoma a una ancestría específica sobre estas localidades. Para esto, se determinó

previamente el faseo cromosómico mediante SHAPEIT2 (Delaneau et al., 2012), utilizando las poblaciones de 1000 Genomas como referencia. Luego, se llevo a cabo la ancestría local mediante RFMix (Maples et al., 2013b) usando tres grupos de referencia: Africano (YRI), Europeo (IBS) y Amerindio (Pehuenche/Huilliche). Finalmente, vamos a disponer para cada individuo de sus ancestrías específicas para cada región de su genoma, tal como se puede apreciar en la Figura 55 para un individuo de la localidad de Puerto Edén (Kawéskar).

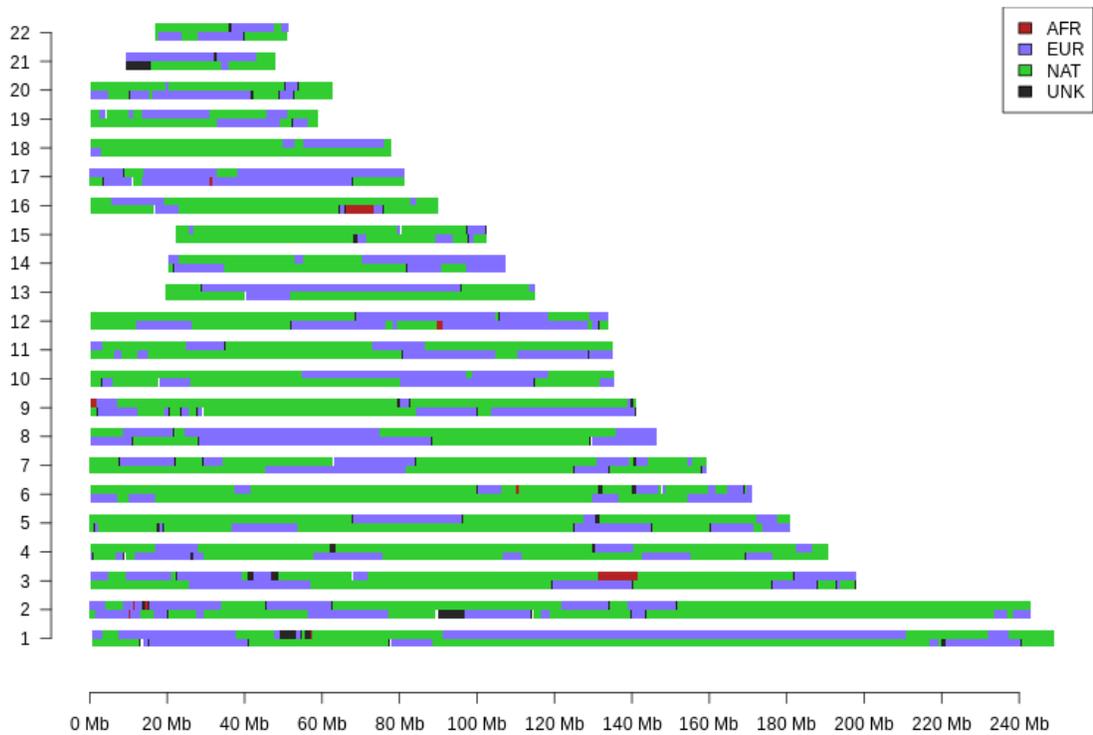


Figura 55. Análisis de Ancestría Local sobre el genoma de un individuo. En el eje vertical se disponen los 22 cromosomas, mientras que en el eje horizontal se grafica el tamaño de cada cromosoma. Las ancestrías inferidas corresponden a la Africana (rojo), Europea (morado) y Nativos (verde). Aquellas regiones del genoma que no pudieron ser asignadas correctamente a alguna ascendencia fueron dejadas como 'desconocido' (negro).

En orden de evaluar la correcta inferencia de las ancestrías entre las aproximaciones globales (Admixture) y locales (RFMix) se procedió a realizar una correlación entre estos dos métodos para las ancestrías Amerindias, Europeas y

Africanas. La Figura 56 muestra una correlación superior al 98% tanto para la ascendencia Europea como la Amerindia, mientras que la Africana posee una correlación del 48%, siendo la más baja, y probablemente debido a la escasa proporción de esta ascendencia en nuestras muestras.

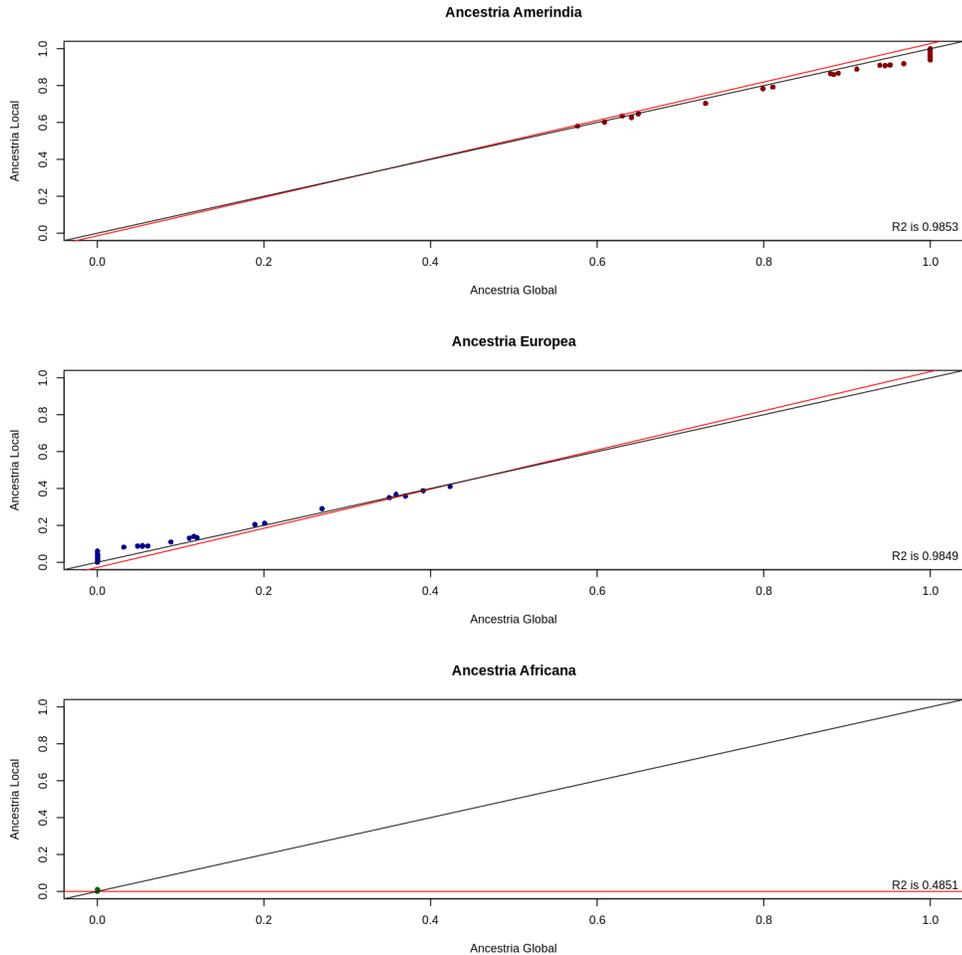


Figura 56. Test de Correlación para la inferencia de Ancestría Global y Local en las poblaciones de la Patagonia Austral. En el gráfico superior se muestra la ascendencia Amerindia, el gráfico del medio la ascendencia Europea, mientras que el gráfico inferior la ascendencia Africana.

Debido a que en estas muestras se volvió a identificar un componente de origen Aymara, se procedió a realizar una estimación de Ancestría Local, esta vez removiendo el componente Africano –dado su bajo componente- y reemplazarlo por esta ascendencia tomando a los individuos de Puno como referencia. Posterior a esto realizó el test de correlación para estimar la afinidad y precisión por parte de RFMix

para distinguir la variabilidad dentro del componente Amerindio en el genoma de un individuo. En la **Figura 57** podemos ver que la ancestría Europa no se ve afectada dado que posee una correlación del 96% entre ambos métodos. Sin embargo, para las ancestrías de origen Amerindio, vemos que las correlaciones caen en un 86 y 65% para los Mapuche y Aymara respectivamente.

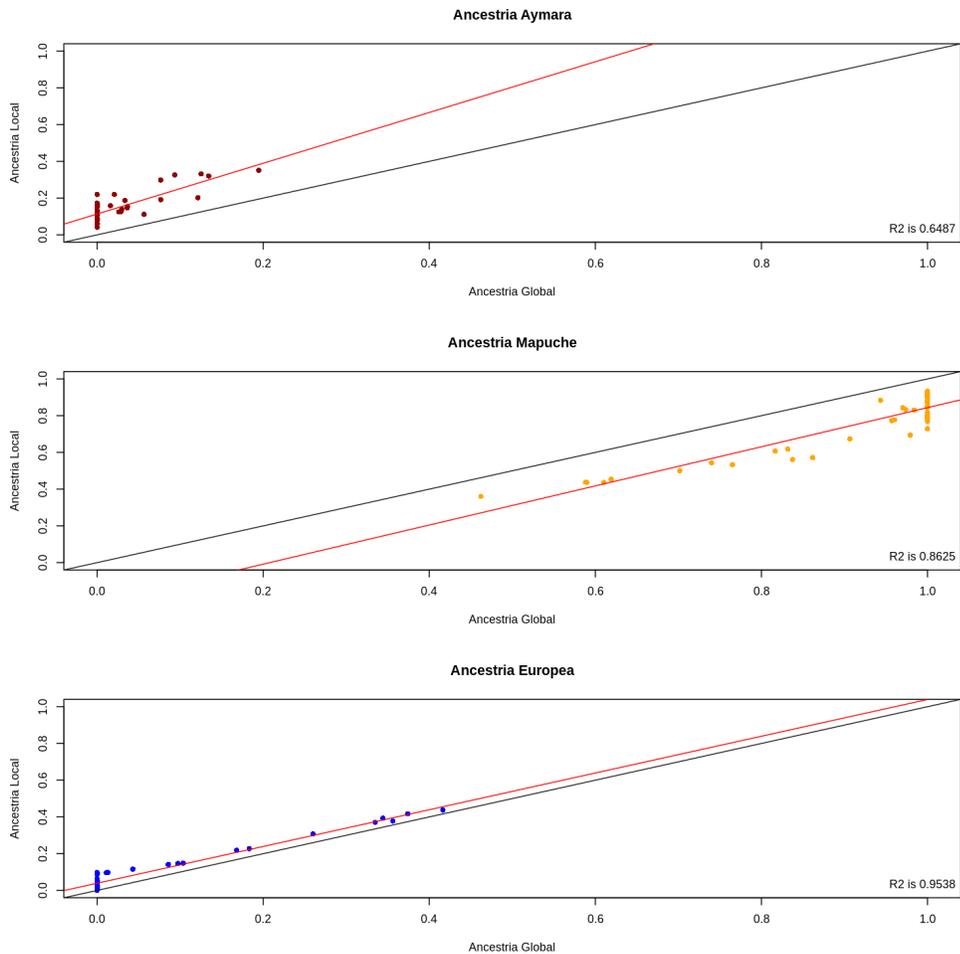


Figura 57. Test de Correlación para la inferencia de Ancestría Global y Local en las poblaciones de la Patagonia Austral. En el gráfico superior se muestra la ancestría Centro Sur de Chile, el gráfico del medio la ancestría Aymara, mientras que el gráfico inferior la ancestría Europea.

5.8.1.4. Modelamiento de los procesos de mestizaje en la Patagonia

Se analizaron las longitudes de cada segmento dentro del genoma de un individuo mediante el análisis de TRACTs (Gravel, Zakharia, Moreno-Estrada, et al., 2013) con el propósito de determinar el tiempo inicial de mezcla para diversas

ancestrías, proporciones de mezcla inicial y sus cambios en el tiempo debido a eventos de migración. Para este caso, solo se evaluó el modelo de mezcla inicial con un pulso posterior (**ppx_xxp**). Este modelo se llevó a cabo para diferentes grupos de ancestrías; por un lado se evaluó la interacción entre los componentes Europeo, Amerindio y Africano, mientras que por otro lado se removi6 este último componente y se reemplazó por un grupo de origen Aymara, esto con la finalidad de analizar la variabilidad dentro del componente Amerindio previo al contacto Europeo.

5.8.1.4.1. Componentes Europeo, Nativo y Africano

Tal como podemos ver en la **Figura 58**, los tiempos de mezcla inicial entre el componente Europeo y Amerindio va de 8 a 12 generaciones atrás (considerando 25 años por generación), siendo las poblaciones Yámana y Kawéskar con menores tiempo de mezcla (8 y 9 respectivamente). En cambio, podemos apreciar que tanto los individuos de Chiloé como Tehuelches poseen un tiempo inicial de 11 y 12 generaciones respectivamente. Con respecto al componente Europeo, podemos decir que se encuentra mas frecuente en las poblaciones Kawéskar con un 15% aproximadamente.

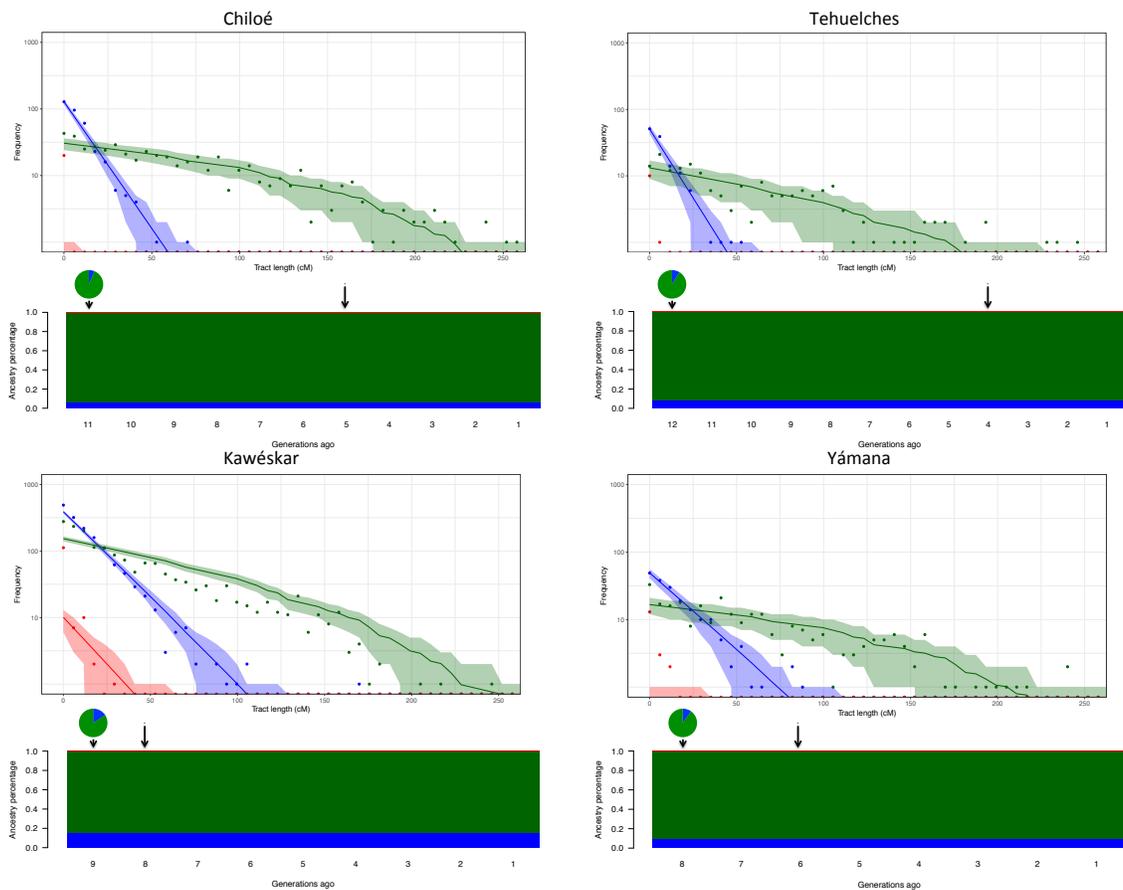


Figura 58. Análisis de TRACTs y tiempos de mestizaje en las poblaciones de la Patagonia para las ancestrías Europeas, Amerindias y Africanas. Se muestran los modelos de TRACTs que mejor se ajustan para los datos observados en las poblaciones Chono, Tehuelches, Kawéskar y Yámana. El decaimiento de la ancestría (arriba) compara la distribución de las longitudes de tracts observados y esperados basado en el modelo de migración propuesto. El área sombreada corresponde al 68,7% del intervalo de confianza. Los puntos indican los valores observados. El modelo de migración (parte inferior del gráfico) muestra el cambio en las proporciones de ancestría en el tiempo. De izquierda a derecha se observa la escala en generaciones, los círculos en la parte superior indican la ocurrencia de los eventos de mezcla, siendo el tamaño del círculo la magnitud de los migrantes que llegaron de cada población ancestral. Los colores verde, azul y rojo representan a las ancestrías Nativa, Europea y Africana respectivamente.

5.8.1.4.2. Componentes Aymara, Mapuche y Europeo

Como se mencionó anteriormente, se evaluó la interacción de los componentes Aymara y Mapuche previo al contacto Europeo mediante TRACTs (Gravel, Zakharia, Moreno-Estrada, et al., 2013). La Figura 59 revela que los tiempos de mezcla entre las ancestrías Aymara y del Centro Sur de Chile son mayores en comparación con el análisis anterior. Este rango va desde incluso 18 generaciones atrás (450 años) en individuos de Chiloé. Para el caso de los individuos

pertenecientes a los Kawéskar se puede ver una mezcla hace 15 generaciones atrás (375 años). Por último, el tiempo de llegada de los Europeos a la Patagonia Austral se vio reflejado hace no más de 9 y 8 generaciones para los Kawéskar y Yámana respectivamente, coincidiendo bastante con lo que arrojó el análisis anterior.

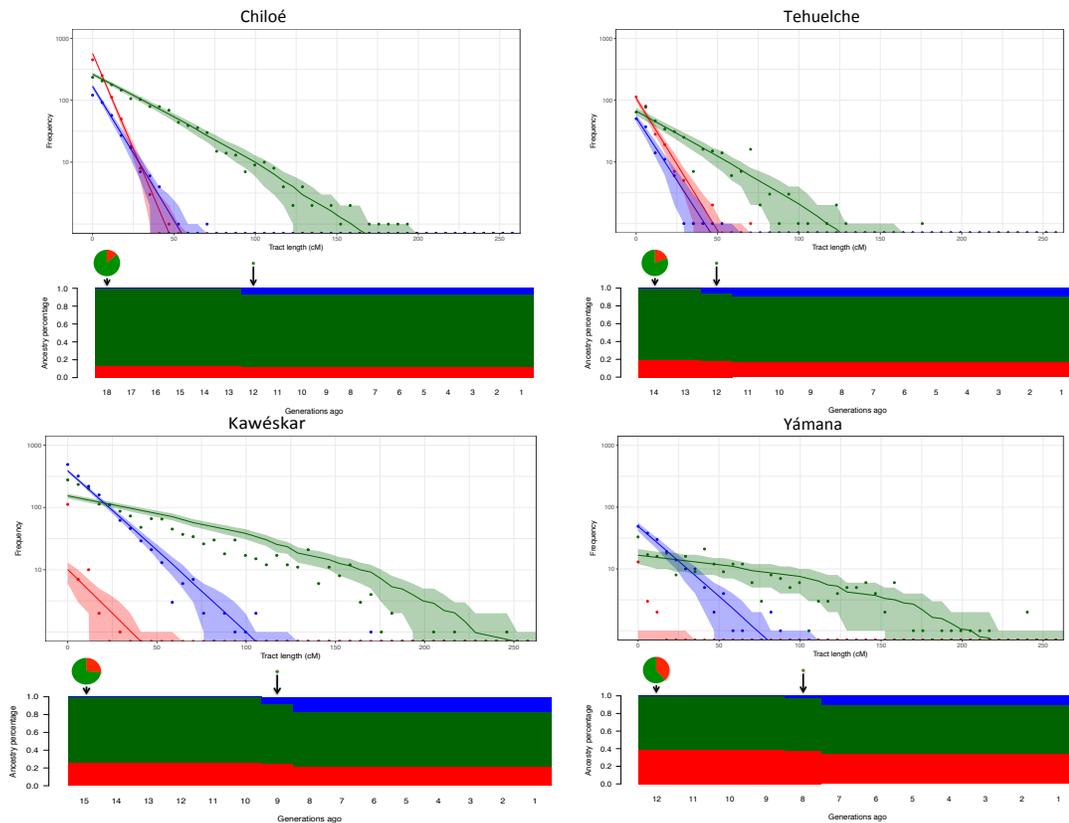


Figura 59. Análisis de TRACTs y tiempos de mestizaje en las poblaciones de la Patagonia para las ancestrías Aymara, Centro Sur de Chile y Europea. Se muestran los modelos de TRACTs que mejor se ajustan para los datos observados en las poblaciones Chono, Tehuelches, Kawéskar y Yámana. El decaimiento de la ancestría (arriba) compara la distribución de las longitudes de tracts observados y esperados basado en el modelo de migración propuesto. El área sombreada corresponde al 68,7% del intervalo de confianza. Los puntos indican los valores observados. El modelo de migración (parte inferior del gráfico) muestra el cambio en las proporciones de ancestría en el tiempo. De izquierda a derecha se observa la escala en generaciones, los círculos en la parte superior indican la ocurrencia de los eventos de mezcla, siendo el tamaño del círculo la magnitud de los migrantes que llegaron de cada población ancestral. Los colores verde, azul y rojo representan a las ancestrías Nativa, Europea y Aymara respectivamente.

5.8.2. Análisis de Selección

5.8.2.1. Señales de Selección pre-contacto Europeo

5.8.2.1.1. Population Branch Statistic (PBS)

Se llevó a cabo el análisis de PBS para los grupos cazadores recolectores marítimos de la Patagonia Austral, conformado por las poblaciones Kawéskar y Yámana. Estos grupos se contrastaron con los Pehuenches y Huilliches, mientras que como grupo externo se utilizaron a los Ibéricos de 1000G (IBS). Se removieron todos aquellos SNPs que tuvieran datos perdidos. En la Tabla XXX se muestra un resumen con el número de marcadores y ventanas de 200kb utilizadas, además de los que pasaron el filtro por sobre el 1% de los datos. En este punto se eliminaron a cinco individuos, debido a su grado de parentesco registrado durante la toma de muestras en la localidad de Puerto Edén, asociados a individuos Kawéskar.

Tabla XXX. Número de SNPs, ventanas y genes analizadas para los cazadores recolectores marítimos de la Patagonia Austral, junto con los marcadores sobre el 1% de la distribución de los datos para PBS.

	CRM	
SNPs	Nº SNPs	350707
	Nº Genes	21063
	Nº SNPs > 1%	3509
	Nº Genes > 1%	2361
Ventanas	Nº Ventanas 200kb	13180
	Nº Genes	19933
	Nº Ventanas > 20 SNPs	8333
	Prop SNPs > 1% en Ventanas	92
	Nº Genes	207

CRM representa a los cazadores recolectores marítimos de la Patagonia Austral (Kawéskar y Yámana).

La Figura 60 muestra un Manhattan plot para los valores de PBS sobre los grupos cazadores recolectores marítimos en la Patagonia Austral. Sobre esta misma figura, podemos ver que los mayores peaks sobre el genoma corresponden a los cromosomas 6, 8, 11, 12, 15 y 16. De acuerdo al top 10 de los SNPs con mayor valor de PBS en la Tabla XXXI podemos identificar solo cuatro genes (SLCO3A1, ZDHHC14, CHST11 y SYT7), los cuales están asociados con transporte transmembrana de sodio,

procesamiento de rRNA, desarrollo cerebral y desarrollo del músculo cardíaco, entre otros.

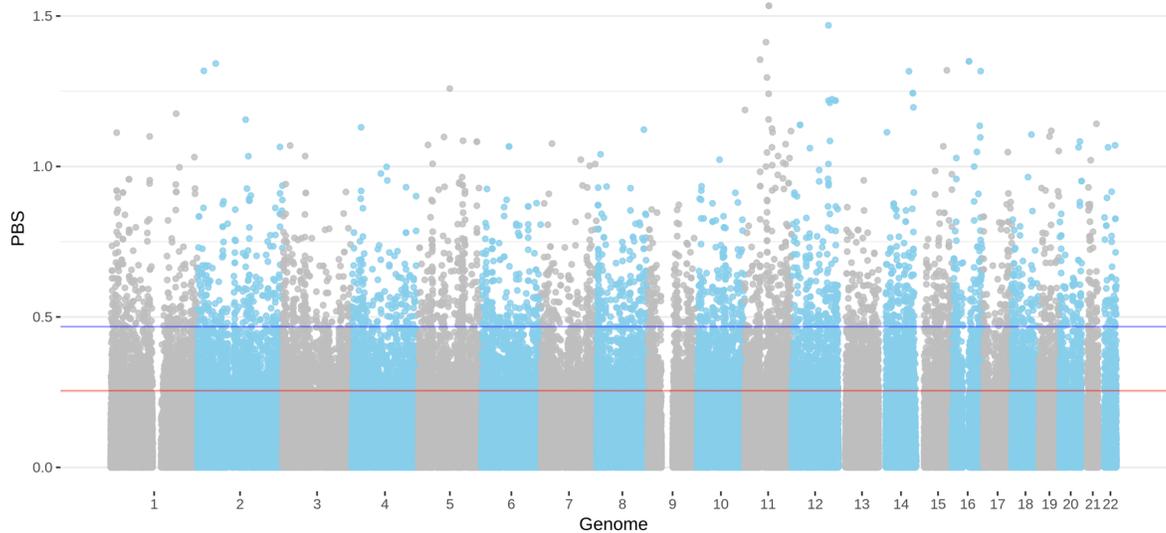


Figura 60. Manhattan Plot para los valores de PBS sobre los cazadores recolectores marítimos de la Patagonia Austral (IBS; Pehuenche/Huilliches; Cazadores Recolectores Marítimos). Las líneas horizontales roja y azul representan el 5 y 1% del percentil respectivamente sobre la distribución de los datos a lo largo del genoma.

Tabla XXXI. Top 10 de los valores de PBS sobre el 1% de los datos observados para los grupos cazadores recolectores marítimos mediante el análisis de SNPs.

Ranking	Cromosoma	Posición	rsID	PBS	Genes
1	15	92487470	rs371469142	1,58840084	SLC03A1
2	15	89564843	rs78969361	1,584315892	Sin Información
3	6	157844283	rs147469339	1,579909329	ZDHHC14
4	8	10591446	rs180736441	1,571213466	LOC102723313
5	11	69593176	rs11263566	1,533969059	Sin Información
6	11	69601039	rs11263574	1,533969059	Sin Información
7	12	105132530	rs10861274	1,468869968	CHST11
8	11	61352303	rs138168158	1,413142882	SYT7, LOC105369331
9	11	44522240	rs79770766	1,354971058	Sin Información
10	16	48853013	rs7203159	1,349315282	Sin Información

Por otro lado, el análisis realizado por ventanas de 200 kb en donde se muestra la proporción de SNPs con un valor de PBS > 1% dentro de cada ventana (Figura 61) muestra otras posibles señales de selección en los cromosomas 2, 5, 11, 12, 14 y 22 donde la proporción de SNPs con PBS > 1% supera el 25%. En la Tabla XXXII se muestra el top 10 de las ventanas con mayor proporción de SNPs > 1%, de estas, cabe recalcar dos genes ubicados en el cromosoma 12; el gen **CHPT1** el cual esta asociado con procesos metabólicos de lípidos. Además, se detectó que aparece la misma ventana bajo selección para el cromosoma 11 (123,6-123,8 Mb) que en el análisis de PBS para la Patagonia con los datos de Axiom. Dentro de esta ventana se puede destacar que resaltan los genes **ZNF202** y **GRAMD1B** relacionados con el metabolismo de lípidos y de la molécula de colesterol respectivamente.

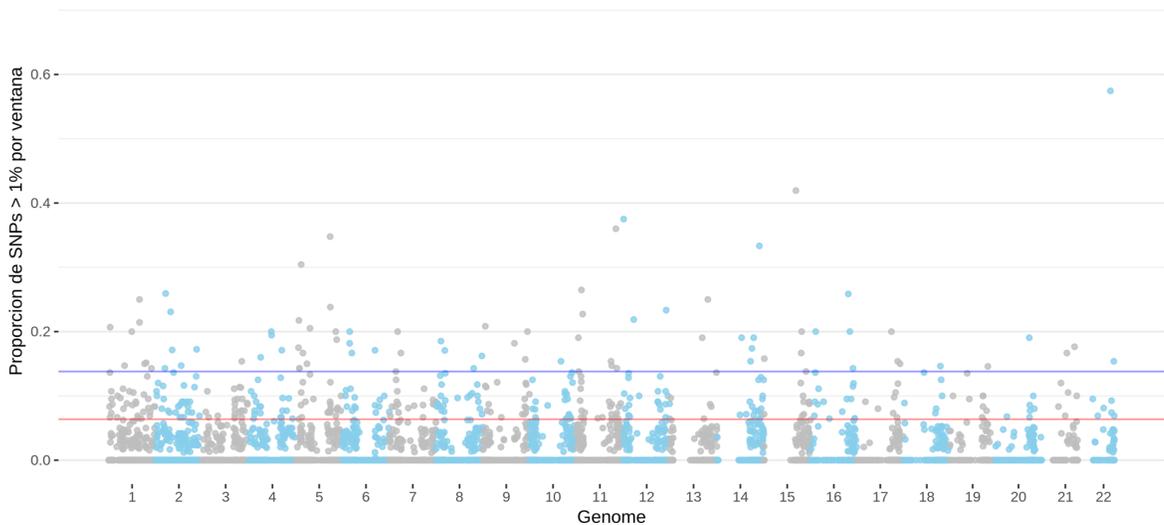


Figura 61. Proporción de SNPs con valores de PBS sobre el 1% en ventanas de 200kb sobre las poblaciones de la Patagonia Austral (IBS; Pehuenche/Huilliches; Cazadores Recolectores Marítimos). Las líneas horizontales roja y azul representan el 5 y 1% del percentil respectivamente sobre la distribución de los datos a lo largo del genoma.

Tabla XXXII. Top 10 de los valores de PBS sobre el 1% de los datos observados para los grupos cazadores recolectores marítimos mediante el análisis de ventanas de 200Kb.

Ranking	Cromosoma	Inicio	Término	Prop SNPs >1%	Genes
1	22	32400000	32600000	0,743589744	FBX07, SYN3, RTCB, BPIFC, LL22NC03-104C7.1, LOC105373003
2	22	32200000	32400000	0,574468085	RTCB, SLC5A4, RFPL3, RFPL2, RFPL3S, SLC5A4-AS1, CPSF1P1, LOC339666, LOC107985561, LOC100422292
3	15	29200000	29400000	0,419354839	NSMCE3, FAM189A1, LOC100420707
4	12	101600000	101800000	0,375	GNPTAB, SYCP3, MYBPC1, CHPT1 , RNY1P16, RNU6-172P, RNU6-101P, ST13P22, RNA5SP368, LOC105369938, LOC105369937
5	11	64400000	64600000	0,36	SLC22A12, SLC22A11, LINC02724, LOC105369341
6	5	39600000	39800000	0,347826087	INTS6P1
7	14	78800000	79000000	0,333333333	NRXN3, LOC107984629
8	5	125400000	125600000	0,304347826	LINC02240
9	11	123600000	123800000	0,264705882	SCN3B, ZNF202 , GRAMD1B , OR6X1, RNU-21P, LOC105369544, LOC105369543
10	2	148000000	148200000	0,259259259	ORC4, MBD5, LOC105373673

5.8.2.1.2. integrated Haplotype Score (iHS)

Se evaluaron las señales de selección basado en desequilibrio de ligamiento mediante iHS por SNPs y por ventanas de 200 kb para las poblaciones cazadoras recolectoras marítimas de la Patagonia Austral. Previo al análisis, se determino a optimizar el numero de individuos a analizar, así como también el número de SNPs a utilizar eliminando todos los datos perdidos en la muestra. Se aplicaron diferentes tasas de pérdida por individuo a través de PLINK (--mind), y luego se evaluó el numero de loci con datos perdidos. Los resultados se encuentran en la Figura 62 y vemos que el número de SNPs sin datos perdidos aumenta a medida que incrementamos el rango de componente Amerindio, pero en detrimento del número de individuos. Analizamos dos rangos; con un filtro del 80% de componente Amerindio encontramos un total de 101761 SNPs para 21 individuos, y con un filtro del 85% de componente Amerindio hay un total de 343114 SNPs, pero el número de

individuos desciende a 15. En la Figura 63 vemos la distribución del número de SNPs por ventanas para evaluar la representatividad de los SNPs seleccionados en el genoma. Ambos rangos no presentan grandes diferencias, por lo que nos quedamos para este análisis con un filtro de 85% y 15 individuos.

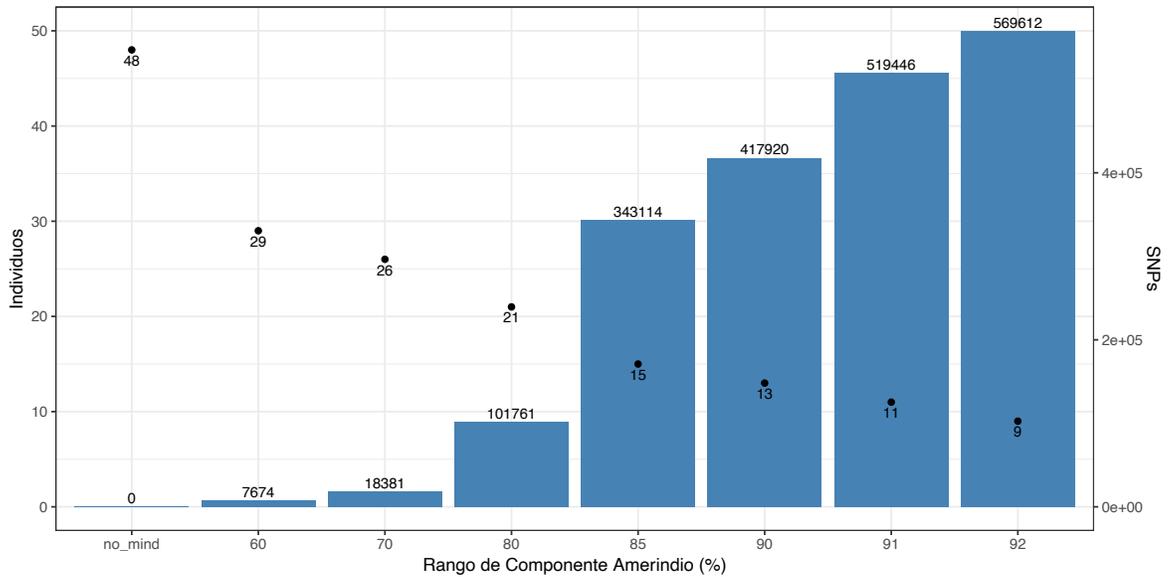


Figura 62. Número de individuos y SNPs para cada rango de componente Amerindio en grupos cazadores recolectores marítimos de la Patagonia. En el eje X se muestra los porcentajes de componente Amerindio utilizados, el S.A representa la ausencia de filtros por componente Amerindio (*sin mind*). El eje Y izquierdo muestra el número de individuos (puntos), mientras que el eje Y derecho revela el número de SNPs (columnas).

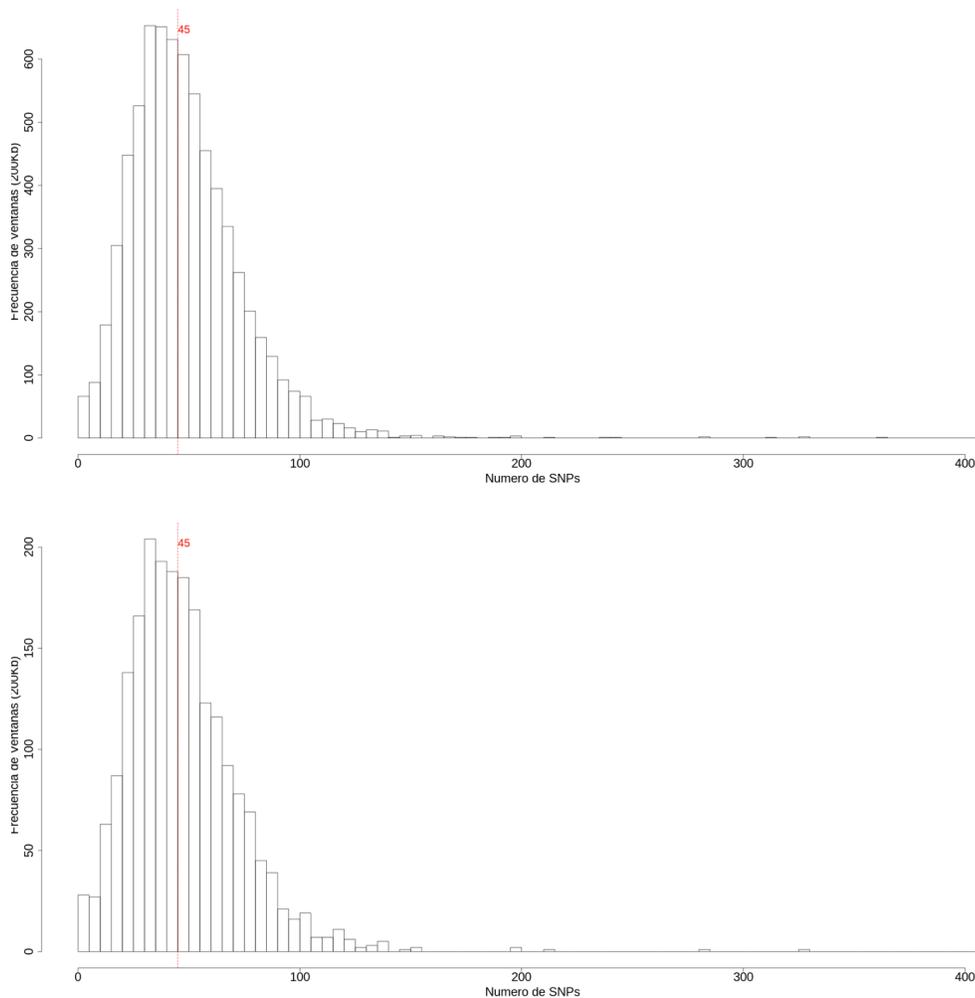


Figura 63. Frecuencia de SNPs por ventana de 200 Kb para diferentes rangos de Ancestría Amerindia. En el eje horizontal vemos el número de SNPs contenidos en cada ventana, mientras que en el eje vertical la frecuencia de ventanas. En A se observa la distribución de los marcadores a lo largo del genoma utilizando un rango de al menos un 77% de ancestría amerindia y en la figura inferior (B) se incluyeron aquellos individuos con al menos un 78% de ancestría amerindia. La línea vertical roja punteada indica la mediana para la distribución de los datos.

Se realizó el análisis de iHS una vez determinado el número de individuos y marcadores a evaluar. El resumen de los filtros utilizados a nivel de SNPs y ventanas de 200 Kb se encuentran descritos en la Tabla XXXIII. En la Figura 64 se muestra la distribución de los valores de iHS a nivel de SNPs a lo largo del genoma para los individuos incorporados en nuestra muestra. De un total de 2436 SNPs con un valor sobre $|iHS| > 2.5$, se observa que los peaks más elevados corresponden a los

cromosomas 2, 8, 12 y 16. El top 10 con los valores de $|iHS| > 2.5$ se encuentran en la Tabla XXXIV, donde cabe mencionar al gen **SLC25A12** en el cromosoma 12 el cual esta asociado con la gluconeogénesis y con la regulación positiva para la biosíntesis de ATP. Sin embargo, hay otro gen que cobra mayor importancia: **TRPM8**, el cual esta ubicado en el cromosoma 2 y esta asociado a estímulos frente a temperatura, regulación positiva de la termogénesis inducida por el frío, entre otras, todas ligadas con mecanismos de respuesta a frío.

Tabla XXXIII. Número de SNPs, Ventanas y Genes analizados para iHS en poblaciones cazadores recolectores marítimas de la Patagonia.

	SNPs				Ventanas				
	Nº SNPs	Nº Genes	Nº SNPs > 2.5	Nº Genes > 2.5	Nº Ventanas 200Kb	Nº Genes	Nº Ventanas > 20 SNPs	Prop SNPs (>2.5) > 1% en Ventanas	Nº Genes
CRM	143075	6147	2436	793	6382	5924	2190	36	57

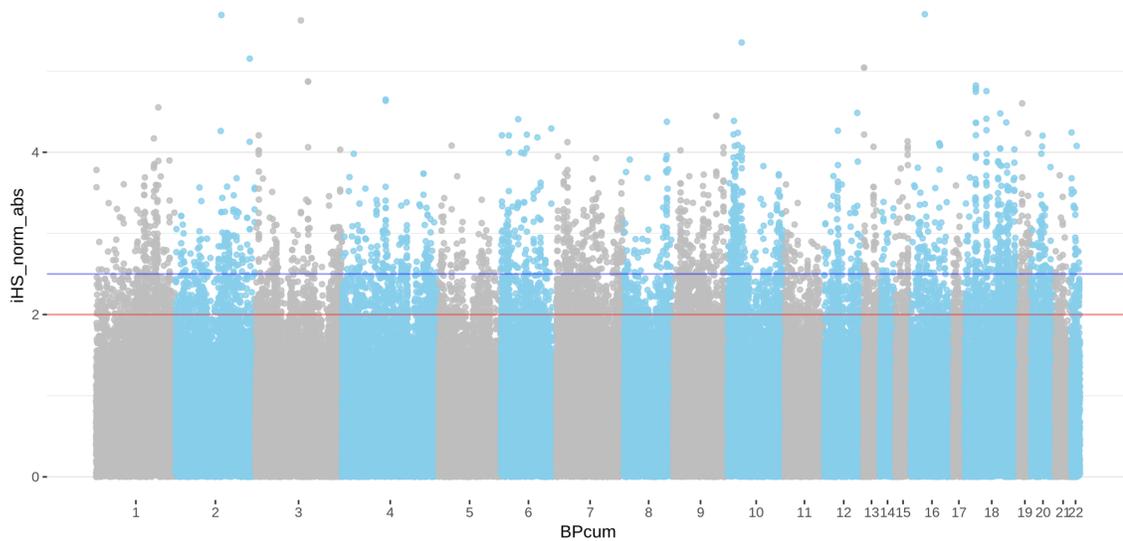


Figura 64. Valores de iHS por SNPs en poblaciones cazadores recolectores marítimas de la Patagonia. Los valores fueron normalizados y se muestra para todo el genoma. Las líneas horizontales roja y azul representan un puntaje de corte de $|iHS| > 2$ e $|iHS| > 2.5$ respectivamente.

Tabla XXXIV. Top 10 de los valores de |iHS| sobre el 1% de los datos observados para los grupos cazadores recolectores marítimos de la Patagonia mediante el análisis por SNPs.

Ranking	Cromosoma	Posición	rsID	iHS	Genes
1	16	17415682	rs8057681	5,7011	XYLT1
2	2	172728492	rs35208355	5,69129	SLC25A12
3	3	120806790	rs9848328	5,6253	STXBP5L
4	10	11785657	rs35706434	5,35276	ECHDC3
5	2	234914079	rs17869069	5,15394	TRPM8
6	13	44424210	rs1466005	5,04326	CCDC122
7	13	44431660	rs9525855	5,04326	CCDC122
8	3	133449190	rs9858383	4,87145	TF
9	3	133450720	rs6787177	4,87145	TF
10	18	9254785	rs17498752	4,82383	ANKRD12

En la Figura 65; **Error! No se encuentra el origen de la referencia.** vemos la proporción de SNPs con valores de $|iHS| > 2.5$ dentro de cada ventana de 200Kb. El resultado indica que la mayoría de las ventanas con una mayor proporción están distribuidas entre el cromosoma 10 y 18. Los resultados de las ventanas dentro del top 10 con las proporciones mas altas, se encuentran en la Tabla XXXV, donde se observa efectivamente que las mayores puntuaciones se distribuyen entre estos cromosomas. La mayoría de las ventanas caen en regiones de loci intergénicos y o regiones sin caracterizar. No obstante, aparece el gen UCN3 dentro de una ventana bajo selección para el cromosoma 10, el cual esta relacionado con la respuesta frente a hipoxia, niveles de nutrientes, respuesta de glucosa y a periodos de hambruna.

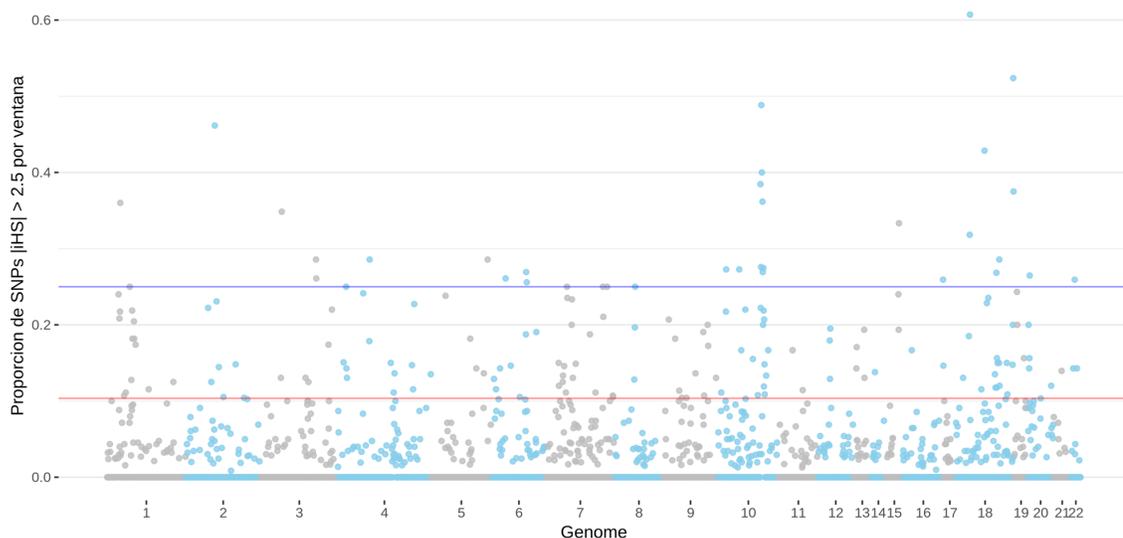


Figura 65. Proporción de SNPs con valores de $|iHS| > 2.5$ dentro de cada ventana de 200Kb en poblaciones cazadoras recolectoras de la Patagonia. Los valores fueron normalizados y se muestra para todo el genoma. Las líneas horizontales roja y azul representan el top 5 y 1% de los datos respectivamente.

Tabla XXXV. Top 10 de los valores de $|iHS|$ sobre el 1% de los datos observados para los grupos cazadores recolectores marítimos de la Patagonia mediante el análisis por ventanas de 200Kb.

Ranking	Cromosoma	Inicio	Termino	Prop SNPs > $ iHS > 2.5$	Genes
1	18	27600000	27800000	0,607142857	LOC105372042
2	18	9200000	9400000	0,523809524	TWSG1, ANKRD12, LOC107985188, LOC102724397
3	10	5200000	5400000	0,488372093	UCN3, AKR1C4, TUBAL3, LINC02561, RPL26P28, ARL4AP3, AKR1C7P, LOC111946241
4	2	177000000	177200000	0,461538462	RNA5SP112, KRT8P40, LOC112806061, LOC105373760, LOC391465
5	18	44800000	45000000	0,428571429	SETBP1, MIR4319, LOC111718495, LOC114827855 GDI2, ASB13, TASOR2, LINC02677, RN7SL445P,
6	10	5600000	5800000	0,4	LINC02678, NRBF2P5, LOC107984199 MANCR, LINC00705, LOC111818965, LOC107984197, LOC105376375, LOC105376374,
7	10	4600000	4800000	0,384615385	LOC105376373 RALBP1, TWSG1, PPP4R1, IRX4P1, RNU2-27P,
8	18	9400000	9600000	0,375	LOC105371980 IL2RA, PFKFB3, RBM17, MIR3155A, MIR3155B, RPL32P23, RN7SKP78, LOC107984201,
9	10	6000000	6200000	0,361702128	LOC101928080
10	1	166600000	166800000	0,36	FMOP9P, FMO11P, FMO10P, RPL4P2, CNN2P10

5.8.2.2. Señales de Selección post-contacto Europeo

5.8.2.2.1. Detección de Loci bajo Selección mediante Ancestría

Se realizó este análisis con la finalidad de visualizar el cambio de las frecuencias de ancestrías Amerindia, Europea y Africana lo largo del genoma en los grupos cazadores recolectores marítimos de la Patagonia, incluyendo únicamente a las localidades de Punta Arenas y Puerto Edén, ambas pertenecientes a grupos Kawéskar. En la Tabla XXXVI se pueden ver los promedios junto con sus respectivas desviaciones y errores estándar sobre un total de 869.562 SNPs. A partir de esto, podemos ver que el promedio mas alto corresponde a la ancestría Amerindia con un 86.8%, mientras que la ancestría Europea posee un promedio de 13%. Al igual que lo visto en este estudio, el componente Africano se encuentra por muy debajo en comparación con las dos primeras ancestrías con un 0.3%.

Tabla XXXVI. Promedio, desviación y error estándar para las ancestrías Amerindias, Europeas y Africanas para las poblaciones Kawéskar y Yámana en la Patagonia Austral.

		CRM (KAW,YAM)
Ancestría Amerindia	Media	0.8676018
	Desviación Estándar	0.04348014
	Error Estándar	0.01087004
Ancestría Europea	Media	0.129665
	Desviación Estándar	0.04307604
	Error Estándar	0.01076901
Ancestría Africana	Media	0.002733189
	Desviación Estándar	0.007728416
	Error Estándar	0.001932104

CRM corresponde a los cazadores recolectores marítimos de la Patagonia Austral, compuesto por los Kawéskar y Yámana.

En la Figura 66 podemos apreciar la distribución de las ancestrías Amerindia y Europea a lo largo del genoma para los cazadores recolectores marítimos de la Patagonia. Para el caso de los Amerindios (Figura superior), vemos que un total de 502 genes se encuentran por sobre el 1% de la distribución de los datos (Tabla XXXVII), donde los peaks mas altos corresponden a los cromosomas 2, 3, 4, 11 y 22.

Para la ancestría Europea (Figura inferior) se registraron un total de 396 genes sobre el top 1% de la distribución de los datos (Tabla XXXVII). Los peaks mas altos se encuentran en los cromosomas 6, 8, 14 y 15.

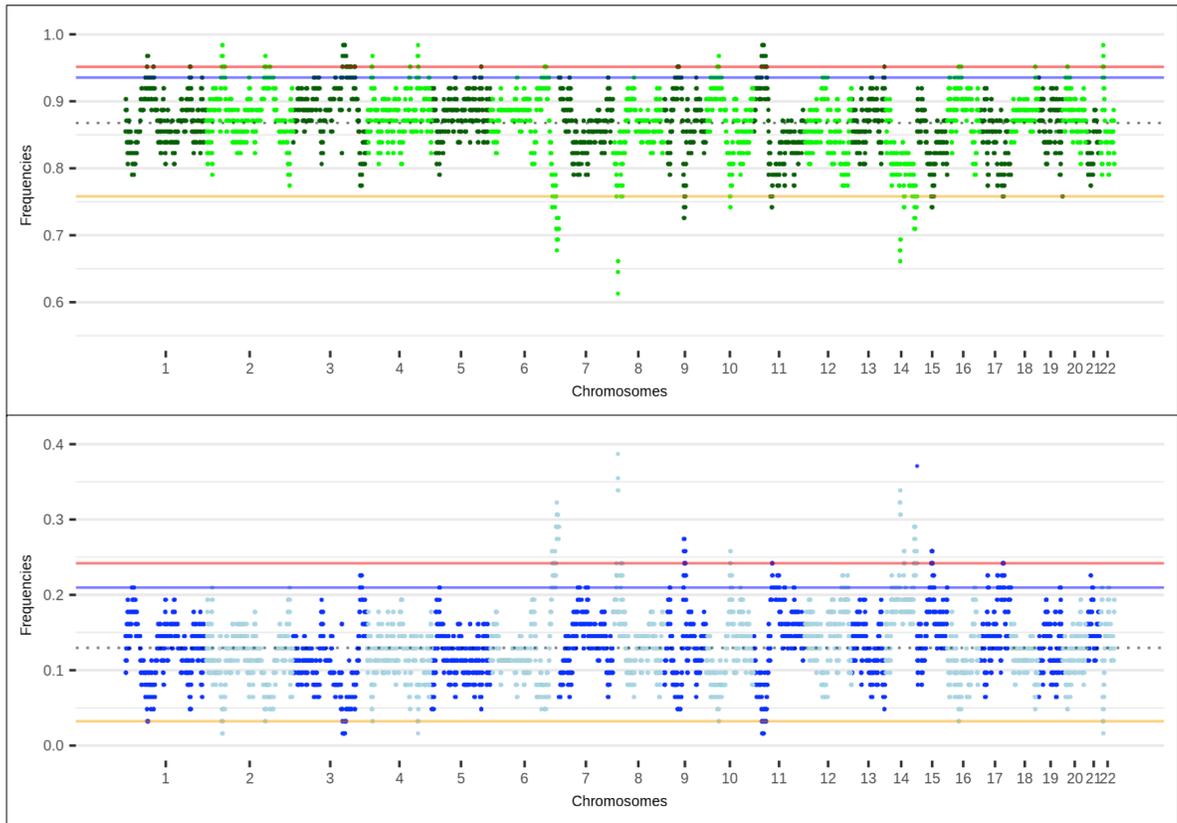


Figura 66. Detección de loci bajo selección a lo largo del Genoma basado en ancestrías para las poblaciones Kawéskar y Yámana de la Patagonia Austral. Las líneas horizontales azul y roja representan el top 5 y 1% de los datos respectivamente. La línea horizontal anaranjada representa el top 1% inferior de la distribución de los datos. La figura superior muestra la distribución de la ancestría Amerindia a lo largo del Genoma, mientras que la figura inferior revela la distribución de la ancestría Europea a lo largo del Genoma.

Tabla XXXVII. Resumen con el número de SNPs y genes para el análisis de Detección de Loci bajo Selección para las ancestrías Amerindias y Europeas.

	Nº SNPs	Nº Genes	Nº SNPs > top 1%	Nº Genes > top 1%
Ancestría Amerindia	869562	24507	22772	502
Ancestría Europea	869562	24507	12579	396

5.8.3. Enriquecimiento de procesos biológicos

5.8.3.1. Population Branch Statistic (PBS)

Se realizó el análisis de enriquecimiento para los grupos de la Patagonia Austral para un total de 2361 genes que sobre el 1% de la distribución de los datos. En la Tabla XXXVIII se describen 56 procesos biológicos generales significativos (P-value < 0.05), de estos destacan varios relacionados con canales para el transporte de iones de Potasio y Calcio, así como también procesos asociados con el desarrollo del sistema nervioso.

Tabla XXXVIII. Enriquecimiento de procesos biológicos para los grupos cazadores recolectores de la Patagonia Austral a partir de los genes por sobre el 1% identificados a partir de PBS.

GO	GO biological process complete	O	E	ER	P-value	FDR
GO:0043268	positive regulation of potassium ion transport	13	3.94	3.30	6.82E-04	4.48E-02
GO:1901379	regulation of potassium ion transmembrane transport	20	7.45	2.69	3.98E-04	3.06E-02
GO:0072073	kidney epithelium development	28	12.45	2.25	3.17E-04	2.58E-02
GO:0070588	calcium ion transmembrane transport	41	19.58	2.09	5.56E-05	6.62E-03
GO:0006813	potassium ion transport	31	15.32	2.02	7.30E-04	4.75E-02
GO:0003007	heart morphogenesis	47	23.94	1.96	6.91E-05	7.69E-03
GO:0010817	regulation of hormone levels	77	48.41	1.59	2.44E-04	2.09E-02
GO:0032101	regulation of response to external stimulus	110	72.14	1.52	5.60E-05	6.62E-03
GO:0050801	ion homeostasis	104	71.61	1.45	4.79E-04	3.51E-02

GO= código de proceso biológico. O= número de genes observados en cada categoría GO. E= número de genes esperado para cada categoría GO. ER= Enriquecimiento de cada proceso biológico. P-value= significancia para cada proceso (<0.05). FDR= valores significativos de p-valor con FDR <0.05 de Benjamin & Hochberg (1995).

Entre las diferentes rutas metabólicas que se han encontrado para este análisis, destaca nuevamente la secreción y síntesis de cortisol (p-value < 0.05), esta categoría ya se había descrito para el set de datos de Axiom y podría representar un escenario importante para las muestras de la Patagonia. También se encontraron otras rutas metabólicas como la secreción de renina, biosíntesis de O-glicanos, arraste circadiano y sinapsis glutamatérgica entre otros (Figura 67).

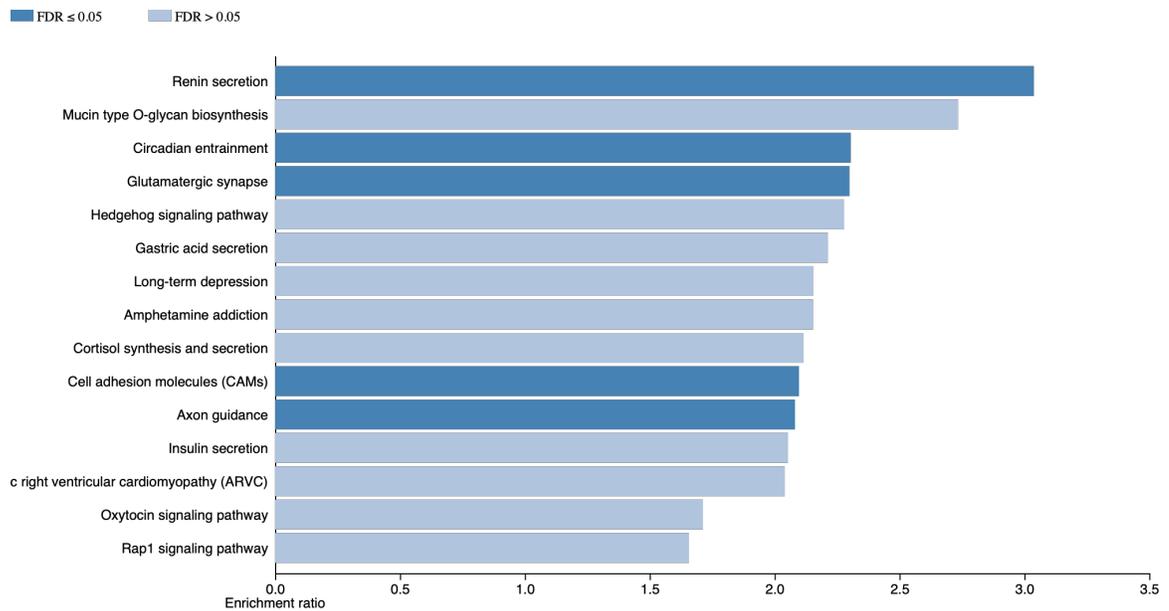


Figura 67. Esquema con la representación de rutas biológicas enriquecidas para los análisis de PBS sobre los SNPs en los individuos de la Patagonia Austral. Cada barra horizontal representa una ruta biológica y el ancho corresponde a la proporción de enriquecimiento. El NES (Normalized Enrichment Score) indica la distribución de las categorías a lo largo de una lista de genes rankeados (valor de PBS). Valores extremos de NES indican un set de genes con un alto ranking y pertenecientes a una misma categoría, representando una regulación positiva (en azul). Los colores de tonalidades fuertes representan categorías con corrección a posterior (FDR < 0.05).

El análisis realizado para la proporción de SNPs por sobre el 1% de la distribución de los datos en ventanas de 200Kb arrojó un total de 207 genes. Estos genes están relacionados para más de 400 procesos biológicos significativos (P-value < 0.05). En la Tabla XXXIX podemos ver algunos de los procesos biológicos de posible interés para nuestro estudio, como los relacionados con el transporte de ácidos grasos de cadena corta, procesos metabólicos para el ácido retinoico, respuesta a la molécula del cortisol y respuesta conductual a periodos de hambruna. Para la identificación de rutas metabólicas (p-value 0.05) destacan varios procesos ligados a la ingesta de alimentos, como la degradación de ácidos grasos, digestión y absorción de carbohidratos y de vitaminas, metabolismo del almidón, entre otros.

Tabla XXXIX. Enriquecimiento de procesos biológicos en ventanas de 200 kb para los grupos cazadores recolectores marítimos de la Patagonia Austral, a partir de los genes por sobre el 1% identificados a partir de PBS.

GO	GO biological process complete	O	E	ER	P-value	FDR*	Genes
GO:0015913	short-chain fatty acid import	1	.02	51.57	2.85E-02	2.09E00	SLC5A8
GO:0015912	short-chain fatty acid transport	1	.02	51.57	2.85E-02	2.08E00	SLC5A8
GO:0034653	retinoic acid catabolic process	1	.04	25.78	4.71E-02	1.87E00	CYP26B1
GO:0042594	behavioral response to starvation	1	.02	62.23	2.37E-02	1.00E00	GALP
GO:0051414	response to cortisol	1	.04	24.89	4.69E-02	1.00E00	SLIT2

FDR*=valores no significativos

GO= código de proceso biológico. O= número de genes observados en cada categoría GO. E= número de genes esperado para cada categoría GO. ER= Enriquecimiento de cada proceso biológico. P-value= significancia para cada proceso (<0.05). FDR= valores significativos de p-valor con FDR <0.05 de Benjamin & Hochberg (1995).

5.8.3.2. *integrated Haplotype Score (iHS)*

El análisis de enriquecimiento para iHS se realizó para un total de 793 genes que contienen en su información genética a SNPs con valores de $|iHS| > 2.5$. Al igual que en los análisis anteriores, los resultados muestran un total de 148 procesos biológicos enriquecidos (P-value < 0.05). Entre estos, quizás los mas interesantes corresponden a desarrollo del sistema neuronal, desarrollo del timo, respuesta inmune, entre otros (Tabla XL).

Tabla XL. Enriquecimiento de procesos biológicos para los grupos cazadores recolectores de la Patagonia Austral a partir de los genes con valores de $|iHS| > 2.5$ por SNPs.

GO	GO biological process complete	O	E	ER	P-value	FDR*	Genes
GO:0000272	polysaccharide catabolic process	4	.93	4.31	2.90E-02	3.36E00	AOAH; PYGB; RB1CC1; CHIT1
GO:0048538	thymus development	5	1.74	2.87	4.95E-02	3.83E00	MAPK1; PRKDC; FGF10; MAFB; PBX1
GO:0045087	innate immune response	9	19.39	.46	1.66E-02	3.51E00	TICAM1; KRT16; PRKDC; LYN; DHX58; KIF16B; MBL2; RSAD2; CD1D

FDR*=valores no significativos

GO= código de proceso biológico. O= número de genes observados en cada categoría GO. E= número de genes esperado para cada categoría GO. ER= Enriquecimiento de cada proceso biológico. P-value= significancia para cada proceso (<0.05). FDR= valores significativos de p-valor con FDR <0.05 de Benjamin & Hochberg (1995).

La identificación de rutas metabólicas arrojó nuevamente la síntesis y secreción del cortisol para el análisis de iHS, esta ruta se ha encontrado anteriormente con los análisis de PBS. Destaca también la ruta metabólica asociada al péptido apelina, el cual fue detectado anteriormente para los valores de PBS (p-value < 0.05) y que esta involucrado en diferentes procesos fisiológicos como la angiogénesis, funciones cardiovasculares, proliferación celular y regulación en el metabolismo de energía. Además, este péptido se encuentra relacionado con algunas patologías como diabetes, obesidad, enfermedades cardiovasculares, etc.

Para el análisis realizado con la proporción de SNPs con valor de $|iHS| > 2.5$ en ventanas de 200Kb se pudieron identificar 57 genes, los cuales arrojaron un total de 165 procesos biológicos (p-value < 0.05). Los más importantes corresponden a algunos procesos relacionados con el metabolismo de la carnitina y de ácidos grasos, así como también la reducción de la ingesta de comida debido a un exceso de dieta (Tabla XLI).

Tabla XLI. Enriquecimiento de procesos biológicos en ventanas de 200kb para los grupos cazadores recolectores de la Patagonia, a partir de los genes por sobre el 1% identificados a partir de iHS.

GO	GO biological process complete	O	E	ER	P-value	FDR*	Genes
GO:0002023	1. reduction of food intake in response to dietary excess	1	.01	> 100	1.31E-02	1.00E00	GDF15
GO:0050957	2. equilibrioception	1	.02	56.55	2.18E-02	1.00E00	PCDH15
GO:0051791	3. medium-chain fatty acid metabolic process	1	.03	32.31	3.47E-02	1.00E00	CROT
GO:0032376	4. positive regulation of cholesterol transport	1	.03	32.31	3.47E-02	1.00E00	ABCB4
GO:0009437	5. carnitine metabolic process	1	.03	32.31	3.47E-02	1.00E00	CROT
GO:0033540	6. fatty acid beta-oxidation using acyl-CoA oxidase	1	.04	28.27	3.89E-02	1.00E00	CROT
GO:0055088	7. lipid homeostasis	2	.25	8.08	2.67E-02	1.00E00	NR5A2, ABCB4

FDR*=valores no significativos

GO= código de proceso biológico. O= número de genes observados en cada categoría GO. E= número de genes esperado para cada categoría GO. ER= Enriquecimiento de cada proceso biológico. P-value= significancia para cada proceso (<0.05). FDR= valores significativos de p-valor con FDR <0.05 de Benjamin & Hochberg (1995).

5.8.3.3. Detección de Loci bajo Selección mediante Ancestría

Este análisis permitió clasificar un total de 502 genes por sobre 1% de los datos, de los cuales se encontraron cuatro procesos biológicos sobrerrepresentados significativamente (p-value < 0.05) (Tabla XLII). Estos procesos se encuentran

estrechamente ligados, asociados principalmente con la biología molecular de nucleótidos. En lo que respecta a rutas metabólicas (p -value < 0.05) cabe mencionar resultados bastante similares con lo realizado para Axiom al estar asociadas con diversas categorías asociadas al metabolismo de piruvato, secreción de insulina, diabetes mellitus tipo I, etc.

Tabla XLII. Enriquecimiento de procesos biológicos para los grupos cazadores recolectores marítimos de la Patagonia a partir de los genes por sobre el 1% identificados en aquellos Loci bajo selección mediante Ancestría Amerindia.

GO	GO biological process complete	O	E	ER	P-value	FDR
GO:0035589	1. G protein-coupled purinergic nucleotide receptor signaling pathway	6	.22	27.26	5.64E-07	4.42E-03
GO:0035588	2. G protein-coupled purinergic receptor signaling pathway	7	.34	20.58	2.74E-07	4.30E-03
GO:0035590	3. purinergic nucleotide receptor signaling pathway	6	.40	14.99	9.03E-06	3.54E-02
GO:0035587	4. purinergic receptor signaling pathway	7	.52	13.45	2.90E-06	1.52E-02

GO= código de proceso biológico. O= número de genes observados en cada categoría GO. E= número de genes esperado para cada categoría GO. ER= Enriquecimiento de cada proceso biológico. P-value= significancia para cada proceso (<0.05). FDR= valores significativos de p-valor con FDR <0.05 de Benjamin & Hochberg (1995).

Para el caso de la ancestría Europea, se identificaron 395 genes sobre el 1%, el cual nos arroja un total de 285 procesos biológicos (p -value < 0.05). Algunos de los procesos biológicos importantes para nuestro estudio que aparecieron para la ancestría Europea fueron respuesta a la luz UV-C, regulación positiva frente a procesos glicolíticos y procesos metabólicos a ácidos grasos de cadena larga. Mas aún, durante la identificación de rutas metabólicas (p -value 0.05) pudimos observar la biosíntesis de ácidos grasos insaturados y la elongación de ácidos grasos aparecen como las categorías mas enriquecidas (**Figura 68**).

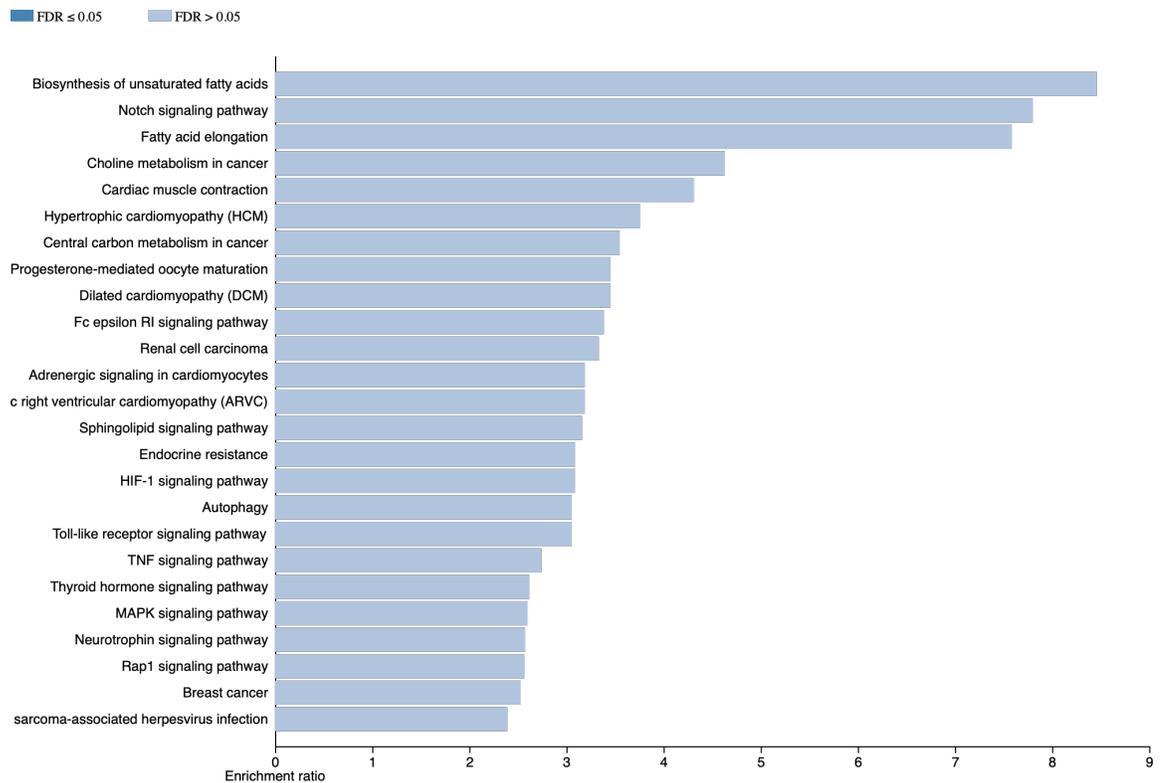


Figura 68. Esquema con la representación de rutas biológicas enriquecidas para la detección de Loci bajo selección mediante ancestría en los individuos de la Patagonia Austral. Cada barra horizontal representa una ruta biológica y el ancho corresponde a la proporción de enriquecimiento. El NES (Normalized Enrichment Score) indica la distribución de las categorías a lo largo de una lista de genes rankeados (valor de PBS). Valores extremos de NES indican un set de genes con un alto ranking y pertenecientes a una misma categoría, representando una regulación positiva (en azul). Los colores de tonalidades fuertes representan categorías con corrección a posteriori (FDR < 0.05).

5.8.4. Genes Candidatos a Selección Positiva en la Patagonia Austral

A continuación, se muestra el set de genes que se encuentran dentro del top 10 para las señales más altas que arrojaron los análisis de PBS e iHS, tanto para SNPs como ventanas de 200Kb (Tabla XLIII). Estos genes corresponden a aquellos que tienen una directa relevancia en torno a sus mecanismos y procesos biológicos frente a respuesta de frío y dieta. Como podemos ver, los genes se concentran en los cromosomas 2, 11 y 12. De estos, la mayoría está asociado a mecanismo de frío (**SLC25A12**, **TRPM8**), mientras que el resto está relacionado con metabolismo de lípidos, incluyendo el colesterol (**CHPT1**, **ZNF202** y **GRAMD1B**). Dentro de este set de genes destacan nuevamente **ZNF202** y **GRAMD1B**, los cuales se ubican en una

región bajo selección para el cromosoma 11 (123,6-123,8Mb). Esta misma ventana se encuentra entre las señales más potentes de selección por parte de los análisis realizados mediante PBS con los datos de Axiom, y tanto para los cazadores recolectores como cazadores recolectores marítimos de la Patagonia (Figura 69). Por otro lado, destaca el gen UCN3, el cual se encuentra asociado con hipoxia celular, respuesta a los niveles de glucosa, entre otros.

Tabla XLIII. Descripción de los genes candidatos dentro del top 10 para frío y dieta en la Patagonia con los análisis realizados con PBS e iHS (SNPs y ventanas de 200Kb).

Crom	rsID	Inicio	Termino	Gen	Marcador		Grupo	Análisis
					SNPs	Ventana (200Kb)		
12		101600000	101800000	CHPT1		X	CRM	PBS
11		123600000	123800000	ZNF202, GRAMD1B		X	CRM	PBS
2	rs17869069			TRPM8	X		CRM	iHS
12	rs35208355			SLC25A12	X		CRM	iHS
10		5200000	5400000	UCN3		X	CRM	iHS

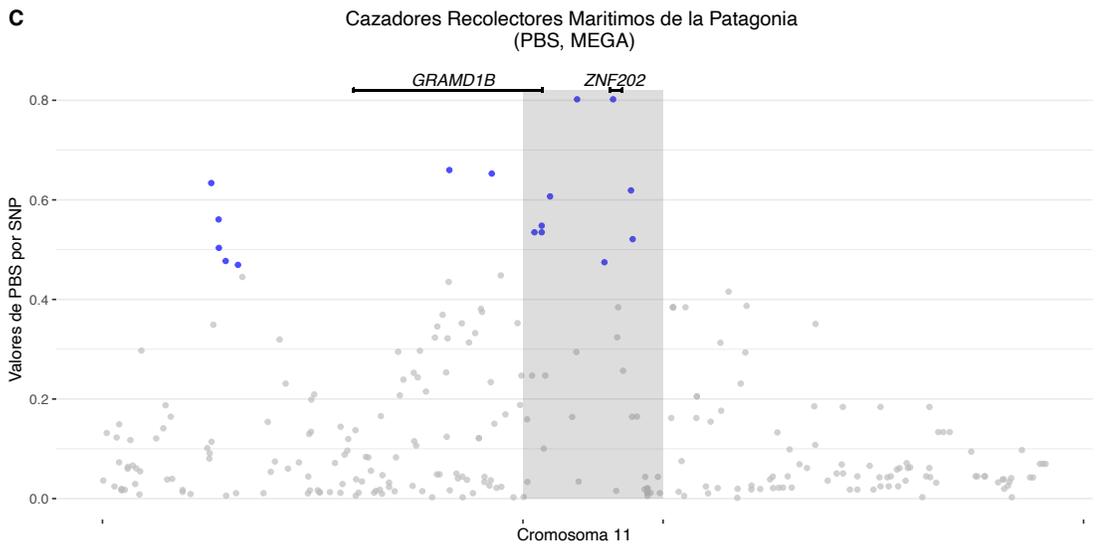
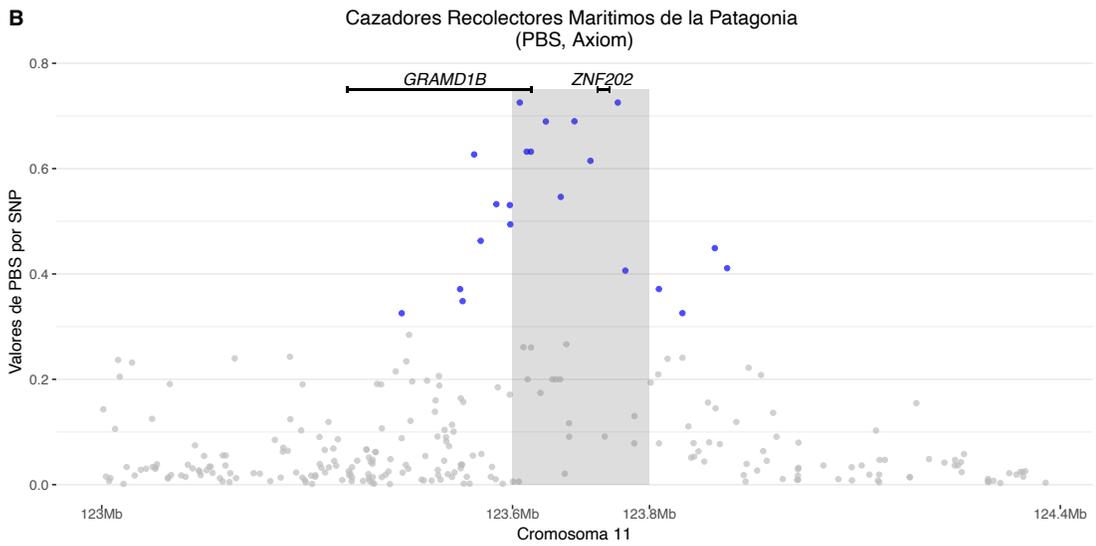
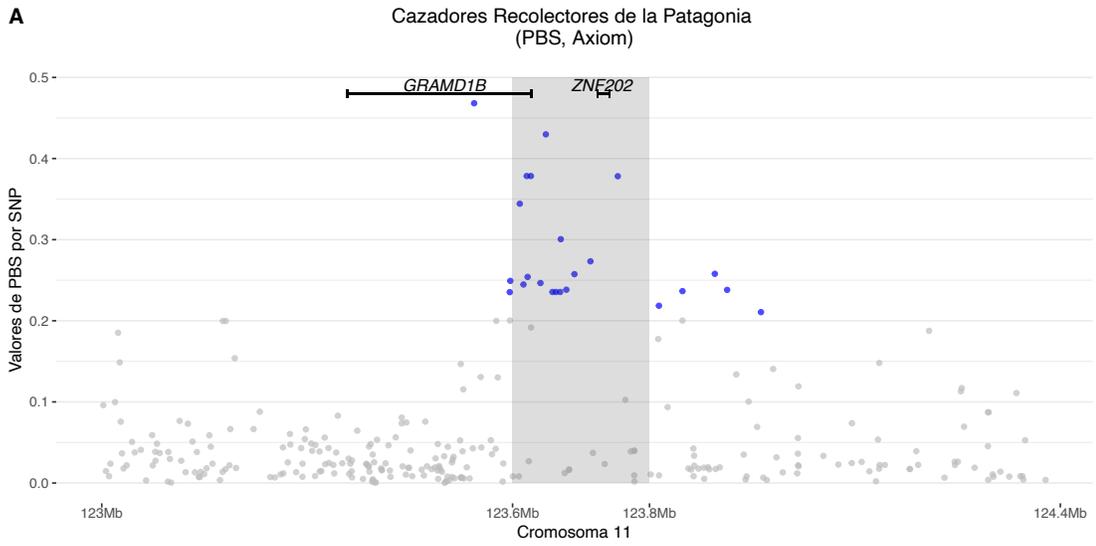


Figura 69. Señal de Adaptación para la región del Cromosoma 11 (123,6-123,8Mb) para las comparaciones de Axiom y MEGA. Los análisis corresponden a PBS donde los puntos corresponden a cada SNP. Los marcadores de color azul representan la distribución del 1% de los datos. El área gris corresponde a la ventana bajo selección. Las líneas negras representan el tamaño de los genes GRAMD1B y ZNF202.

5.8.4.1. Comparación de Regiones bajo Selección con Genes asociados a Frío y Dieta

Se evaluaron para los grupos de la Patagonia Austral las distancias de aquellas ventanas con valores de PBS sobre el 1% con los genes descritos previamente para frío, dieta. Esto con el propósito de determinar la existencia de regiones cromosómicas cercanas a genes bajo selección en diversas poblaciones, principalmente de Siberia (Figura 70). Acá podemos ver principalmente la presencia de una región bajo selección en nuestros datos y que se encuentra en el cromosoma 11, próximo a los genes de la familia **FADS**, **LRP5** y **CPT1A**.

Posterior a esta comparación se procedió a evaluar con mayor precisión esta zona del cromosoma 11 (61-70Mb), pero esta vez con los valores de PBS por SNPs. Tal como se puede apreciar en la Figura 71, existe una region con SNPs los cuales se encuentran por sobre el 1% de los valores de PBS y que se encuentran muy cercanos a genes descritos previamente. Este es el caso para el gen **CPT1A** (68.7-68.8Mb) el cual esta relacionado con mecanismos de respuesta a frío y dieta marina en poblaciones de Siberia (Cardona et al., 2014; Clemente et al., 2014) y que posee una ventana de 200Kb a 0.2Mb de distancia, sin mencionar la presencia de dos SNPs que se encuentran dentro de este gen. También es posible detectar tres SNPs cercanos al gen **FADS1**, asociado a metabolismo de lípidos (Fumagalli et al., 2015).

Resulta importante destacar el patrón observado en este cromosoma, el cual se repite con los análisis anteriores realizados tanto para cazadores recolectores como cazadores recolectores marítimos de la Patagonia. Este patrón se observa de igual manera para las poblaciones de la Patagonia Austral, con una región cromosómica que va desde los 63.3Mb a los 65.2Mb. Esta region es similar a la encontrada anteriormente y se pueden ver los mismos genes, de los cuales algunos

se encuentran asociados al metabolismo de glucógeno, y mas importante aún al metabolismo de lípidos. En etse último proceso destacan los genes **PLAAT3**, **PLAAT4** y **PLAAT5**.

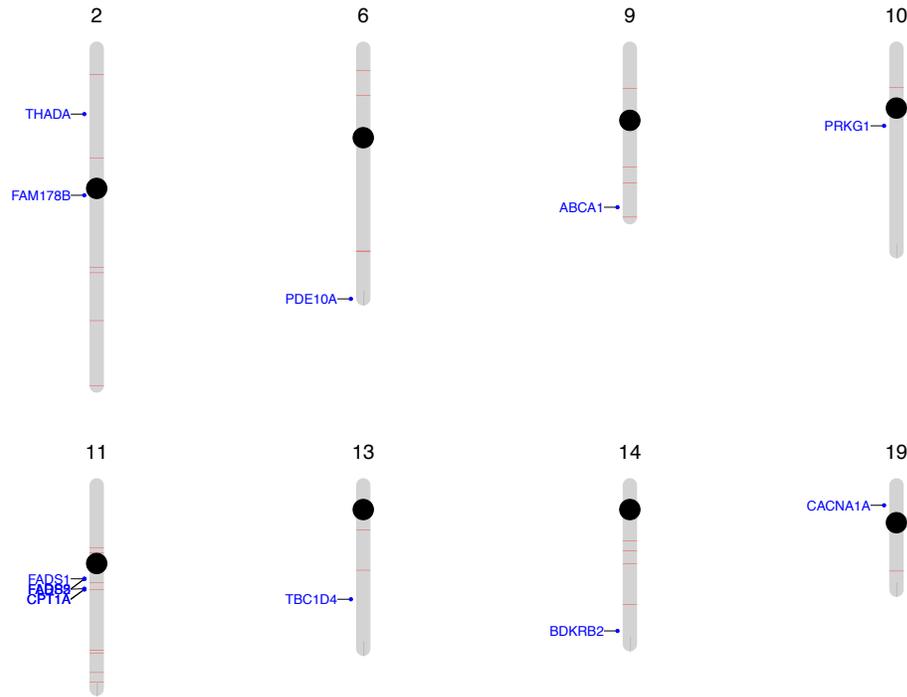


Figura 70. Comparación entre los genes asociados previamente a frío y dieta con las ventanas sobre el 1% de los valores de PBS. En azul se encuentran los genes asociados a frío y dieta descritos previamente. Las bandas horizontales rojas dentro de cada cromosoma corresponden a ventanas de 200Kb sobre el 1% de los datos.

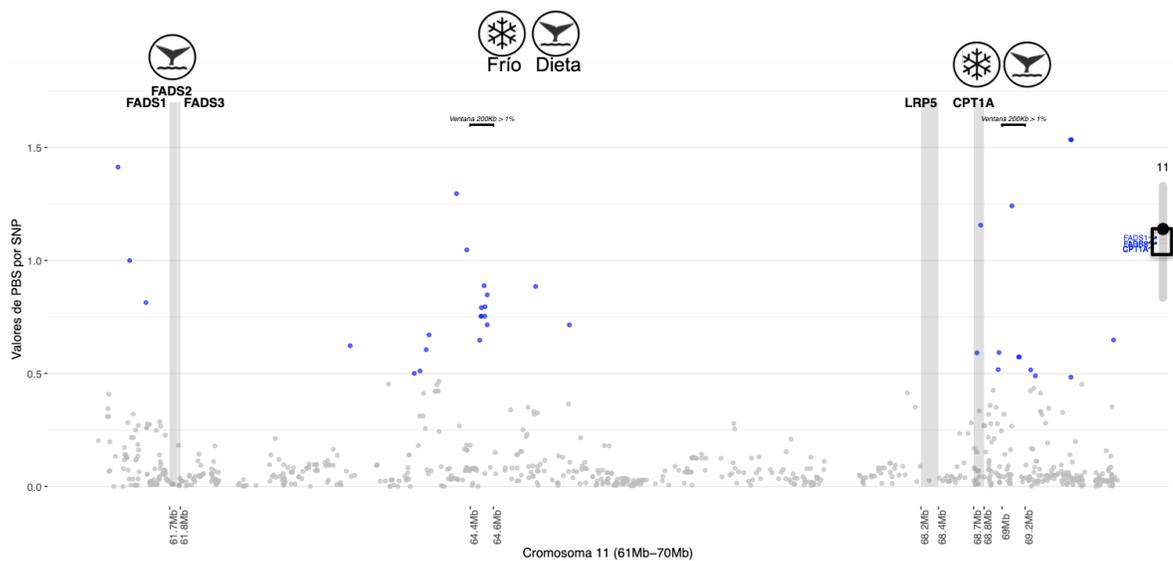


Figura 71. Distribución de los valores de PNS por SNPs entre las regiones 61-70Mb en el cromosoma 11. La distribución de los valores por sobre el 1% de los datos se encuentra en color azul. Los recuadros en gris señalan los genes FADS1, FADS2, FADS3, LRP5 y CPT1A. Las líneas horizontales en el margen superior muestra las ventanas de 200Kb para aparecen cercanas a los genes de interés. Al costado derecho se muestra el cromosoma 11 con el recuadro en negro donde se amplificó el análisis. Los símbolos que se encuentran al lado de cada gen señala el fenotipo asociado.

A continuación, se procedió a evaluar las señales de PBS por SNPs para los cromosomas donde se encuentran todos los otros genes descritos previamente para frío, dieta marina o buceo. Esto se realizó para las poblaciones de la Patagonia Austral y se puede ver en la **Figura 72**, donde existen algunos SNPs que se encuentran cerca de genes relacionados con mecanismos de buceo como **PDE10A**, **BDKRB2** y **CACNA1A**, todas estas descritas por M. A. Ilardo et al., (2018). Por otro lado, para el gen **PRKG1** asociado a frío en poblaciones de Siberia se puede ver una región de varios SNPs con valores sobre el 1% que se encuentra a 2 Mb de distancia de este gen.

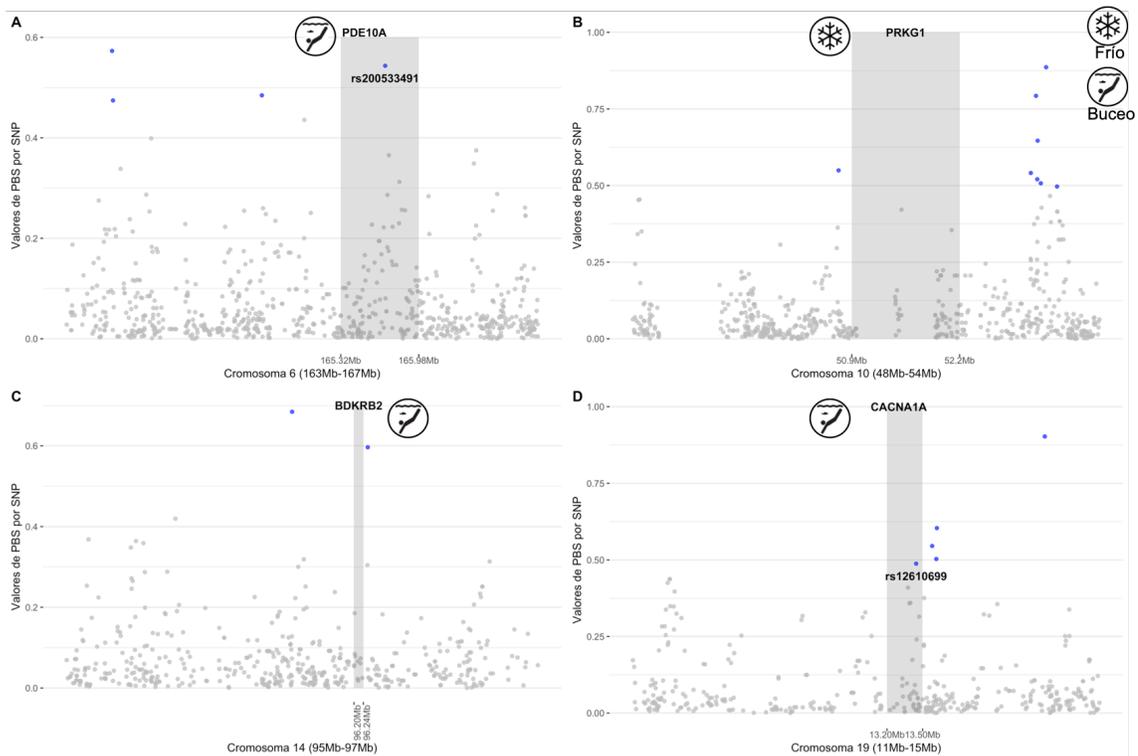


Figura 72. Comparación de valores de PBS por SNPs entre los cazadores recolectores marítimos de la Patagonia en regiones cromosómicas de genes descritos previamente en literatura para otras poblaciones nativas. Aquellos valores por sobre la distribución del 1% se encuentra en azul. El rectángulo gris corresponde al gen descrito a comparar. Cada símbolo al costado del gen corresponde a si esta asociado al frío, dieta y/o buceo.

Se realizó el mismo análisis con los resultados de iHS, sin embargo no fue posible detectar ventanas cercanas a las regiones descritas previamente en la literatura. Mas aún, quisimos evaluar la distribución de los SNPs con valores de $|iHS| > 2.5$ sobre estos genes, no obstante se encontró que todos los genes descritos se encontraban en regiones del genoma sin información debido al proceso de remoción de datos perdidos para este análisis. Esto se puede apreciar mediante la representación de algunos cromosomas en la Figura 73.

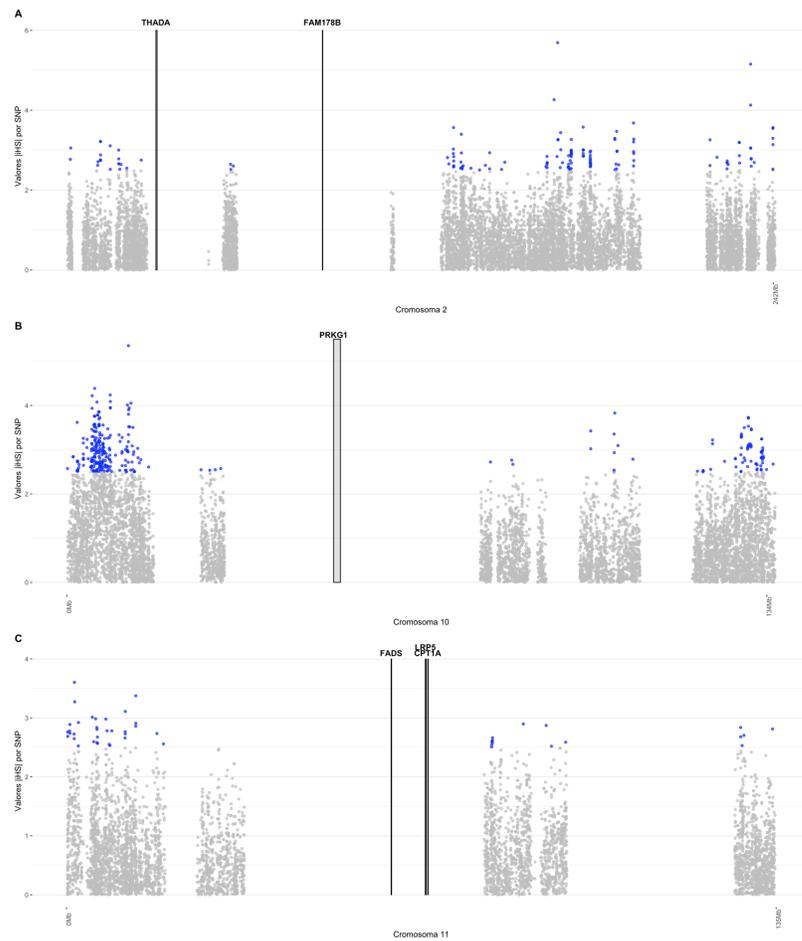


Figura 73. Análisis de iHS para SNPs en la Patagonia Austral. Los puntos de color azul corresponden a los valores de $|iHS| > 2.5$. Se muestran únicamente los cromosomas 2, 10 y 11 donde se ubican genes descritos previamente en Siberia bajo selección para frío y dieta. Los recuadros negros verticales corresponden a la ubicación física de estos genes.

5.8.4.2. Procesos y Rutas Metabólicas en la Patagonia Austral

Las rutas metabólicas que se destacan en los cazadores recolectores marítimos de la Patagonia Austral corresponde nuevamente a la síntesis y secreción del cortisol. Esta ruta ya había sido descrita anteriormente con los datos de Axiom, no obstante la descripción de genes que se encuentra participando en dicha ruta son en su mayoría distintos (Tabla XLIV), salvo por los genes PBX1, CREB1, SCARB1, PLCB1 y PLCB4 que aparecen en ambas comparaciones.

Tabla XLIV. Genes involucrados en la ruta metabólica de síntesis y secreción del cortisol en los individuos de la Patagonia.

Gen	Nombre Gen	Cromosoma	Inicio	Termino
PBX1	PBX homeobox 1	1	164559360	164886047
CREB1	cAMP responsive element binding protein 1	2	207529892	207605989
CACNA1D	calcium voltage-gated channel subunit alpha1 D	3	53495049	53813151
ITPR1	inositol 1,4,5-trisphosphate receptor type 1	3	4493348	4847840
ADCY1	adenylate cyclase 1	7	45574140	45723116
CREB3L2	cAMP responsive element binding protein 3 like 2	7	137874979	138002101
NR5A1	nuclear receptor subfamily 5 group A member 1	9	124481236	124507420
PLCB3	phospholipase C beta 3	11	64251523	64269452
CACNA1C	calcium voltage-gated channel subunit alpha1 C	12	1969552	2697949
SCARB1	scavenger receptor class B member 1	12	124777628	124863973
GNAS	GNAS complex locus	20	58839681	58911196
PLCB1	phospholipase C beta 1	20	8132265	8884900
PLCB4	phospholipase C beta 4	20	9068710	9480816

Por ultimo, una de las rutas que llama la atención y que se encontró en los análisis de iHS corresponde al señalamiento de la molécula de apelina (FDR > 0.05) (Tabla XLV). Esta molécula se desenvuelve en muchas funciones fisiológicas, de las cuales destacan mecanismos de vasodilatación, vasoconstricción y en regulación de la lipólisis en los adipocitos (Figura 74).

Tabla XLV. Genes involucrados en la ruta metabólica de señalamiento de apelina en los individuos de la Patagonia Austral.

Gen	Nombre Gen	Cromosoma	Inicio	Termino
GNB1	G protein subunit beta 1	1	1785285	1891117
ADCY10	adenylate cyclase 10	1	167809080	167916389
ITPR1	inositol 1,4,5-trisphosphate receptor type 1	3	4493348	4847840
CAMK4	calcium/calmodulin dependent protein kinase IV	5	111223653	111498503
APLNR	apelin receptor	11	57233577	57237453
KRAS	KRAS proto-oncogene, GTPase	12	25204789	25251003
BECN1	beclin 1	17	42810132	42824316
KLF2	Kruppel like factor 2	19	16324826	16328662
PLCB1	phospholipase C beta 1	20	8132265	8884900
PLCB4	phospholipase C beta 4	20	9068710	9480816

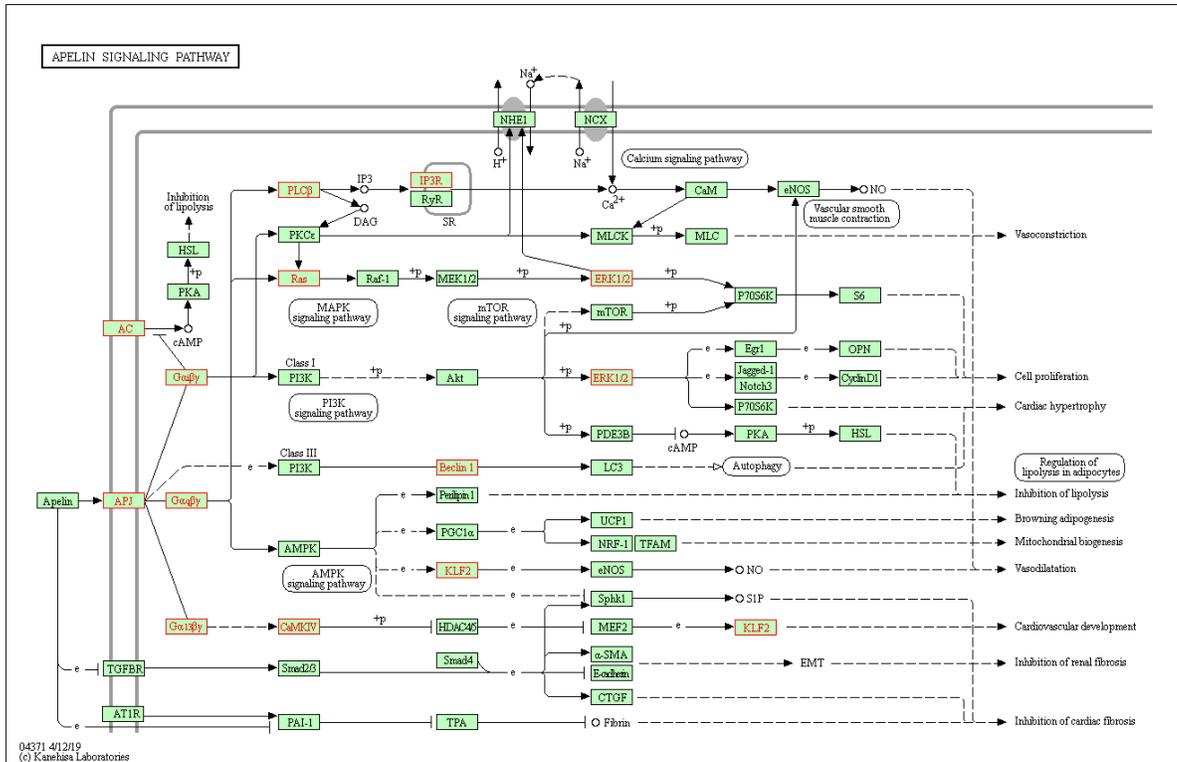


Figura 74. Ruta metabólica descrita para el señalamiento de la apelina (hsa04371) extraído a partir del Mapa de rutas KEGG (www.kegg.jp). Los recuadros en rojo corresponden a los genes sobre-expresados en los cazadores recolectores de la Patagonia.

6. Discusión

6.1. Estructuración Poblacional en Individuos de la Patagonia a través de Ancestrías

El análisis de las poblaciones actuales de la Patagonia constituye un desafío al momento de evaluar la estructuración genética y las posibles señales de selección asociadas a las poblaciones nativas que vivieron en el pasado en esta región. Esto se debe a que los grupos humanos que viven en la actualidad en el continente americano son producto de un proceso de mestizaje ocurrido principalmente entre los nativos americanos y los europeos, hace más de 500 años atrás (Valenzuela, 2011).

Esto se ve reflejado en una primera instancia al momento de realizar los análisis de estructuración genética mediado por ancestrías. Cuando observamos la Ancestría Global y el Análisis de Componentes Principales en las poblaciones de la Patagonia podemos ver que los principales componentes ancestrales que participan y/o explican nuestros datos corresponden a los Europeos y Nativos Americanos. Para las poblaciones de la Patagonia, en promedio, el aporte del primer componente es de un 28% para nuestras muestras, un valor bajo si lo comparamos con otros datos de ancestría en diferentes poblaciones de Chile que poseen un rango entre el 49-57% aproximadamente (Eyheramendy et al., 2015). Esta relación cambia aún más para el caso de las poblaciones Kawéskar y Yámana de la Patagonia Austral (genotipificados con MEGA) donde vemos que el componente Europeo no supera el 18%.

La diferencia de estos valores podría tener relación con las diferentes interacciones de los asentamientos europeos y sus constantes conflictos con las poblaciones nativas donde actualmente se encuentran las poblaciones urbanas (Via et al., 2011; S. Wang et al., 2008). No obstante, cabe mencionar que el origen de la muestra de la Patagonia es distinta, puesto que se eligieron individuos *a priori* con posible descendencia directa de estos grupos Nativos. Esto nos puede dar otra perspectiva acerca de la muestra, con un aporte Europeo que dado los filtros

posteriores puede ser prueba de eventos en el pasado, probablemente durante el contacto con los Europeos hace 500 años atrás. Las localidades como Laitec, Puerto Edén o Villa Ukika no serían un foco principal de desarrollo urbano ni de inmigración, y su grado de aislamiento es considerable, por lo que es esperable que no hayan sufrido cambios en su composición genética posterior al proceso de Conquista. Por ende, el componente Europeo encontrado en estas localidades independiente de la elección de los individuos puede reflejar los primeros contactos entre ambos grupos ancestrales desde el momento de su llegada. Esto podría estar representado particularmente en el caso de Villa Ukika y Puerto Edén, los cuales poseen un 5.4% y 4.5% respectivamente.

Al momento de descomponer la ancestría Nativa Americana, y dejando por un momento de lado el componente Europeo, podemos ver claramente una subestructuración conformada por un componente asociado a Aymara (Puno) y otro relacionado a los Mapuches (Pehuenche/Huilliche). Esta situación la vemos para la Ancestría Global (K=5) en los individuos de la Patagonia, donde poseen una elevada afinidad con los grupos Nativos Mapuche, indicando posiblemente un componente ancestral en común para estas poblaciones, como se ha mencionado anteriormente en análisis de ADN mitocondrial (de Saint Pierre, Bravi, et al., 2012).

En relación a esto, se encontró evidencia de un componente andino, el cual está mayormente representado en Tehuelche (9.8%) y Yámana (10,9%), mientras que aparece escasamente en individuos Chono (Laitec, 0.6%) en la Patagonia Occidental y prácticamente ausente en Pehuenche y Huilliche del sur de Chile. Estos componentes Mapuche y Aymara, provenientes de dos rutas que se separaron al momento de entrar al continente sudamericano contribuirían para el poblamiento de la Patagonia (de Saint Pierre et al., 2012; de Saint Pierre et al., 2012b; Rothhammer & Dillehay, 2009). Sin embargo, cabe mencionar que esta asociación por parte de la ancestría global (K=5) entre las ancestrías Mapuche y Aymara hace referencia mas bien a un grupo patagónico ancestral que no se encuentra representado en nuestra base de datos para este tipo de análisis, y que pudiera estar

dando cuenta de diferentes procesos del poblamiento de la Patagonia. De hecho, se ha estimado que la separación entre las poblaciones del altiplano andino con las costeras ocurrió entre los 9200 y 8200 años atrás (Lindo et al., 2018), tiempo por el cual ya se había poblado la Patagonia.

De acuerdo a esto último, se pueden plantear diferentes hipótesis en relación al poblamiento de la Patagonia: i) un rápido poblamiento temprano de la Patagonia a través de una ruta costera del Pacífico por individuos que aun no estarían diferenciados genéticamente entre Mapuche o Andino. Estas poblaciones se mantuvieron en pequeños tamaños, pudiendo reflejar aún su composición ancestral al momento de poblar la Patagonia. Por otro lado, tanto las poblaciones andinas como las Mapuche experimentaron grandes crecimientos poblacionales, lo que les permitiría acumular mayores diferencias entre estas posteriormente. ii) las poblaciones de la Patagonia presentan en su genoma la composición genética de un grupo ancestral formado por una mezcla principalmente entre el grupo costero proveniente del pacífico con un grupo proveniente del Altiplano Andino. Por ultimo, iii) una más reciente que indicaría que el componente andino correspondería a un contacto posterior al poblamiento de la Patagonia. No obstante, para poder responder estas hipótesis es necesario abordarlo con diferentes aproximaciones metodológicas que permitan acercarnos a una explicación acerca del poblamiento de la Patagonia.

En relación a estas distintas hipótesis, un rápido poblamiento costero mediante el pacífico cobra mayor importancia debido a una investigación basada en el análisis de ADN antiguo en poblaciones Kawéskar y Yámana de la Patagonia Austral (De la Fuente et al 2018). En estos individuos se encontró un componente en bajas proporciones asociado a diferentes grupos nativos de origen Andinos y del norte de Suramérica. Esto indica que la ancestría Andina visto en nuestros datos corresponde mas bien a un componente ancestral de las poblaciones Sudamericanas y que es la que se encuentra presente en los individuos de la Patagonia. Además, de acuerdo a análisis en genomas mitocondriales no fue posible encontrar ningún linaje que

conecte a las poblaciones de la Patagonia de manera directa con las poblaciones Andinas, sino mas bien de un grupo ancestral tanto para la Patagonia Chilena como Argentina (de Saint Pierre et al., 2012).

La incorporación de individuos provenientes de la Patagonia Austral permitió detectar un nuevo componente de origen Nativo Americano. Este componente se diferencia de las ancestrías del centro sur de Chile y del Altiplano Andino, estando presente casi exclusivamente en individuos Yámana (Villa Ukika) y Kawéskar (Puerto Edén y Punta Arenas). El resultado de esta diferenciación genética para las poblaciones descendientes de cazadores recolectores marítimos de la Patagonia estaría dada por el aislamiento que sufrieron estos grupos producto de los cambios climáticos durante el Último Máximo Glacial, entre estos, el aumento del nivel del mar y la formación del Estrecho de Magallanes hace 8,000 años atrás (Quattrocchio et al., 2011). Esta situación separó principalmente a las poblaciones Yámana del resto del continente, lo que generó una acumulación de variantes genéticas, y por ende un nuevo componente ancestral. El hecho de presenciar esta ancestría Austral en los grupos Kawéskar indicaría una comunicación entre estos grupos canoeros al oeste de la Patagonia. Finalmente, a pesar de la descripción de esta nueva caracterización genética, es necesario poder evaluar en el resto de las poblaciones de la Patagonia, dado que solo se incluyeron dos individuos Tehuelches y cinco de Laitec para este set de datos. En esta última localidad se encontró una fracción de este componente, lo que pudiera estar indicando una conexión entre el límite norte de la distribución Kawéskar con los extintos Chono.

6.2. Procesos de Mestizaje en las Poblaciones de la Patagonia post-contacto Europeo

La estimación de las ancestrías globales para los individuos nos ayuda a obtener una aproximación frente a cuales son los principales componentes ancestrales que participan en la muestra. No obstante, es importante también poder inferir las ancestrías de los distintos segmentos cromosómicos asociados a una ancestría en específico. Tal análisis se realizó con RFMix (Maples et al., 2013) y se tomaron como

referencia las ancestrías Europeas (IBS) y Nativas Americanas (Pehuenche/Huilliche) como eje central. Como un modo de evaluar la precisión de este método se realizó una correlación con lo que se obtuvo para la ancestría global. Esto se realizó para dos aproximaciones: 1) con las ancestrías europeas, nativo americanas y africanas y 2) con las ancestrías europeas, Mapuche y Aymara. Para la primera comparación los valores fueron del 99%, indicando una elevada correlación para este análisis salvo para la ancestría africana. Esto último puede deberse a que las frecuencias para este componente son demasiado bajas en nuestros individuos, posiblemente no diferentes de cero. Para la segunda aproximación vemos una disminución en la correlación para el componente Aymara y Mapuche con un 93.9 y 72.7%. Como se puede apreciar en los resultados, esta sobre-estimación para el componente Aymara se debe a que el análisis de Ancestría local, implementado por RfMix no permite asignar de manera óptima las ancestrías Aymara/Mapuche puesto su poco tiempo de divergencia. Es por esto que no sería recomendable utilizar este análisis en poblaciones ancestrales al menos con 9000 años de divergencia.

Los diversos escenarios demográficos que han acontecido en Latinoamérica posterior a la llegada de los españoles han demostrado ser complejos y diferentes dependiendo de la región donde esta ocurrió (Gravel, Zakharia, Moreno-Estrada, et al., 2013; Homburger et al., 2015). Por esta razón, conocer los posibles cambios en la estructura genética de los grupos de cazadores recolectores de la Patagonia tras el contacto Europeo son de vital importancia para entender como se generó el proceso de mezcla y como pueden responder frente a diferentes rasgos complejos, incluso a respuestas evolutivas. Para esto, se aplicó el modelo extendido markoviano implementado en TRACTs para comparar los datos observados con las predicciones realizadas a partir de diferentes modelos demográficos considerando varios escenarios de migración. El ajuste de los modelos arrojó diferencias para las poblaciones en la Patagonia, lo que en una primera aproximación nos indica la existencia de diferentes escenarios migratorios.

El inicio del mestizaje mas tardío en las poblaciones de la Patagonia Chilena ocurrió en los Kawéskar hace 11 generaciones atrás (275 años), sin embargo el análisis para esta población esta hecha solo para 4 individuos de Punta Arenas. Por esta razón, se tomara en cuenta el tiempo de mestizaje entregado por los individuos Kawéskar de Punta Arenas y Puerto Edén genotipificados por MEGA, los cuales arrojaron un tiempo de 9 generaciones atrás (225 años). Para las localidades con ascendencia Chono (Laitec) y Villa Ukika se determinó un tiempo de mestizaje de 10 (250 años) y 9 (225 años) generaciones respectivamente. Podemos ver que para estas localidades de la Patagonia Chilena, el tiempo de mestizaje es bastante similar, no obstante, difieren a los encontrado para Santiago (Centro de Chile), la cual remonta el inicio del mestizaje hace 14 generaciones atrás (350 años) con un pulso Africano y posteriormente uno de origen Europeo (Homburger et al., 2015). Distinto es lo que se encontró para otras poblaciones Chilenas con un tiempo de mestizaje de 10 generaciones (Eyheramendy et al., 2015), no obstante, debido a que no se consideraron modelos demográficos, es posible que flujos europeos posteriores hayan contribuido a una sub-estimación del inicio de la mezcla. En el caso de la población Tehuelche, el tiempo de mezcla inicial es de 11 generaciones atrás (275 años) con un posterior pulso Africano y posteriormente Europeo. Esto guarda relación con los datos encontrados para otras ciudades de Argentina (Homburger et al., 2015).

Todo esto nos podría suponer que el mestizaje fue mucho más tardío en las poblaciones de la Patagonia (exceptuando Tehuelches), indicando un contacto que pudo darse por migraciones posteriores al proceso de conquista, dado que estas zonas no eran el foco de asentamientos españoles, también puede ser producto de un contacto distinto mediado por otros grupos europeos. Además, las frecuencias de estos grupos ancestrales se han mantenido constante desde el inicio del mestizaje, lo cual nos estaría indicando que los segmentos europeos que vemos actualmente, corresponderían al pool génico que dio origen a las poblaciones mestizas, y que se ajustan además con las frecuencias que vemos para la ancestría global.

Se sabe que los primeros contactos entre los grupos Kawéskar y extranjeros datan del siglo XVI. Desde entonces, comenzaron a frecuentar la zona cazadores de pieles, balleneros, piratas y corsarios, atraídos por los galeones españoles que estaban obligados a pasar por el Estrecho de Magallanes. Posteriormente, a partir del siglo XVIII el paso de los navegantes se llevó a cabo por el Cabo de Hornos. No es sino a fines del siglo XVIII donde el contacto entre extranjeros y Kawéskar se fueron haciendo mas frecuentes (Gleisner & Sara Montt, 2014). Estas descripciones se ajustan de manera bastante optima con la inclusión del componente Europeo en los diferentes grupos de la Patagonia. De acuerdo a los datos genéticos, la mezcla en la Patagonia se originó a principios del 1700 (siglo XVIII), lo cual concuerda con lo descrito por los historiadores para esta región del continente.

A partir de este punto, se tuvo que tomar una estrategia distinta al momento de evaluar señales de selección en los grupos de la Patagonia, puesto que la mayoría de estos están pensados para individuos con solo un tipo de ancestría (Acuna-Alonzo et al., 2010; Schlebusch et al., 2015; Valverde et al., 2015), exceptuando aquellos que evalúan señales de selección mediado por ancestrías, como por ejemplo posterior a la llegada de los Europeos (Bhatia et al., 2014; Pierron et al., 2018; Reynolds et al., 2019). Puesto que el objetivo de este trabajo es evaluar regiones bajo selección pre y post contacto con los Europeos, se tuvo que rescatar únicamente aquellos segmentos del genoma de origen Nativo Americano. Este flujo metodológico se realizó de manera novedosa en este trabajo y constituye una parte fundamental para este estudio. Es indudable que las poblaciones Nativas sufrieron una reducción importante desde la llegada de los Europeos (K. Patterson & Runge, 2002; R. Thornton, 1997), perdiendose mucha información genética de estas poblaciones. Es por esta razón, que poder rescatar el genoma Nativo Americano de los individuos actuales para inferir escenarios que involucren señales de selección previo al contacto Europeo son de gran utilidad para poder comprender la historia evolutiva de las poblaciones nativas. Este es uno de los alcances fundamentales que podemos

desprender de este estudio, y que necesita extenderse al resto de las poblaciones mestizas a lo largo del continente.

6.3. Análisis de Selección

Dada la naturaleza de nuestras poblaciones actuales, se hace necesario poder implementar diferentes metodologías al momento de detectar señales de selección. En este sentido, tanto PBS como iHS nos permiten obtener información acerca de posibles señales de selección que pudieron haber estado jugando un rol en el pasado. Por otro lado, también se hace necesario poder evaluar nuevas regiones bajo selección producto del contacto Europeo durante la conquista. Para este último caso, el mapeo genómico en busca de loci en selección resulta ser el más eficaz para este panorama (Bhatia et al., 2014). En una primera aproximación, los resultados indican que se han detectado diferentes regiones del genoma bajo selección, dependiendo si es pre o post contacto con los Europeos. Esto nos da una idea en donde las presiones selectivas eran diferentes a las que se generaron posteriormente a la llegada de los Europeos en los grupos cazadores recolectores de la Patagonia.

6.3.1. Señales de Selección pre-contacto Europeo

6.3.1.1. Cazadores Recolectores de la Patagonia

De acuerdo a los resultados que se obtuvieron para PBS e iHS, podemos identificar varias señales de selección en los grupos cazadores recolectores de la Patagonia previo al momento de la llegada de los Europeos. Dentro del top 10 de las señales con mayor puntaje pudimos encontrar varias de estas relacionadas con nuestras hipótesis frente al frío y dieta marina, estando la mayoría de ellas concentrada en el cromosoma 2.

Cuando nos referimos a los grupos cazadores recolectores de la Patagonia, vemos que existen varios genes bajo selección que pudieran estar destacando sobre el resto. Por ejemplo, se ha demostrado en ratones que el gen **ASXL2** ayuda a la

adipogénesis y regula la homeostasis de lípidos, glucosa y el esqueleto, su alteración se encuentra asociado con enfermedades asociados a los huesos y obesidad. La deficiencia en la expresión de este gen origina en ratones lipodistrofia y falta de respuesta frente a dieta elevadas en grasas (Izawa et al., 2015). Si bien no existe demasiada información acerca de la repercusión en humanos, es posible que pueda estar jugando un rol asociado con el metabolismo de grasas y desarrollo de adipocitos en las poblaciones de la Patagonia. Por otro lado, un estudio señala una asociación entre el gen **LRP2** con los niveles de colesterol y colesterol LDL en el plasma sanguíneo, además se encontraron variantes asociados a pacientes con hipercolesterolemia (Mii et al., 2007). Esto podría estar directamente relacionado con la exposición al frío, puesto que ya se ha descrito anteriormente al gen **LRP5** bajo selección positiva en poblaciones del noreste de Siberia (Clemente, Cardona, Inchley, Peter, Jacobs, Pagani, et al., 2014).

De acuerdo a los resultados, existe una región en el cromosoma 11 (123.6-123.8Mb) que se encuentra bajo selección tanto para los cazadores recolectores como cazadores recolectores marítimos de la Patagonia. Esta ventana contiene a los genes **GRAMD1B** y **ZNF202**. Para este último gen (Zinc Finger Protein 202), se ha descrito que corresponde a un represor transcripcional que se une a la región regulatoria de muchos genes relacionados con el procesamiento de lípidos y homeostásis energética. Además, estaría relacionado con dislipidemia, diabetes, y obesidad (Schmitz et al., 2004; Wagner et al., 2000). Distintas variantes se han asociado en la regulación de los niveles de colesterol HDL (lipoproteínas de alta densidad) en diferentes poblaciones de Estados Unidos y en el norte de Europa (Razzaghi et al., 2012; Stene et al., 2006), así como también, en aterosclerosis severa, enfermedad isquémica coronaria (Stene et al., 2008) y con una fuerte asociación con un trastorno en el metabolismo caracterizado por la ausencia total de apolipoproteína y una concentración muy baja de lipoproteínas de alta densidad (HDL) en el plasma (Porsch-Özcürümez et al., 2001). De acuerdo a lo anterior,

podemos inferir un rol importante de este gen en el metabolismo de lípidos y equilibrio energético en las poblaciones de la Patagonia.

Por otro lado, el gen **ADAM17** ubicado en el cromosoma 2 participa en múltiples procesos asociados con regulación celular sobre diversos substratos: citoquinas, factores de crecimiento, y sus receptores así como también las moléculas de adhesión se activan o inactivan por su clivaje con **ADAM17** (Monika Gooz, 2010). También ha sido asociado mediante GWAS a diferentes patologías como cáncer, resistencia a la insulina, riesgo de obesidad, enfermedades cardiovasculares, artritis reumatoide y la enfermedad de Alzheimer (Junyent et al., 2010; Monika Gooz, 2010; Zunke & Rose-John, 2017). En relación a esto, una de sus funciones cobra especial interés, puesto que este gen actúa en el tejido adiposo mediante la regulación del metabolismo de lípidos (Liu et al., 2009; Menghini et al., 2013). Específicamente, se encontró una variante (i33708A>G) que interactúa directamente con riesgo de obesidad de acuerdo a la ingesta de una dieta con (n-6) ácidos grasos poli-insaturados (PUFA) (Junyent et al., 2010). Esta interacción estaría directamente involucrada con los mecanismos de respuesta al frío, y posiblemente en conjunto con el consumo de ácidos grasos poli-insaturados. La termogénesis inducida por el frío es un mecanismo que permite generar calor mediante un aumento en el metabolismo corporal en respuesta a temperaturas ambientales frías (Brychta & Chen, 2017). En relación a lo anterior, el gen **ADAM17** podría jugando un rol importante en estos procesos metabólicos, actuando en respuesta a las presiones selectivas en la Patagonia. Señales de selección para este gen no se han descrito por lo que podría ser una respuesta nueva en estos grupos cazadores recolectores.

6.3.1.2. Cazadores Recolectores Marítimos de la Patagonia

Continuando con los cazadores recolectores marítimos de la Patagonia, encontramos dos señales bajo selección solamente en estos grupos, los cuales corresponden a **GFRAL** (cromosoma 6) y **WNT9B** (cromosoma 17). De estos, ha habido más información para el primer gen y que puede estar confiriendo una clara

respuesta selectiva en estos grupos. **GFRAL** corresponde a una familia de receptores en el cerebro el cual se ve regulado directamente por GDF15 (Breit, Tsai, & Brown, 2017), un biomarcador de estrés celular que participa en múltiples procesos biológicos, incluyendo homeostásis energética, regulación de peso corporal, cáncer y enfermedades crónicas (Mullican & Rangwala, 2018). Se ha encontrado que en condiciones de enfermedad o estrés, los animales utilizan una ruta neuronal alternativa para responder a cambios metabólicos frente a una alteración sobre la demanda energética (J. Y. Hsu et al., 2017). Los niveles de GDF15 aumentan significativamente cuando los ratones están expuestos crónicamente a alimentos elevados en grasa, generando una respuesta celular de estrés integrada frente al exceso de dieta (Patel et al., 2019). Esto provoca que la interacción de “emergencia” GDF15-GFRAL modele las respuestas frente a condiciones de estrés metabólico, regulando el consumo de comida, gasto energético y peso corporal (J. Y. Hsu et al., 2017; Yang et al., 2017).

6.3.1.3. Poblaciones Nativas Kawéskar y Yámana (Patagonia Austral)

Ahora bien, cuando queremos adentrarnos en el análisis de las poblaciones canoeras Kawéskar y Yámana, encontramos varias señales de selección dentro de las mas fuertes en nuestros datos. Estas se encuentran en genes que parecen tener un efecto directo en la metabolización de lípidos, transporte de electrones en la mitocondria, entre otros, sin embargo no hay suficientes evidencias que permitan asociarlos directamente como responsables de adaptación. Sin embargo existen otros que pueden darnos argumentos necesarios para proponerlos como candidatos directos a selección. Por ejemplo se ha descrito que la expresión de los niveles de **SLC25A12** descendieron con el consumo de ácidos grasos de origen marinos en poblaciones asiáticas (Rundblad et al., 2018). Por otro lado, se conocen algunas características del gen **UCN3**, el cual participa en la degradación de lípidos bajo condiciones de hipoxia (Xiong et al., 2014). Los UCNs (urocortina) corresponden a péptidos sintetizados endógenamente en el corazón y el sistema vasculatorio,

participando en la protección de estrés cardiovascular, como la hipoxia (Takahashi, 2012). En adición, este gen ha aparecido como un candidato a selección en poblaciones Andinas enfrentados a una hipoxia moderada (Eichstaedt, et al., 2015), por lo que el hecho que aparezca con una potente señal de adaptación en poblaciones cazadoras recolectoras marítimas de la Patagonia nos podría estar indicando un proceso de adaptación biológica en respuesta a los bajos niveles de de oxígeno producto de las inmersiones y prácticas de buceo en estas poblaciones. De hecho, para los grupos Kawéskar se ha descrito que las mujeres eran excelentes buceadoras, y se lanzaban al mar principalmente en busca de erizos y moluscos, permaneciendo por más de una hora en el agua (Gleisner & Sara Montt, 2014). Esto sugiere que durante el proceso de cambio a vida marítima, hubo una fertilidad diferenciada de algunos genotipos con variantes en genes que permitieron una mejor respuesta relacionadas con el buceo.

La señal de selección sobre el gen **TRPM8** es una de las que más llama la atención debido a que ha sido descrito como el principal protagonista en la respuesta a ambientes fríos en los mamíferos, el cual se originó en el ancestro común más reciente de los Tetrápodos (Blanquart et al., 2019; Saito & Shingai, 2006). A nivel molecular, la sensación de temperatura se debe a una activación de los canales iónicos de la familia TRP, dentro del cual **TRPM8** juega un rol en la sensación de frío y posteriores termorregulaciones fisiológicas (Bautista et al., 2007). Su activación en los nervios aferentes inician potenciales de acción que se propagan a través del sistema nervioso central (SNC). En homeotermos, el SNC compara el exterior con la temperatura del cuerpo antes de inducir el proceso de termogénesis (Rial & Zardoya, 2009). Diversas mutaciones en este gen son responsables por una reducción en la sensación de frío en roedores que hibernan en comparación con aquellos que no lo hacen (Matos-Cruz et al., 2017). En distintas poblaciones humanas, se detectó la variante rs10166942 ubicado a 1Kb río arriba de este gen y que presenta diferencias alélicas extremas, las cuales se correlacionan con latitud y temperatura. Estas frecuencias poseen un rango de frecuencias del 5% en Nigeria

contra un 88% en Finlandia, lo cual sugiere un evento de adaptación local hacia los sectores septentrionales de Europa y Asia (Key et al., 2018). Ahora bien, dado que no hay registros en torno a las frecuencias de esta variante en poblaciones de Siberia, América, incluso en nuestros datos, no podemos corroborar si la selección se llevo a cabo en esta misma variante, o mas bien corresponde a una mutacion *de novo* en las poblaciones de la Patagonia. No obstante, la variante que encontramos bajo selección, rs17869069 de naturaleza intrónica se encuentra a 1Mb de distancia de rs10166942. En relación a esto, los autores señalan que si bien hay una marcada diferencia alélica para el alelo T de rs10166942 en **TRPM8** en un gradiente latitudinal, no se puede descartar que la selección se ejerció sobre una región desconocida, pero fuertemente ligada con este alelo. Lo anterior nos podría dar una clara señal que la región que estamos detectando en la Patagonia Austral sea la misma. Las variaciones en la temperatura a lo largo del mundo han modelado fuertemente la distribución de variantes bajo selección en humanos (A. M. Hancock et al., 2011), esto nos permite respaldar estos patrones de variación incluyendo una de las últimas regiones en ser colonizadas por el humano moderno.

6.3.1.4. Procesos Biológicos y Rutas Metabólicas en las Poblaciones de la Patagonia

Cuando vemos los procesos biológicos sobrerrepresentados podemos ver varios relacionados con el procesamiento de lípidos, glicógeno, regulación de insulina, diferenciación de tejidos celulares como el adiposo, regulación positiva y negativa de la termogénesis. Estos procesos se ajustan con los genes descritos anteriormente, y su interacción tendría una estrecha respuesta con mecanismos de frío y dieta marítima. Sin embargo, no solo existen asociaciones directas, sino que también podemos ver algunos procesos biológicos que participan como intermediarios. Por ejemplo, varios procesos están relacionados con la respuesta celular al ácido retinoico, un metabolito que se encuentra mas indirectamente relacionado con el metabolismo de lípidos (Okla et al., 2017). También, se detectaron

procesos asociados con el metabolismo y síntesis de la molécula carnitina, entregados principalmente por los análisis de iHS. Esta es una molécula endógena asociada con el metabolismo de ácidos grasos y esencialmente con el transporte de cadenas de ácidos grasos hacia la matriz mitocondrial, permitiendo así metabolizar la grasa y obtener energía en las células adiposas. Varios alimentos son fuente de carnitina, principalmente las carnes rojas, aunque también está presente en pescados y aves (Pekala et al., 2011). Anomalías con esta molécula también puede derivar en resistencia a la insulina y diabetes mellitus tipo 2 (Bene et al., 2018).

A pesar de esto, existen muchos procesos biológicos sobrerrepresentados con características que a primera vista no estaban relacionados directamente con nuestras hipótesis, lo que puede dar cuenta de que estamos frente a otras señales de selección, o mas bien son parte de la misma respuesta adaptativa que estamos evaluando, pero ejerciendo un efecto intermediario o indirecto. Por esta razón, una de las proyecciones necesarias corresponde a la determinación de rutas biológicas que puedan incorporar varias funciones que por separado no estén relacionadas, pero que en conjunto puedan otorgarnos mayor información para un metabolismo en específico.

Una de las principales rutas arrojadas por la mayoría de las comparaciones corresponde a la síntesis y secreción de la molécula de cortisol, la cual está relacionada con la adaptación al estrés y el metabolismo de ácidos grasos y glucosa. El estrés guarda relación con la disrupción en la homeostasis del individuo lo que se desencadena en un estado de excitación, alerta y vigilancia. Esta “respuesta al estrés” se refleja en la rápida activación del sistema nervioso simpático, el cual lidera la liberación de noradrelina y adrenalina de la médula adrenal, y por consiguiente se genera un aumento de glucocorticoides en la sangre (Kloet et al., 2005). La alteración en la cantidad de consumo en ácidos grasos insaturados n-3 (Omega-3) está relacionado con la regulación el estrés biológico asociado al cortisol, una elevada cantidad de estos ácidos grasos, encontrados principalmente en pescados y mariscos reduce el estrés en tres sistemas principales: i) el sistema inmune-inflamatorio, ii) el

eje adrenal-pituitario-hipotalámico y iii) el sistema nervioso autónomo (Thesing et al., 2019). Esto se caracteriza por un impedimento en la producción de glucocorticoides (Xie et al., 2016).

Esto sin duda nos puede estar diciendo que las condiciones ambientales de frío gatillaron respuestas evolutivas sobre diferentes genes que participan en la ruta metabólica del cortisol en las poblaciones de la Patagonia. Este proceso se encuentra ligado a su vez con la regulación del colesterol y la ingesta de PUFAs (Omega-3) a través de la dieta. Por ende, los mecanismos de respuestas frente a estas condiciones ambientales extremas y modos de vida estarían regulados por una red de genes. Para las poblaciones de la Patagonia se puede indicar una intrincada red de rutas metabólicas asociadas con el colesterol, consumo y metabolismo de ácidos grasos, y modificaciones frente al estrés biológico. Esta red genética se detectó igualmente para algunos genes bajo selección que en su conjunto participan en el proceso de termogénesis.

6.3.1.5. Comparación de Regiones bajo Selección en la Patagonia con genes asociados a condiciones de frío y dieta

En los resultados se compararon las señales bajo selección para cada ventana de 200Kb con las regiones en donde se encontraron genes descritos previamente en otras poblaciones nativas, principalmente de Siberia. Durante el aislamiento de estos grupos en el ártico, se generaron ciertas señales de selección asociadas a frío y dieta marina. Estos genes corresponden principalmente a la región **FADS (FADS1, FASD2, FASD3), THADA, PRKG1, CPT1A y LRP5** (Cardona et al., 2014; Clemente et al., 2014; Fumagalli et al., 2015). Los resultados mas importantes tienen relación con el patrón observado para el cromosoma 11, específicamente entre los 61-70Mb, región donde se encuentra la mayoría de los genes asociados a frío y dieta en las poblaciones de Siberia (**FADS1, FADS2, FADS3, LRP5, CPT1A**). Podemos observar la presencia de múltiples señales bajo selección en regiones próximas a los genes **FADS** y **CPT1A**, lo cual se ve mucho mas intensificado en grupos cazadores recolectores marítimos

de la Patagonia (representados por Axiom y MEGA), sin embargo de menor intensidad y fuera del Top 10 de las señales de selección mas fuertes en nuestras poblaciones. Esta región del cromosoma ha sido descrita como una zona importante frente a mecanismos de respuesta asociados al frío y al tipo de dieta marítima (Cardona et al., 2014; Clemente et al., 2014; Fumagalli et al., 2015). De hecho, pudimos identificar un bloque de aproximadamente 2,8Mb (62,7-65,5Mb) bajo selección presente en todos los análisis que se realizaron. En esta región se encontraron varios genes, asociados a metabolismo de lípidos, hipoxia, glucosa y glicógeno, de los cuales quizás los mas importantes correspondan a **PLAAT4**, **PLAAT4** y **PLAAT5**.

Se han encontrado diferencias en las frecuencias de algunos haplotipos asociados a la región **FADS** muestran una señal de selección no solo en el Ártico, sino que también en diversas poblaciones nativas a lo largo del continente Americano, incluyendo poblaciones de la Amazonia las cuales poseen condiciones extremadamente diferentes en torno al clima, cultura y dieta (Eduardo et al., 2017). Los patrones observados para este cromosoma en los grupos cazadores recolectores de la Patagonia, y en el resto del continente podrían corresponder a una huella de un fuerte evento pasado de adaptación ocurrido en Beringia, previo a la expansión de los ancestros Nativos durante el poblamiento de América y Groenlandia. Las dataciones mas antiguas de ocupación humana en Sudamérica se encuentran en Monte Verde hace 15,000 años cal AP (Dillehay et al., 2008), por otro lado, los datos en genomas mitocondriales indicarían una rápida expansión desde Beringia a Tierra del Fuego que tomó cerca de 2000 años (de Saint Pierre et al., 2012), lo que nos permitiría explicar la presencia de estas señales en las poblaciones de la Patagonia. Sin desmedro a lo anterior, la presencia de algunos SNPs bajo selección en los genes **FADS3** y **CPT1A** podrían participar en respuestas adaptativa al frío o en la dieta marítima en estos grupos de la Patagonia, producto que estas señales no se alcanzaron a perder durante el poblamiento Americano. Sin embargo estas señales no serían lo suficientemente fuertes como para explicar los procesos de adaptación

frente a las condiciones de frío o dieta de tipo marítimo posterior al poblamiento de la Patagonia en comparación con los nuevos genes candidatos que se describieron anteriormente (**ADAM17, ZNF202, GFRAL y TRPM8**).

Por otro lado, la detección de regiones bajo selección cercanas a genes como **BDKRB2** y **PDE10A**, ambos asociados a funciones fisiológicas relacionadas con buceo (M. Ilardo & Nielsen, 2018), nos podría estar indicando que se generaron respuestas similares de adaptación frente al buceo en poblaciones que divergieron hace más de 60,000 años. No obstante, es necesario poder corroborar estos análisis y evaluar si estas regiones se encuentran ligadas con estos genes. Sin embargo, como se mencionó anteriormente se encontraron funciones biológicas en genes bajo selección, las cuales están asociadas con regulación en la presión arterial y desarrollo del oído interno. Esto nos hace suponer preliminarmente que las poblaciones de la Patagonia poseen mecanismos adaptativos en respuesta al buceo.

Para terminar, es importante mencionar los alcances al momento de enmascarar los segmentos del genoma de origen Europeo o Africano para los análisis de selección, especialmente con iHS. Esto se debe principalmente a que este análisis utiliza la información de los SNPs vecinos en una distancia específica (Sabeti et al., 2007), contraponiéndose con el enmascaramiento y traduciéndose en una reducción de posibles señales de selección, o ausencia de estas. Aun con esto en consideración, se pudieron detectar fuertes señales, especialmente en nuevos genes candidatos y otras regiones cercanas a los que ya se habían descrito previamente. Para validar esto, se requiere analizar otras muestras, así como también genomas completos y/o antiguos. En lo que respecta a iHS, el filtro relacionado con los datos perdidos mermó de sobremanera el número de individuos para el análisis, por lo que no se removieron posteriormente individuos por parentesco para no ver afectado el tamaño de la muestra. Sin duda alguna, esto genera otra vulnerabilidad en los supuestos para los análisis de selección. Sin embargo, existen argumentos para validar las señales de selección que estamos detectando, por ejemplo vemos una región bajo selección que se repite entre los datos analizados por Axiom y MEGA

(ZNF202, GRAMD1B), por otro lado, probablemente las señales que veamos en estos datos representen a la familia de esta localidad, lo cual debido al bajo tamaño poblacional de los grupos Kawéskar, pudiese también estar representando una señal para estos grupos. En el otro extremo, el análisis de PBS no se vería afectado por el enmascaramiento, por lo que se propone como un método óptimo para analizar señales de selección asociados a una ancestría en particular en poblaciones mestizas.

6.3.2. Señales de Selección post-contacto Europeo

Sin duda alguna la llegada de los europeos al continente americano generó muchos cambios en las poblaciones nativas, productos de los cambios en la estructura social, guerras durante la conquista e introducción de patógenos (R. Thornton, 1997; Via et al., 2011). Esto trajo consigo de manera inmediata una drástica reducción en el tamaño poblacional de más del 90% en estos grupos nativos en la mayor parte del continente (Salzano & Maria, 2002).

Frente a esto, la región del cromosoma 6 que contiene los genes asociados con HLA (Antígenos Leucocitarios Humanos) ha jugado un rol clave frente a la respuesta inmune de patógenos durante los últimos 500 años a lo largo de todo el continente Americano. En algunas poblaciones nativas de la costa noroeste de Norte América se encontró una señal de selección en el gen HLA-DQA1 previa al contacto Europeo, y que pudo haberse convertido en señal de selección negativa después del contacto (Nakagome et al., 2016; Reynolds et al., 2019). En otras poblaciones de naturaleza andina, se han encontrado variantes con una ventaja adaptativa en el gen CD83, asociada con respuesta a la viruela (Lindo et al., 2018). En lo que respecta a diversas poblaciones de carácter mestiza, se ha descrito un exceso de ancestría Africana asociado a la región HLA mediante el uso de micro-satélites, indicando una fuerte señal de selección, posterior a la llegada de los Europeos (Deng et al., 2016). Por último, los mecanismos de respuesta adaptativa para las poblaciones actuales de la Patagonia también se encuentran asociados al cromosoma 6, con una región enriquecida para el componente Nativo Americano relacionado con HLA. Los

principales procesos biológicos y rutas metabólicas asociadas a la región HLA se encuentran fuertemente asociadas con presentación de antígenos, asma, enfermedades autoinmunes, virales y bacterianas.

Las enfermedades infecciosas son una de las causas más importantes de mortalidad en las poblaciones humanas. Las adaptaciones genéticas humanas están controladas en su mayoría por la presencia de virus, bacterias o parásitos (Fumagalli et al., 2011). Dicho esto, la elevada diversidad en esta región del cromosoma 6 permite responder fuertes eventos de selección en tiempos relativamente cortos e independientes de la ancestría específica de alguna población. Sin afectar lo anterior, también se han encontrado otros genes en distintas cromosomas que han permitido ejercer una respuesta adaptativa inmune en diferentes poblaciones Nativas de los Andes, México y el sureste de Estados Unidos (Lindo et al., 2018; Reynolds et al., 2019).

El escenario que aparentemente está ocurriendo en la Patagonia puede estar dado por una fuerte presión selectiva debido a la carga patogénica relacionada con los Europeos. Por ende, distintas variantes asociadas a la región HLA en los grupos cazadores recolectores de la Patagonia pudieron haber conferido una selección positiva frente a diversas infecciones como la viruela, sarampión y otras enfermedades infecciosas durante los últimos 300 años. Esta respuesta adaptativa inmune provocó que los individuos que vemos en la actualidad descenderían de antepasados que participaron en el proceso de mestizaje y que además, sobrevivieron a estas epidemias. De acuerdo a los registros, las principales causas de muertes en la población Kawéskar eran las enfermedades venéreas, principalmente la Sífilis. Sin embargo, también causaron estragos la Tuberculosis y el Sarampión (Gleisner & Sara Montt, 2014).

Para terminar, es importante mencionar que uno de los genes involucrados en la respuesta inmune en diversas poblaciones humanas corresponde a **ADAM17**. Este gen está involucrado en la entrada viral y participa en la inflamación intestinal gatillada por receptores Toll-like (Fumagalli et al., 2011). Esto apuntaría además, a

la importancia de este gen en las poblaciones de la Patagonia, puesto que no solo estaría participando en un rol asociado a respuesta a frío o dieta, sino que también pudo haber conferido una respuesta adaptativa inmune frente a los patógenos traídos por los Europeos.

6.4. Historia de los Procesos de Adaptación Biológica en las Poblaciones de la Patagonia: Señales pre y post-contacto Europeo.

En base a lo mencionado anteriormente, resulta apropiado poder resumir los principales eventos de adaptación biológica frente a frío y dieta marina que ocurrieron en las poblaciones descendientes de cazadores recolectores de la Patagonia (Figura 75). A pesar de que estas poblaciones poseen regiones bajo selección cercanas a los genes **FADS**, **CPT1A** y **LRP5**, estas corresponderían a huellas de adaptación en el pasado, ahora bien, es necesario poder evaluar si realmente estas regiones están jugando un rol de adaptación biológica en estas poblaciones. En lo que respecta a los genes candidatos encontrados en este estudio, podemos indicar procesos de adaptación en los primeros grupos que colonizaron la Patagonia, potencialmente asociados a los genes **ASLX2**, **LRP2**, **ADAM17** y **ZNF202**, los cuales estarían confiando mecanismos de respuesta relacionado con el metabolismo de lípidos, regulación en los niveles de colesterol en la sangre y homeostásis del equilibrio energético.

El proceso de diferenciación entre cazadores recolectores marítimos y terrestres de la Patagonia generó una alteración en el sistema de dieta, principalmente para el primer grupo, antes de diferenciarse hace más de 7.000 años atrás (Orquera et al., 2011; Orquera & Piana, 2009). El aumento en una dieta rica en productos del mar, los que incluían frecuentemente carnes de lobo de mar, peces y mariscos, y en menor medida carne de ballena, de huemul, aves y huevos (Gleisner & Sara Montt, 2014) pudo haber provocado una alteración en el metabolismo frente al consumo de alimentos y la elevada cantidad de ácidos grasos. Las principales señales bajo selección presente en estos grupos estarían ligados al gen **GFRAL**, el cual pudo haber

participado en este proceso de adaptación biológica durante el cambio en el sistema de dietas. Otras señales de selección que se pudieron detectar, esta vez de manera específica en grupos Kawéskar y Yámana de la Patagonia Austral fueron los genes **TRPM8** asociados a mecanismos de respuesta a frío. El gen **UCN3** estaría participando en la regulación del sistema cardiovascular en respuesta a hipoxia, y por lo tanto relacionado con el buceo, como práctica de recolección de productos marítimos por parte de estos grupos. No obstante, una de las señales bajo selección más robusta correspondería al gen **ZNF202**, puesto que se mantendría durante la diferenciación de los cazadores recolectores marítimos de la Patagonia.

Estas condiciones de adaptación, tanto cultural como biológica se vieron interrumpidas por el contacto con los Europeos mediante las embarcaciones que cruzaban por los canales de la Patagonia, y posteriormente por la inclusión de los Españoles durante el proceso de Conquista hace aproximadamente 300 años atrás. Esto trajo consigo que la conformación de las nuevas poblaciones de la Patagonia, esta vez mestizas, tuvieran en sus genomas una región bajo selección de ancestría Nativa asociada principalmente a la región HLA, y en respuesta a las enfermedades traídas por los Europeos.

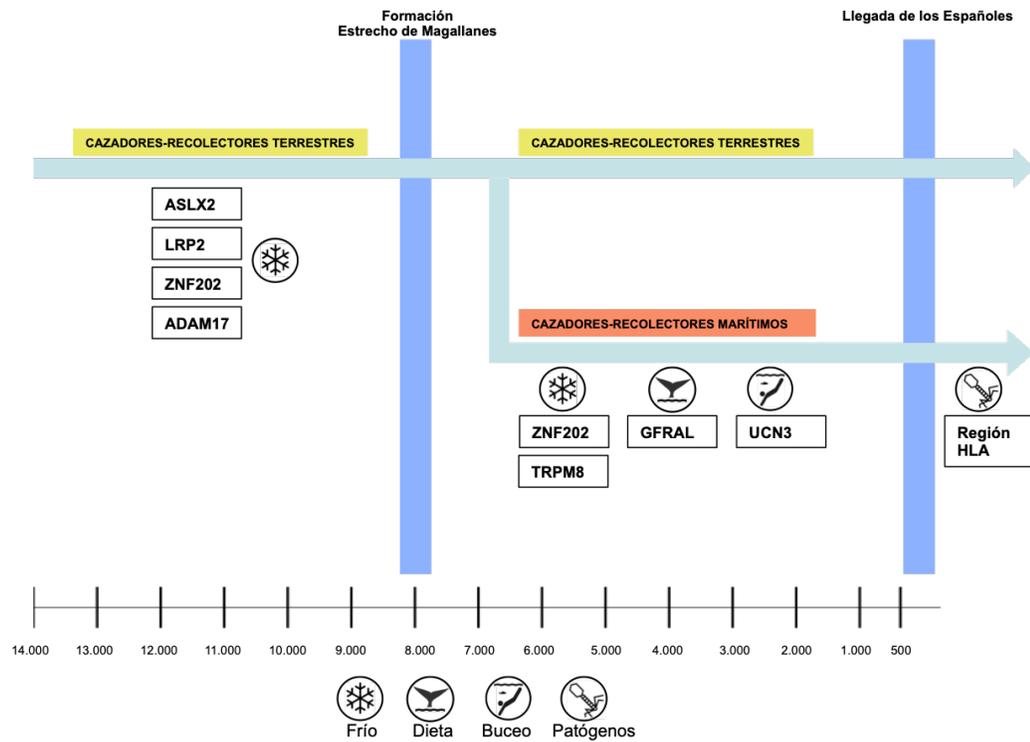


Figura 75. Esquema con la historia de los procesos evolutivos por parte de las poblaciones de la Patagonia. El eje inferior indica el tiempo en años, desde la izquierda (14.000 años atrás) el más antiguo y la derecha (500 años atrás) siendo lo más reciente. Las barras verticales azules representan los hitos más importantes para los eventos de selección. Las flechas horizontales representa la separación entre los cazadores recolectores terrestres y marítimos de la Patagonia. Los recuadros corresponden a los genes candidatos bajo selección asociados a frío, dieta, buceo o patógenos.

6.5. Consideraciones Finales

De acuerdo a lo descrito anteriormente, la mayoría de los genes que se encuentran bajo selección y procesos biológicos descritos se encuentran relacionados ya sea con el metabolismo de lípidos y/o regulación en los niveles de colesterol HLD o LDL en los individuos de la Patagonia. Estos mecanismos de respuesta adaptativa a las condiciones ambientales de la Patagonia (frío y dieta marítima) parecieran interactuar entre sí y ser mucho más complejas. Por ejemplo, la dieta y el frío son reguladores de la termogénesis adaptativa mediante la activación del sistema nervioso simpático (Van Marken Lichtenbelt & Daanen, 2003). Para la dieta, el consumo de ácidos grasos en condiciones ambientales extremas son de vital importancia dado que corresponden a uno de los substratos –aparte de la

glucosa- para la producción de calor y para la fluidez de la membrana celular en respuesta a cambios de temperatura (Ma et al., 2015; Weber & Haman, 2005). Hoy en día, existe evidencia que el aceite proveniente de pescados, rico en n-3 PUFAs, ayuda a la estimulación de la termogénesis en el tejido adiposo pardo (BAT) e incluso en el tejido adiposo blanco (WAT) (Okla et al., 2017).

Estas dietas elevadas en ácidos grasos previenen la ganancia de peso inducido por dieta y la adiposidad, sin embargo, variaciones en la termogénesis en poblaciones actuales puede desencadenar directamente en riesgo de obesidad y diabetes tipo 2 (Kaisanlahti, 2019). Asociado a esto, la mayoría de los genes que están bajo selección en nuestros datos están relacionadas con obesidad, problemas cardiovasculares, dislipidemia, resistencia a la insulina, diabetes tipo 2, etc. Esto es un indicador importante para las poblaciones actuales no solo de la Patagonia, lo que presumiblemente en el pasado significaba una optimización en la ingesta, almacenamiento y metabolización de lípidos productos de la dieta, ahora se traduce en problemas de salud pública. Este problema lo vemos reflejado en poblaciones actuales de México para genes como **ABCA1** (Acuna-Alonzo et al., 2010) y no solo podría tratarse de adaptaciones locales, sino que de eventos adaptativos asociados a genes como **FADS**, **CPT1A** o **LRP5** en las poblaciones fundadoras de América, y que hoy en día aun se encuentran estas señales en las poblaciones mestizas de América.

No obstante, es importante mencionar aquellos alcances necesarios para poder ratificar de manera optima los resultados encontrados en este trabajo. Por un lado, y en relación a lo mencionado anteriormente, se requiere la inclusión de genomas de individuos de la Patagonia, así como también muestras de ADN antiguo con tal de evaluar las señales de selección encontrados en este estudio. Por otro lado, es importante poder evaluar si la señales de selección que vemos en la Patagonia se encuentran presentes en aquellas poblaciones que utilizamos como referencia, tales como los Mapuches y los Aymara. A pesar que esto lo podemos evaluar mediante el análisis de PBS (Population Branch Statistic) enfocado en los grupos de la Patagonia, se requiere realizar lo mismo tomando como foco a estos grupos de referencia.

La inclusión de una nueva base de datos asociados con la Patagonia puede ayudar a aunar los datos de microarreglos a través de la imputación de estos, y así evaluar los procesos de adaptación biológica en estas poblaciones. Además, este estudio realiza una nueva aproximación de enmascaramiento del genoma No Nativo, por lo que no sólo podríamos analizar individuos con una elevada ancestría Nativa, sino que también individuos mestizos, y por lo tanto recuperar información valiosa de las poblaciones originarias incorporada originalmente en las poblaciones actuales.

7. Conclusiones

De acuerdo a los análisis de ancestría global, las contribuciones mas importantes en las poblaciones descendientes de cazadores recolectores de la Patagonia fueron entre las ancestrías Nativo Americanas y Europea. El componente nativo fue similar entre los Patagónicos Chilenos y Argentinos, sin embargo no pudo ser caracterizado debido a la falta de una población ancestral a la Patagonia en nuestros datos de referencia. El análisis a partir de los datos de MEGA, con más SNPs, indicó la presencia un nuevo componente ancestral único en grupos Yámana y Kawéskar, desapareciendo el aparente componente Andino. Distintas hipótesis pueden explicar el origen de las poblaciones patagónicas: 1) un grupo ancestral que pobló la Patagonia mediante una ruta costera del Pacífico; 2) una población compuesta por una mezcla entre grupos Mapuches y en menor grado por individuos de origen Andino; 3) el componente Andino de los genomas amerindios patagónicos es el producto de una mezcla posterior al proceso de poblamiento de la Patagonia. Si bien el conjunto de nuestros resultados dan mayor soporte a la primera hipótesis, se requerirán otros estudios para poder contrastar estas hipótesis.

Se encontraron distintas señales de selección para las poblaciones de la Patagonia involucrados directamente con termogénesis en respuesta a frío e interacción con el elevado consumo de ácidos grasos. Asimismo, se encontraron procesos biológicos asociados con la regulación en la termogénesis, metabolismo de ácidos grasos, regulación de ingesta de comida, entre otros.

Los procesos de adaptación biológica frente a la respuesta a frío en los grupos cazadores recolectores de la Patagonia están dados principalmente por los genes **ASXL2**, **LRP2**, **ADAM17** y **ZNF202**, los cuales se han descrito como partícipes del metabolismo de lípidos y colesterol.

Posteriormente, los grupos cazadores recolectores marítimos de la Patagonia presentaron por su parte señales de selección para los genes **ZNF202**, **GFRAL** y **UCN3**, relacionados con adaptación biológica a frío, dieta marítima y buceo.

Se identificó un SNP bajo selección en las poblaciones de la Patagonia Austral, la cual está dentro del gen **TRPM8**, asociado con la regulación de la temperatura y se ha descrito como principal causante de respuesta adaptativa a ambientes fríos en múltiples poblaciones a lo largo del mundo.

Fue posible detectar señales de selección en la región del cromosoma 11, conteniendo a los genes **FADS**, **LRP5** y **CPT1A** al igual que en poblaciones nativas de Siberia y Groenlandia. Sin embargo, sus señales no son lo suficientemente fuertes y están mas bien dando cuenta de una huella de selección en el pasado que ocurrió en Beringia por parte de grupos amerindios ancestrales.

Las poblaciones actuales presentaron un exceso de enriquecimiento de origen Nativo, lo que indica una señal de selección en la región **HLA** del cromosoma 6 asociado a una respuesta inmune tras la llegada de los Europeos a la Patagonia.

8. Bibliografía

- Achilli, A., Perego, U. a, Bravi, C. M., Coble, M. D., Kong, Q.-P., Woodward, S. R., ... Bandelt, H.-J. (2008). The phylogeny of the four pan-American MtDNA haplogroups: implications for evolutionary and disease studies. *PLoS One*, 3(3), e1764. <https://doi.org/10.1371/journal.pone.0001764>
- Achilli, A., Perego, U. a, Lancioni, H., Olivieri, A., Gandini, F., Hooshiar Kashani, B., ... Torroni, A. (2013). Reconciling migration models to the Americas with the variation of North American native mitogenomes. *Proceedings of the National Academy of Sciences of the United States of America*, 110(35), 14308–14313. <https://doi.org/10.1073/pnas.1306290110>
- Acuna-Alonzo, V., Flores-Dorantes, T., Kruit, J. K., Villarreal-Molina, T., Arellano-Campos, O., Hunemeier, T., ... Canizales-Quinteros, S. (2010). A functional ABCA1 gene variant is associated with low HDL-cholesterol levels and shows evidence of positive selection in Native Americans. *Human Molecular Genetics*, 19(14), 2877–2885. <https://doi.org/10.1093/hmg/ddq173>
- Alexander, D. H., & Lange, K. (2011). Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinformatics*, 12(1), 246. <https://doi.org/10.1186/1471-2105-12-246>
- Báez, O. R., Moraga, M., Méndez, C., & Cherkinsky, A. (2015). Maritime Hunter-Gatherers in the Chonos Archipelago (43°50'–46°50' S), Western Patagonian Channels. *The Journal of Island and Coastal Archaeology*, (March), 1–25. <https://doi.org/10.1080/15564894.2014.1001920>
- Bank, C., Ewing, G. B., Ferrer-admettla, A., Foll, M., & Jensen, J. D. (2014). Thinking too positive ? Revisiting current methods of population genetic selection inference. *Trends in Genetics*, 30(12), 540–546. <https://doi.org/10.1016/j.tig.2014.09.010>

- Barros Valenzuela, A. (1975a). Capítulo 2. Yámana: los Primitivos más australes de la tierra. In *Australes Aborígenes de América* (Lord Cochr, p. 63 pp.). Santiago, Chile.
- Barros Valenzuela, A. (1975b). Capítulo 5. En el país de Ayayema. In *Aborígenes Australes de América* (Lord Cochr, p. 63 pp.). Santiago, Chile.
- Barros Valenzuela, A. (1975c). *Capítulo 8. Chonos. Aborígenes Australes de América* (Lord Cochr). Santiago, Chile.
- Bautista, D. M., Siemens, J., Glazer, J. M., Tsuruda, P. R., Basbaum, A. I., Stucky, C. L., ... Julius, D. (2007). The menthol receptor TRPM8 is the principal detector of environmental cold. *Nature*, *448*(7150), 204–208.
<https://doi.org/10.1038/nature05910>
- Bene, J., Hadzsiev, K., & Melegh, B. (2018). Role of carnitine and its derivatives in the development and management of type 2 diabetes. *Nutrition Diabetes*, *7*(8), 8.
- Berón, M. A., Luna, L. H., & Barberena, R. (2009). Isotopic Archaeology in the Western Pampas (Argentina): Preliminary Results and Perspectives. *International Journal of Osteoarchaeology*, *19*, 250–265. <https://doi.org/10.1002/oa>
- Bhatia, G., Tandon, A., Patterson, N., Aldrich, M. C., Ambrosone, C. B., Amos, C., ... Price, A. L. (2014). Genome-wide scan of 29,141 African Americans finds no evidence of directional selection since admixture. *American Journal of Human Genetics*, *95*(4), 437–444. <https://doi.org/10.1016/j.ajhg.2014.08.011>
- Bhatia, G., Tandon, A., Patterson, N., Aldrich, M. C., Ambrosone, C. B., Amos, C., ... Reich, D. (2014). Genome-wide Scan of 29 , 141 African Americans Finds No Evidence of Directional Selection since Admixture. *The American Journal of Human Genetics*, *95*(4), 437–444. <https://doi.org/10.1016/j.ajhg.2014.08.011>
- Bigham, A., Bauchet, M., Pinto, D., Mao, X., Akey, J. M., Mei, R., ... Shriver, M. D. (2010).

Identifying signatures of natural selection in Tibetan and Andean populations using dense genome scan data. *PLoS Genetics*, 6(9).

<https://doi.org/10.1371/journal.pgen.1001116>

Bigham, A. W., Mao, X., Mei, R., Brutsaert, T., Wilson, M. J., Julian, C. G., ... Shriver, M. D. (2009). Identifying positive selection candidate loci for high-altitude adaptation in Andean populations. *Human Genomics*, 4(2), 79–90.

<https://doi.org/6010W7357574W752> [pii]

Blanquart, S., Borowiec, A. sophie, Delcourt, P., Figeac, M., Emerling, C. A., Meseguer, A. S., ... Bidaux, G. (2019). Evolution of the human cold/menthol receptor, TRPM8. *Molecular Phylogenetics and Evolution*, 136(January), 104–118.

<https://doi.org/10.1016/j.ympev.2019.04.011>

Bonhomme, M., Chevalet, C., Servin, B., Boitard, S., Abdallah, J., Blott, S., & SanCristobal, M. (2010). Detecting Selection in Population Trees: The Lewontin and Krakauer Test Extended. *Genetics*, 186(1), 241–262.

<https://doi.org/10.1534/genetics.110.117275>

Borrero, L. A., Barberena, R., Franco, N. V., Charlin, J., & Tykot, R. (2009). Isotopes and Rocks: Geographical Organisation of Southern Patagonian Hunter-Gatherers, 19, 309–327. <https://doi.org/10.1002/oa>

Botigué, L. R., Henn, B. M., Gravel, S., Maples, B. K., Gignoux, C. R., & Corona, E. (2013). Gene flow from North Africa contributes to differential human genetic diversity in southern Europe, 110(29), 11791–11796.

<https://doi.org/10.1073/pnas.1306223110>

Breit, S. N., Tsai, V. W. W., & Brown, D. A. (2017). Targeting Obesity and Cachexia: Identification of the GFRAL Receptor–MIC-1/GDF15 Pathway. *Trends in Molecular Medicine*, 23(12), 1065–1067. <https://doi.org/10.1016/j.molmed.2017.10.005>

- Brychta, R., & Chen, K. (2017). Cold-induced thermogenesis in humans. *European Journal of Clinical Nutrition*, 71(3), 345–352.
- Cardona, A., Pagani, L., Antao, T., Lawson, D. J., Eichstaedt, C. a., Yngvadottir, B., ... Kivisild, T. (2014a). Genome-wide analysis of cold adaptation in indigenous Siberian populations. *PLoS ONE*, 9(5).
<https://doi.org/10.1371/journal.pone.0098076>
- Cardona, A., Pagani, L., Antao, T., Lawson, D. J., Eichstaedt, C. A., Yngvadottir, B., ... Kivisild, T. (2014b). Genome-wide analysis of cold adaptation in indigenous Siberian populations. *PLoS ONE*, 9(5).
<https://doi.org/10.1371/journal.pone.0098076>
- Cavalli-Sforza, L. L. (1969). Human Diversity. In *Proc 12th Intl Cong Genet* (pp. 405–416).
- Clemente, F. J., Cardona, A., Inchley, C. E., Peter, B. M., Jacobs, G., Mitt, M., ... Ma, R. (2014). REPORT A Selective Sweep on a Deleterious Mutation in CPT1A in Arctic Populations, 584–589. <https://doi.org/10.1016/j.ajhg.2014.09.016>
- Clemente, F. J., Cardona, A., Inchley, C. E., Peter, B. M., Jacobs, G., Pagani, L., ... Kivisild, T. (2014). A Selective Sweep on a Deleterious Mutation in CPT1A in Arctic Populations. *The American Journal of Human Genetics*, 95(5), 584–589.
<https://doi.org/10.1016/j.ajhg.2014.09.016>
- Crisci, J. L., Poh, Y.-P., Bean, A., Simkin, A., & Jensen, J. D. (2012). Recent Progress in Polymorphism-Based Population Genetic Inference. *Journal of Heredity*, 103(2), 287–296. <https://doi.org/10.1093/jhered/esr128>
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., ... Durbin, R. (2011a). The variant call format and VCFtools. *Bioinformatics*, 27(15), 2156–2158. <https://doi.org/10.1093/bioinformatics/btr330>

- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., ... Durbin, R. (2011b). The Variant Call Format and VCFTools. *Bioinformatics*.
- De, C., Ávila-arcos, M. C., Galimany, J., Carpenter, M. L., & Homburger, J. R. (2018). Genomic insights into the origin and diversification of late maritime hunter-gatherers from the Chilean Patagonia. <https://doi.org/10.1073/pnas.1715688115>
- de Saint Pierre, M., Bravi, C. M., Motti, J. M. B., Fuku, N., Tanaka, M., Llop, E., ... Moraga, M. (2012). An Alternative Model for the Early Peopling of Southern South America Revealed by Analyses of Three Mitochondrial DNA Haplogroups. *PLoS ONE*, 7(9), e43486. <https://doi.org/10.1371/journal.pone.0043486>
- de Saint Pierre, M., Gandini, F., Perego, U. A., Bodner, M., Gómez-Carballa, A., Corach, D., ... Olivieri, A. (2012). Arrival of Paleo-Indians to the Southern Cone of South America: New Clues from Mitogenomes. *PLoS ONE*, 7(12), 1–9. <https://doi.org/10.1371/journal.pone.0051311>
- Delaneau, O., Marchini, J., & Zagury, J.-F. (2012). A linear complexity phasing method for thousands of genomes. *Nature Methods*, 9(2), 179–181. <https://doi.org/10.1038/nmeth.1785>
- Deng, L., Ruiz-Linares, A., Xu, S., & Wang, S. (2016). Ancestry variation and footprints of natural selection along the genome in Latin American populations. *Scientific Reports*, 6, 1–7. <https://doi.org/10.1038/srep21766>
- Dillehay, T. D. (1997). *Monte Verde: a late Pleistocene settlement in Chile*. Smithsonian Institution Press.
- Dillehay, T. D. (2009). Probing deeper into first American studies. *Proceedings of the National Academy of Sciences of the United States of America*, 106(4), 971–978. <https://doi.org/10.1073/pnas.0808424106>

- Dillehay, T. D., Ramírez, C., Pino, M., Collins, M. B., Rossen, J., & Pino-Navarro, J. D. (2008). Monte Verde: Seaweed, Food, Medicine, and the Peopling of South America, *320*(May), 784–786.
- Dulik, M. C., Zhadanov, S. I., Osipova, L. P., Askapuli, A., Gau, L., Gokcumen, O., ... Schurr, T. G. (2012). Mitochondrial DNA and y chromosome variation provides evidence for a recent common ancestry between Native Americans and indigenous Altaians. *American Journal of Human Genetics*, *90*(2), 229–246.
<https://doi.org/10.1016/j.ajhg.2011.12.014>
- Eduardo, C., Amorim, G., Nunes, K., Meyer, D., Comas, D., Cátira, M., ... Hünemeier, T. (2017). Genetic signature of natural selection in first Americans, (4).
<https://doi.org/10.1073/pnas.1620541114>
- Eichstaedt, C. A., Antao, T., Cardona, A., Pagani, L., Kivisild, T., & Mormina, M. (2015). Mutation Research / Fundamental and Molecular Mechanisms of Mutagenesis Positive selection of AS3MT to arsenic water in Andean populations. *Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis*, *780*, 97–102.
<https://doi.org/10.1016/j.mrfmmm.2015.07.007>
- Eichstaedt, C. A., Antão, T., Cardona, A., Pagani, L., Kivisild, T., & Mormina, M. (2015). Genetic and phenotypic differentiation of an andean intermediate altitude population. *Physiological Reports*, *3*(5), 1–15.
<https://doi.org/10.14814/phy2.12376>
- Excoffier, L., Hofer, T., & Foll, M. (2009). Detecting loci under selection in a hierarchically structured population. *Heredity*, *103*(4), 285–298.
<https://doi.org/10.1038/hdy.2009.74>
- Eyheramendy, S., Martinez, F. I., Manevy, F., Vial, C., & Repetto, G. M. (2015). Genetic structure characterization of Chileans reflects historical immigration patterns.

Nature Communications, 6, 1–10. <https://doi.org/10.1038/ncomms7472>

Fagundes, N. J. R., Kanitz, R., & Bonatto, S. L. (2008). A reevaluation of the Native American mtDNA genome diversity and its bearing on the models of early colonization of Beringia. *PLoS One*, 3(9), e3157.

<https://doi.org/10.1371/journal.pone.0003157>

Fagundes, N. J. R., Kanitz, R., Eckert, R., Valls, A. C. S., Bogó, M. R., Salzano, F. M., ... Bonatto, S. L. (2008). Mitochondrial Population Genomics Supports a Single Pre-Clovis Origin with a Coastal Route for the Peopling of the Americas, (March), 583–592. <https://doi.org/10.1016/j.ajhg.2007.11.013>.

Fumagalli, M., Bjerregaard, P., Jørgensen, M. E., & Korneliussen, T. S. (2015). Research reports.

Fumagalli, M., & Sironi, M. (2014). Human genome variability, natural selection and infectious diseases. *Current Opinion in Immunology*, 30(1), 9–16.

<https://doi.org/10.1016/j.coi.2014.05.001>

Fumagalli, M., Sironi, M., Pozzoli, U., Ferrer-Admettla, A., Pattini, L., & Nielsen, R. (2011). Signatures of Environmental Genetic Adaptation Pinpoint Pathogens as the Main Selective Pressure through Human Evolution. *PLoS Genetics*, 7(11), e1002355. <https://doi.org/10.1371/journal.pgen.1002355>

Gaete, N., Navarro, X., Constantinescu, F., Mera, C., Selles, D., Solari, M. E., ... Durán, L. (2004). Una mirada al modo de vida canoero del mar interior desde Piedra Azul. *Chungará (Arica)*, 36, 333–346.

Garreaud, R., Lopez, P., Minvielle, M., & Rojas, M. (2013). Large-scale control on the Patagonian climate. *Journal of Climate*, 26(1), 215–230.

<https://doi.org/10.1175/JCLI-D-12-00001.1>

- Gleisner, C., & Sara Montt. (2014). *Kawésqar*. (Sara Montt & C. Gleisner, Eds.) (Unidad de). Santiago, Chile: Ograma.
- Goebel, T., Waters, M. R., & O'Rourke, D. H. (2008). The late Pleistocene dispersal of modern humans in the Americas. *Science (New York, N.Y.)*, 319(5869), 1497–1502. <https://doi.org/10.1126/science.1153569>
- González-José, R. (2003). *El Poblamiento de la Patagonia. Análisis de la Variación Craneofacial en el Contexto del Poblamiento Americano*. Universidad de Barcelona.
- Gravel, S. (2012). Population Genetics Models of Local Ancestry, 191(June), 607–619. <https://doi.org/10.1534/genetics.112.139808>
- Gravel, S., Zakharia, F., Mccauley, J. L., Byrnes, J. K., Hedges, D. J., Gignoux, C. R., ... Bustamante, C. D. (2013). Reconstructing the Population Genetic History of the Caribbean, 9(11). <https://doi.org/10.1371/journal.pgen.1003925>
- Gravel, S., Zakharia, F., Moreno-Estrada, A., Byrnes, J. K., Muzzio, M., Rodriguez-Flores, J. L., ... Bustamante, C. D. (2013). Reconstructing Native American Migrations from Whole-Genome and Whole-Exome Data. *PLoS Genetics*, 9(12). <https://doi.org/10.1371/journal.pgen.1004023>
- Gutenkunst, R. N., Hernandez, R. D., Williamson, S. H., & Bustamante, C. D. (2009). Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genetics*, 5(10). <https://doi.org/10.1371/journal.pgen.1000695>
- Hancock, a. M., Alkorta-Aranburu, G., Witonsky, D. B., & Di Rienzo, a. (2010). Adaptations to new environments in humans: the role of subtle allele frequency shifts. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1552), 2459–2468. <https://doi.org/10.1098/rstb.2010.0032>

- Hancock, A. M., Alkorta-aranburu, G., Witonsky, D. B., & Rienzo, A. Di. (2010). Adaptations to new environments in humans : the role of subtle allele frequency shifts, 2459–2468. <https://doi.org/10.1098/rstb.2010.0032>
- Hancock, A. M., Witonsky, D. B., Alkorta-Aranburu, G., Beall, C. M., Gebremedhin, A., Sukernik, R., ... Di Rienzo, A. (2011). Adaptations to climate-mediated selective pressures in humans. *PLoS Genetics*, *7*(4). <https://doi.org/10.1371/journal.pgen.1001375>
- Hellenthal, G., Busby, G. B. J., Band, G., Wilson, J. F., Capelli, C., Falush, D., & Myers, S. (2014). A genetic atlas of human admixture history. *Science*, *343*(6172), 747–751. <https://doi.org/10.1126/science.1243518>
- Hermisson, J., & Pennings, P. S. (2005). Soft sweeps: Molecular population genetics of adaptation from standing genetic variation. *Genetics*, *169*(4), 2335–2352. <https://doi.org/10.1534/genetics.104.036947>
- Hofer, T., Ray, N., Wegmann, D., & Excoffier, L. (2009). Large allele frequency differences between human continental groups are more likely to have occurred by drift during range expansions than by selection. *Annals of Human Genetics*, *73*(1), 95–108. <https://doi.org/10.1111/j.1469-1809.2008.00489.x>
- Holsinger, K. E., & Weir, B. S. (2009). Genetics in geographically structured populations: defining, estimating and interpreting F_{ST} . *Nat Rev Genet*, *10*(9), 639–650. Retrieved from <http://dx.doi.org/10.1038/nrg2611>
- Homburger, J. R., Moreno-estrada, A., Gignoux, C. R., Nelson, D., Sanchez, E., Ortiz-tello, P., ... Gravel, S. (2015). Genomic Insights into the Ancestry and Demographic History of South America, 1–26. <https://doi.org/10.1371/journal.pgen.1005602>
- Hoppel, C. (2003). The role of carnitine in normal and altered fatty acid metabolism. *American Journal of Kidney Disorders*, *41*(4), S4–S12.

- Hsu, F., Kent, J. W., Clawson, H., Kuhn, R. M., Diekhans, M., & Haussler, D. (2006). The UCSC known genes. *Bioinformatics*, 22(9), 1036–1046.
<https://doi.org/10.1093/bioinformatics/btl048>
- Hsu, J. Y., Crawley, S., Chen, M., Ayupova, D. A., Lindhout, D. A., Higbee, J., ... Allan, B. B. (2017). Non-homeostatic body weight regulation through a brainstem-restricted receptor for GDF15. *Nature*, 550(7675), 255–259.
<https://doi.org/10.1038/nature24042>
- Ilardo, M. A., Moltke, I., Korneliussen, T. S., Ilardo, M. A., Moltke, I., Korneliussen, T. S., ... Rasmussen, S. (2018). Sea Nomads Article Physiological and Genetic Adaptations to Diving in Sea Nomads. *Cell*, 173(3), 569-573.e15.
<https://doi.org/10.1016/j.cell.2018.03.054>
- Ilardo, M., & Nielsen, R. (2018). ScienceDirect Human adaptation to extreme environmental conditions. *Current Opinion in Genetics & Development*, 53, 77–82.
<https://doi.org/10.1016/j.gde.2018.07.003>
- Izawa, T., Rohatgi, N., Fukunaga, T., Wang, Q. T., Silva, M. J., Gardner, M. J., ... Zou, W. (2015). ASXL2 Regulates Glucose, Lipid, and Skeletal Homeostasis. *Cell Reports*, 11(10), 1625–1637. <https://doi.org/10.1016/j.celrep.2015.05.019>
- Jacovas, Vanessa C, Couto-silva, C. M., Nunes, K., Lemes, R. B., Oliveira, M. Z. De, Salzano, F. M., ... Hünemeier, T. (2018). Selection scan reveals three new loci related to high altitude adaptation in Native Andeans, 1–8.
<https://doi.org/10.1038/s41598-018-31100-6>
- Jacovas, Vanessa Cristina, Rovaris, D. L., Pérez, O., & Azevedo, S. De. (2015). Genetic Variations in the TP53 Pathway in Native Americans Strongly Suggest Adaptation to the High Altitudes of the Andes, 1–15.
<https://doi.org/10.1371/journal.pone.0137823>

- Jin, W., Xu, S., Wang, H., Yu, Y., Shen, Y., Wu, B., & Jin, L. (2011). Genome-wide detection of natural selection in African Americans pre-and post-admixture. *Genome Research*, 519–527. <https://doi.org/10.1101/gr.124784.111>
- Junyent, M., Parnell, L. D., Lai, C. Q., Arnett, D. K., Tsai, M. Y., Kabagambe, E. K., ... Ordovás, J. M. (2010). ADAM17_i33708A>G polymorphism interacts with dietary n-6 polyunsaturated fatty acids to modulate obesity risk in the Genetics of Lipid Lowering Drugs and Diet Network study. *Nutrition, Metabolism and Cardiovascular Diseases*, 20(10), 698–705. <https://doi.org/10.1016/j.numecd.2009.06.011>
- Kaisanlahti, A. (2019). Browning of white fat : agents and implications for beige adipose tissue to type 2 diabetes, 1–10.
- Key, F. M., Abdul-Aziz, M. A., Mundry, R., Peter, B. M., Sekar, A., D’Amato, M., ... Andrés, A. M. (2018). Human local adaptation of the TRPM8 cold receptor along a latitudinal cline. *PLoS Genetics*, 14(5), 1–22. <https://doi.org/10.1371/journal.pgen.1007298>
- Kilian, R., & Lamy, F. (2012). A review of Glacial and Holocene paleoclimate records from southernmost Patagonia (49-55°S). *Quaternary Science Reviews*, 53(C), 1–23. <https://doi.org/10.1016/j.quascirev.2012.07.017>
- Kimura, M. (1955). SOLUTION OF A PROCESS OF RANDOM GENETIC DRIFT WITH A CONTINUOUS MODEL. *Proceedings of the National Academy of Sciences of the United States of America*, 41(3), 144–150. Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC528040/>
- Kloet, E. R. De, Joëls, M., & Holsboer, F. (2005). STRESS AND THE BRAIN : FROM ADAPTATION TO DISEASE, 6(June), 463–475. <https://doi.org/10.1038/nrn1683>
- Lachance, J., & Tishkoff, S. A. (2013). Population Genomics of Human Adaptation. *Annual Review of Ecology, Evolution, and Systematics*, 44(1), 123–143.

<https://doi.org/DOI 10.1146/annurev-ecolsys-110512-135833>

Lawson, D. J., Hellenthal, G., Myers, S., & Falush, D. (2012). Inference of population structure using dense haplotype data. *PLoS Genetics*, *8*(1), 11–17.

<https://doi.org/10.1371/journal.pgen.1002453>

Legoupil, D., & Fontugne, M. (1997). El Poblamiento Marítimo en los Archipiélagos de Patagonia: Núcleos Antiguos y Dispersión Reciente. *ANS. INST. PAT., Ser. Cs. Hs.*, *25*, 75–87.

Legoupil, Dominique. (2005). Recolectores De Moluscos Tempranos En El Sureste De La Isla De Chiloé: Una Primera Mirada. *Magallania (Punta Arenas)*, *33*(1), 51–61.

<https://doi.org/10.4067/S0718-22442005000100004>

Leonard, W. R., Snodgrass, J. J., & Sorensen, M. V. (2005a). Metabolic Adaptation in Indigenous Siberian Populations. *Annual Review of Anthropology*, *34*(1), 451–471.

<https://doi.org/10.1146/annurev.anthro.34.081804.120558>

Leonard, W. R., Snodgrass, J. J., & Sorensen, M. V. (2005b). Metabolic Adaptation in Indigenous Siberian Populations. *Annual Review of Anthropology*, *34*(1), 451–471.

<https://doi.org/10.1146/annurev.anthro.34.081804.120558>

Lewontin, R. C., & Krakauer, J. (1973). Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics*, *74*(1), 175–195.

Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/4711903>

Liao, Y., Wang, J., Jaehnig, E., Shi, Z., & Zhang, B. (2019). WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. *Nucleic Acids Research*.

Lindo, J., Haas, R., Hofman, C., Apata, M., Moraga, M., Verdugo, R. A., ... Rienzo, A. Di. (2018). The genetic prehistory of the Andean highlands 7000 years BP through European contact.

- Liu, Q., Zhang, J., Tran, H., Verbeek, M. M., Reiss, K., Estus, S., & Bu, G. (2009). LRP1 shedding in human brain : roles of ADAM10 and ADAM17, 7, 1–7.
<https://doi.org/10.1186/1750-1326-4-17>
- Ma, D. K., Li, Z., Lu, A. Y., Sun, F., Chen, S., Rothe, M., ... Horvitz, H. R. (2015). Acyl-CoA dehydrogenase drives heat adaptation by sequestering fatty acids. *Cell*, 161(5), 1152–1163. <https://doi.org/10.1016/j.cell.2015.04.026>
- Maples, B. K., Gravel, S., Kenny, E. E., & Bustamante, C. D. (2013a). RFMix : A Discriminative Modeling Approach for Rapid and Robust Local-Ancestry Inference. *The American Journal of Human Genetics*, 93(2), 278–288.
<https://doi.org/10.1016/j.ajhg.2013.06.020>
- Maples, B. K., Gravel, S., Kenny, E. E., & Bustamante, C. D. (2013b). RFMix: A discriminative modeling approach for rapid and robust local-ancestry inference. *American Journal of Human Genetics*, 93(2), 278–288.
<https://doi.org/10.1016/j.ajhg.2013.06.020>
- Massone, M., & Gusinde, M. (1987). *Las Culturas Aborígenes de Chile Austral en el Tiempo* (Museo Chil). Santiago, Chile.
- Matos-Cruz, V., Schneider, E. R., Mastrotto, M., Merriman, D. K., Bagriantsev, S. N., & Gracheva, E. O. (2017). Molecular Prerequisites for Diminished Cold Sensitivity in Ground Squirrels and Hamsters. *Cell Reports*, 21(12), 3329–3337.
<https://doi.org/10.1016/j.celrep.2017.11.083>
- Méndez, C., Barberena, R., Reyes, O., & Nuevo Delaunay, A. (2014). Isotopic ecology and human diets in the forest-steppe ecotone, Aisén region, central-western Patagonia, Chile. *International Journal of Osteoarchaeology*, 24(2), 187–201.
<https://doi.org/10.1002/oa.2337>
- Menghini, R., Fiorentino, L., Casagrande, V., Lauro, R., & Federici, M. (2013). The role of

- ADAM17 in metabolic inflammation. *Atherosclerosis*, 228(1), 12–17.
<https://doi.org/10.1016/j.atherosclerosis.2013.01.024>
- Mi, H., Muruganujan, A., Casagrande, J. T., & Thomas, P. D. (2013). Large-scale gene function analysis with the PANTHER classification system. *Nat. Protocols*, 8(8), 1551–1566. Retrieved from <http://dx.doi.org/10.1038/nprot.2013.092>
- Mii, A., Nakajima, T., Fujita, Y., Iino, Y., Kamimura, K., Bujo, H., ... Katayama, Y. (2007). Genetic association of low-density lipoprotein receptor-related protein 2 (LRP2) with plasma lipid levels. *Journal of Atherosclerosis and Thrombosis*, 14(6), 310–316. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/18174661>
- Mitrofanova, A. V., Ahmetov, I. I., Baranova, T. I., Glotov, O. S., Minigalin, A. D., Glotov, A. S., ... Korf, E. A. (2017). Genetic determination of the vascular reactions in humans in response to the diving reflex. *American Journal of Physiology-Heart and Circulatory Physiology*, 312(3), H622–H631.
<https://doi.org/10.1152/ajpheart.00080.2016>
- Monika Goz. (2010). *ADAM-17: The Enzyme That Does It All Monika*. *Crit Rev Biochem Mol Biol*. (Vol. 45). <https://doi.org/10.3109/10409231003628015>.ADAM-17
- Moreno, P. I. (2004). Millennial-scale climate variability in northwest Patagonia over the last 15 000 yr. *Journal of Quaternary Science*, 19(1), 35–47.
<https://doi.org/10.1002/jqs.813>
- Moreno, P. I., & León, A. L. (2003). Abrupt vegetation changes during the last glacial to Holocene transition in mid - latitude South America. *Journal of Quaternary Science*, 18(8), 787–800.
- Mullican, S., & Rangwala, S. (2018). Uniting GDF15 and GFRAL: Therapeutic Opportunities in Obesity and Beyond. *Trends in Endocrinology and Metabolism*, 29, 8.

- Mulligan, C. J., Kitchen, A., & Miyamoto, M. M. (2008). Updated three-stage model for the peopling of the Americas. *PLoS ONE*, *3*(9), 7–10.
<https://doi.org/10.1371/journal.pone.0003199>
- Nakagome, S., Rasmussen, M., Petzelt, B., Lindo, J., & Huerta-sa, E. (2016). A time transect of exomes from a Native American population before and after European contact. <https://doi.org/10.1038/ncomms13175>
- Nelson, M. R., Bryc, K., King, K. S., Indap, A., Boyko, A. R., Novembre, J., ... Lai, E. H. (2008). The Population Reference Sample , POPRES : A Resource for Population , Disease , and Pharmacological Genetics Research, 347–358.
<https://doi.org/10.1016/j.ajhg.2008.08.005>.
- Nielsen, R., Hellmann, I., Hubisz, M., Bustamante, C., & Clark, A. G. (2007). Recent and ongoing selection in the human genome. *Nature Reviews Genetics*, *8*(11), 857–868. <https://doi.org/10.1038/nrg2187>
- Nielsen, R., & Wakeley, J. (2001). Distinguishing migration from isolation: a Markov chain Monte Carlo approach. *Genetics*, *158*(2), 885–896.
- Ocampo, C., & Rivas, P. (2004). Poblamiento Temprano de los Extremos Geográficos de los Canales Patagónicos: Chiloé e Isla Navarino. *Chungara, Revista de Antropología Chilena*, *36*(1), 317–331.
- Okla, M., Kim, J., Koehler, K., & Chung, S. (2017). Dietary Factors Promoting Brown and Beige Fat. *Adv Nutr*, *8*, 473–83. <https://doi.org/10.3945/an.116.014332>.not
- Olson-Manning, C. F., Wagner, M. R., & Mitchell-Olds, T. (2012). Adaptive evolution: evaluating empirical support for theoretical predictions. *Nat Rev Genet*, *13*(12), 867–877. Retrieved from <http://dx.doi.org/10.1038/nrg3322>
- Orquera, L. A., Legoupil, D., & Piana, E. L. (2011). Littoral adaptation at the southern

end of South America. *Quaternary International*, 239(1–2), 61–69.

<https://doi.org/10.1016/j.quaint.2011.02.032>

Orquera, L. A., & Piana, E. L. (2009). Sea Nomads of the Beagle Channel in Southernmost South America: Over Six Thousand Years of Coastal Adaptation and Stability. *The Journal of Island and Coastal Archaeology*, 4(1), 61–81.

<https://doi.org/10.1080/15564890902789882>

PANARELLO, H., ZANGRANDO, F., TESSONE, A., KOZAMEH, L., & TESTA, N. (2006).

ANÁLISIS COMPARATIVO DE PALEODIETAS HUMANAS ENTRE LA REGIÓN DEL CANAL BEAGLE Y PENÍNSULA MITRE: PERSPECTIVAS DESDE LOS ISÓTOPOS ESTABLES . *Magallania (Punta Arenas)* . scielocl .

Patel, S., Alvarez-Guaita, A., Melvin, A., Rimmington, D., Dattilo, A., Miedzybrodzka, E. L., ... O’Rahilly, S. (2019). GDF15 Provides an Endocrine Signal of Nutritional Stress in Mice and Humans. *Cell Metabolism*, 29(3), 707-718.e8.

<https://doi.org/10.1016/j.cmet.2018.12.016>

Patterson, K., & Runge, T. (2002). Smallpox and the Native American. *The American Journal of the Medical Sciences*, 232(4), 216–222.

Patterson, N., Price, A. L., & Reich, D. (2006). Population structure and eigenanalysis.

PLoS Genetics, 2(12), 2074–2093. <https://doi.org/10.1371/journal.pgen.0020190>

Pearson, O. M., & Millones, M. (2005). Rasgos esqueléticos de adaptación al clima y a la actividad entre los habitantes aborígenes de Tierra del Fuego. *Magallania*, 33(1),

37–50. <https://doi.org/10.4067/S0718-22442005000100003>

Pekala, J., Patkowska-Sokota, B., Bodkowski, R., Jamroz, D., Nowakowski, P., Lochynski, S., & Librowski, T. (2011). L-carnitine--metabolic functions and meaning in humans life. *Current Drug Metabolism*, 12(7), 667–678.

- Perego, U. a, Achilli, A., Angerhofer, N., Accetturo, M., Pala, M., Olivieri, A., ... Torroni, A. (2009). Distinctive Paleo-Indian migration routes from Beringia marked by two rare mtDNA haplogroups. *Current Biology : CB*, *19*(1), 1–8.
<https://doi.org/10.1016/j.cub.2008.11.058>
- Pickrell, J. K., Coop, G., Novembre, J., Kudaravalli, S., Li, J. Z., Absher, D., ... Pritchard, J. K. (2009). Signals of recent positive selection in a worldwide sample of human populations Signals of recent positive selection in a worldwide sample of human populations. *Genome Research*, 826–837. <https://doi.org/10.1101/gr.087577.108>
- Pierron, D., Heiske, M., Razafindrazaka, H., Pereda-Loth, V., Sanchez, J., Alva, O., ... Letellier, T. (2018). Strong selection during the last millennium for African ancestry in the admixed population of Madagascar. *Nature Communications*, *9*(1), 1–9.
<https://doi.org/10.1038/s41467-018-03342-5>
- Porsch-Özcürümez, M., Langmann, T., Heimerl, S., Borsukova, H., Kaminski, W. E., Drobnik, W., ... Schmitz, G. (2001). The Zinc Finger Protein 202 (ZNF202) Is a Transcriptional Repressor of ATP Binding Cassette Transporter A1 (ABCA1) and ABCG1 Gene Expression and a Modulator of Cellular Lipid Efflux. *Journal of Biological Chemistry*, *276*(15), 12427–12433.
<https://doi.org/10.1074/jbc.M100218200>
- Pritchard, J. K., & Di Rienzo, A. (2010). Adaptation - not by sweeps alone. *Nature Reviews. Genetics*, *11*(10), 665–667. <https://doi.org/10.1038/nrg2880>
- Pritchard, J. K., Pickrell, J. K., & Coop, G. (2010). The Genetics of Human Adaptation: Hard Sweeps, Soft Sweeps, and Polygenic Adaptation. *Current Biology*, *20*(4), R208–R215. <https://doi.org/10.1016/j.cub.2009.11.055>
- Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). In f e r e n c e o f P o p u l a t i o n S t r u c t u r e U s i n g M u l t i l o c u s G e n o t y p e D a t a.

<https://doi.org/10.1111/j.1471-8286.2007.01758.x>

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M., Bender, D., ... PC, S. (2007). PLINK: a toolset for whole-genome association and population-based linkage analysis. *American Journal of Human Genetics*, (81).

Quattrocchio, M. E., Volkheimer, W., Borrromei, A. M., & Martínez, M. a. (2011). Changes of the palynobiotas in the Mesozoic and Cenozoic of Patagonia: A review. *Biological Journal of the Linnean Society*, 103(2), 380–396.
<https://doi.org/10.1111/j.1095-8312.2011.01652.x>

Quinlan, A., & Hall, I. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, 26(6), 841–842.

R Core Team. (2017). R: A language and environment for statistical computing. Vienna, Austria.

Razzaghi, H., Santorico, S. A., & Ilyas Kamboh, M. (2012). Population-based resequencing of LIPG and ZNF202 genes in subjects with extreme HDL levels. *Frontiers in Genetics*, 3(JUN), 1–16. <https://doi.org/10.3389/fgene.2012.00089>

Reynolds, A. W., Mata-Míguez, J., Miró-Herrans, A., Briggs-Cloud, M., Sylestine, A., Barajas-Olmos, F., ... Bolnick, D. A. (2019). Comparing signals of natural selection between three Indigenous North American populations. *Proceedings of the National Academy of Sciences*, 201819467.
<https://doi.org/10.1073/pnas.1819467116>

Rial, E., & Zardoya, R. (2009). Minireview Oxidative stress, thermogenesis and evolution of uncoupling proteins, 2–6.
<https://doi.org/10.1186/jbiol155>

Rivas, P., Ocampo, C., & Aspillaga, E. (1999). Poblamiento temprano de los canales

patagonicos: El nucleo ecotonal septentrional. *Anales Del Instituto de La Patagonia, Serie Ciencias Humanasa*, 27, 221–230.

Rothhammer, F., & Dillehay, T. D. (2009). The late pleistocene colonization of South America: An interdisciplinary perspective. *Annals of Human Genetics*, 73(5), 540–549. <https://doi.org/10.1111/j.1469-1809.2009.00537.x>

Rundblad, A., Holven, K. B., Bruheim, I., Myhrstad, M. C., & Ulven, S. M. (2018). Effects of fish and krill oil on gene expression in peripheral blood mononuclear cells and circulating markers of inflammation: a randomised controlled trial. *Journal of Nutritional Science*, 7. <https://doi.org/10.1017/jns.2018.2>

Sabeti, P. C., Reich, D. E., Higgins, J. M., Levine, H. Z. P., Richter, D. J., Schaffner, S. F., ... Lander, E. S. (2002). Detecting recent positive selection in the human genome from haplotype structure. *Nature*, 419(6909), 832–837. Retrieved from <http://dx.doi.org/10.1038/nature01140>

Sabeti, P. C., Varilly, P., Fry, B., Lohmueller, J., Hostetter, E., Cotsapas, C., ... Lander, E. S. (2007). Genome-wide detection and characterization of positive selection in human populations. *Nature*, 449(7164), 913–918. Retrieved from <http://dx.doi.org/10.1038/nature06250>

Sagredo, E. a., Moreno, P. I., Villa-Martínez, R., Kaplan, M. R., Kubik, P. W., & Stern, C. R. (2011). Fluctuations of the Última Esperanza ice lobe (52°S), Chilean Patagonia, during the last glacial maximum and termination 1. *Geomorphology*, 125(1), 92–108. <https://doi.org/10.1016/j.geomorph.2010.09.007>

Saito, S., & Shingai, R. (2006). Evolution of thermoTRP ion channel homologs in vertebrates. *Physiological Genomics*, 27(3), 219–230. <https://doi.org/10.1152/physiolgenomics.00322.2005>

Salzano, F. M., & Maria, B. C. (2002). *The Evolution and Genetics of Latin American*

Populations (Cambridge U).

- Scheinsohn, V. (2003). Hunter-Gatherer archaeology in South America. *Annu. Rev. Anthropol.*, 32(2003), 339–361.
<https://doi.org/10.1146/annurev.anthro.32.061002.093228>
- Schlebusch, C. M., Gattepaille, L. M., Engström, K., Vahter, M., Jakobsson, M., & Broberg, K. (2015). Human Adaptation to Arsenic-Rich Environments. *Molecular Biology and Evolution*, 32(6), msv046-. <https://doi.org/10.1093/molbev/msv046>
- Schlebusch, C. M., Jr, C. M. L., Vahter, M., Engström, K., Tito, R. Y., Obregón-tito, A. J., ... Broberg, K. (2013). Possible Positive Selection for an Arsenic-Protective Haplotype in Humans, (1), 53–58.
- Schmitz, G., Heimerl, S., & Langmann, T. (2004). Zinc finger protein ZNF202 structure and function in transcriptional control of HDL metabolism. *Current Opinion in Lipidology*, 15(2), 199–208. <https://doi.org/10.1097/00041433-200404000-00013>
- Schraiber, J. G., & Akey, J. M. (2015). Methods and models for unravelling human evolutionary history. *Nature Publishing Group*, 16(12), 727–740.
<https://doi.org/10.1038/nrg4005>
- Smith, J. M., & Haigh, J. (1974). The hitch-hiking effect of a favourable gene. *Genetics Research*, 23(01), 23–35. <https://doi.org/10.1017/S0016672300014634>
- Stene, M. C. A., Frikke-Schmidt, R., Nordestgaard, B. G., Grande, P., Schnohr, P., & Tybjærg-Hansen, A. (2008). Functional Promoter Variant in Zinc Finger Protein 202 Predicts Severe Atherosclerosis and Ischemic Heart Disease. *Journal of the American College of Cardiology*, 52(5), 369–377.
<https://doi.org/10.1016/j.jacc.2008.03.059>
- Stene, M. C., Frikke-Schmidt, R., Nordestgaard, B. G., & Tybjærg-Hansen, A. (2006). Zinc

- Finger Protein 202 , genetic variation, and HDL cholesterol in the general population . *Journal of Lipid Research*, 47(5), 944–952.
<https://doi.org/10.1194/jlr.m500521-jlr200>
- Szpiech, Z. A., & Hernandez, R. D. (2014). Selscan: An efficient multithreaded program to perform EHH-based scans for positive selection. *Molecular Biology and Evolution*, 31(10), 2824–2827. <https://doi.org/10.1093/molbev/msu211>
- Takahashi, K. (2012). Distribution of urocortins and corticotropin-releasing factor receptors in the cardiovascular system. *International Journal of Endocrinology*, 2012. <https://doi.org/10.1155/2012/395284>
- Tamm, E., Kivisild, T., Reidla, M., Metspalu, M., Smith, D. G., Mulligan, C. J., ... Malhi, R. S. (2007). Beringian standstill and spread of Native American founders. *PLoS One*, 2(9), e829. <https://doi.org/10.1371/journal.pone.0000829>
- Tessone, A., Zangrando, A. F., Barrientos, G., Goñi, R., Panarello, H., & Cagnoni, M. (2009). Stable Isotopes Studies in the Salitroso Lake Basin (Southern Patagonia, Argentina): Assessing Diet of Late Holocene Hunter-Gatherers. *International Journal of Osteoarchaeology*, 19, 297–308. <https://doi.org/10.1002/oa>
- Thesing, C. S., Bot, M., Milaneschi, Y., Giltay, E. J., & Penninx, B. W. J. H. (2019). Psychoneuroendocrinology Omega-3 polyunsaturated fatty acid levels and dysregulations in biological stress systems. *Psychoneuroendocrinology*, 97(July 2018), 206–215. <https://doi.org/10.1016/j.psyneuen.2018.07.002>
- Thornton, R. (1997). Aboriginal North American Population and Rates of Decline, ca. a.d 1500-1901. *Current Anthropology*, 38, 310–315.
- Thornton, T., Tang, H., Hoffmann, T. J., Ochs-Balcom, H. M., Caan, B. J., & Risch, N. (2012). Estimating kinship in admixed populations. *American Journal of Human Genetics*, 91(1), 122–138. <https://doi.org/10.1016/j.ajhg.2012.05.024>

- Tishkoff, B. S. (2015). As 1282 18, 349(6254), 1282–1284.
- Valenzuela, C. Y. (2011). Human sociogenetics. *Biological Research*, 44(4), 393–404.
<https://doi.org/10.4067/S0716-97602011000400012>
- Valverde, G., Zhou, H., Lippold, S., Filippo, C. De, & Tang, K. (2015). A Novel Candidate Region for Genetic Adaptation to High Altitude in Andean Populations, 1–22.
<https://doi.org/10.1371/journal.pone.0125444>
- Van Marken Lichtenbelt, W. D., & Daanen, H. A. M. (2003). Cold-induced metabolism. *Current Opinion in Clinical Nutrition and Metabolic Care*, 6(4), 469–475.
<https://doi.org/10.1097/00075197-200307000-00018>
- Verdugo, A., & Orostica, K. (2018). chromPlot: Global visualization tool of genomic data. R package version 1.10.0.
- Via, M., Gignoux, C. R., Roth, L. A., Fejerman, L., Galanter, J., Choudhry, S., ... Martínez-Cruzado, J. C. (2011). History shaped the geographic distribution of genomic admixture on the island of puerto rico. *PLoS ONE*, 6(1).
<https://doi.org/10.1371/journal.pone.0016513>
- Villa-Martínez, R., & Moreno, P. I. (2007). Pollen evidence for variations in the southern margin of the westerly winds in SW Patagonia over the last 12,600 years. *Quaternary Research*, 68(3), 400–409.
<https://doi.org/10.1016/j.yqres.2007.07.003>
- Vitti, J. J., Grossman, S. R., & Sabeti, P. C. (2013a). Detecting natural selection in genomic data. *Annual Review of Genetics*, 47, 97–120.
<https://doi.org/10.1146/annurev-genet-111212-133526>
- Vitti, J. J., Grossman, S. R., & Sabeti, P. C. (2013b). Detecting Natural Selection in Genomic Data. *Annual Review of Genetics*, 47(1), 97–120.

<https://doi.org/10.1146/annurev-genet-111212-133526>

Voight, B. F., Kudravalli, S., Wen, X., & Pritchard, J. K. (2006a). A map of recent positive selection in the human genome. *PLoS Biology*, 4(3), 0446–0458.

<https://doi.org/10.1371/journal.pbio.0040072>

Voight, B. F., Kudravalli, S., Wen, X., & Pritchard, J. K. (2006b). A Map of Recent Positive Selection in the Human Genome. *PLoS Biology*, 4(3), e72.

<https://doi.org/10.1371/journal.pbio.0040072>

Wagner, S., Hess, M. A., Ormonde-Hanson, P., Malandro, J., Hu, H., Chen, M., ... Bowen, B. R. (2000). A broad role for the zinc finger protein ZNF202 in human lipid metabolism. *Journal of Biological Chemistry*, 275(21), 15685–15690.

<https://doi.org/10.1074/jbc.M910152199>

Wang, K., Li, M., & Hakonarson, H. (2010). ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Research*, 38(16), 1–7. <https://doi.org/10.1093/nar/gkq603>

Wang, S., Ray, N., Rojas, W., Parra, M. V., Bedoya, G., Gallo, C., ... Ruiz-Linares, A. (2008). Geographic patterns of genome admixture in latin American mestizos. *PLoS Genetics*, 4(3), 1–9. <https://doi.org/10.1371/journal.pgen.1000037>

Weber, J. M., & Haman, F. (2005). Fuel selection in shivering humans. *Acta Physiologica Scandinavica*, 184(4), 319–329. <https://doi.org/10.1111/j.1365-201X.2005.01465.x>

Weir, B. S., & Cockerham, C. (1984). Estimating F-Statistics for the Analysis of Population Structure. *Evolution*, 38(6), 1358–1370.

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. New York: Springer-Verlag.

Wu, Q., Zheng, P., Hu, Y., & Wei, F. (2014a). Genome-scale analysis of demographic

- history and adaptive selection. *Protein and Cell*, 5(2), 99–112.
<https://doi.org/10.1007/s13238-013-0004-1>
- Wu, Q., Zheng, P., Hu, Y., & Wei, F. (2014b). Genome-scale analysis of demographic history and adaptive selection. *Protein and Cell*, 5(2), 99–112.
<https://doi.org/10.1007/s13238-013-0004-1>
- Xie, X., Wang, X., Mick, G. J., Kabarowski, J. H., Wilson, L. S., Barnes, S., ... Luo, X. (2016). Yorkshire Boars, 157(April), 1512–1521. <https://doi.org/10.1210/en.2015-1831>
- Xiong, Y., Qu, Z., Chen, N., Gong, H., Song, M., Chen, X., ... Xu, C. (2014). The local corticotropin-releasing hormone receptor 2 signalling pathway partly mediates hypoxia-induced increases in lipolysis via the cAMP-protein kinase A signalling pathway in white adipose tissue. *Molecular and Cellular Endocrinology*, 392(1–2), 106–114. <https://doi.org/10.1016/j.mce.2014.05.012>
- Yakhini, Z., Eden, E., Navon, R., Steinfeld, I., & Lipson, D. (2009). GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics*, 10(1), 1–7. <https://doi.org/10.1186/1471-2105-10-48>
- Yang, L., Chang, C. C., Sun, Z., Madsen, D., Zhu, H., Padkjær, S. B., ... Jørgensen, S. B. (2017). GFRAL is the receptor for GDF15 and is required for the anti-obesity effects of the ligand. *Nature Medicine*, 23(10), 1158–1166.
<https://doi.org/10.1038/nm.4394>
- Yesner, D. R., Figuerero Torres, M. J., Guichon, R. A., & Borrero, L. A. (2003). Stable isotope analysis of human bone and ethnohistoric subsistence patterns in Tierra del Fuego. *Journal of Anthropological Archaeology*, 22(3), 279–291.
[https://doi.org/10.1016/S0278-4165\(03\)00040-0](https://doi.org/10.1016/S0278-4165(03)00040-0)
- Yi, X., Liang, Y., Huerta-Sanchez, E., Jin, X., Cuo, Z. X. P., Pool, J. E., ... Wang, J. J. (2010). Sequencing of Fifty Human Exomes Reveals Adaptation to High Altitude. *Science*,

329(5987), 75–78. <https://doi.org/10.1126/science.1190371>. Sequencing

Zhang, C., Bailey, D. K., Awad, T., Liu, G., Xing, G., Cao, M., ... Kennedy, G. C. (2006). A whole genome long-range haplotype (WGLRH) test for detecting imprints of positive selection in human populations. *Bioinformatics*, *22*(17), 2122–2128. <https://doi.org/10.1093/bioinformatics/btl365>

Zhou, Q., Zhao, L., & Guan, Y. (2016). Strong Selection at MHC in Mexicans since Admixture. *PLoS Genetics*, *12*(2), 1–17. <https://doi.org/10.1371/journal.pgen.1005847>

Zunke, F., & Rose-John, S. (2017). The shedding protease ADAM17: Physiology and pathophysiology. *Biochimica et Biophysica Acta - Molecular Cell Research*, *1864*(11), 2059–2070. <https://doi.org/10.1016/j.bbamcr.2017.07.001>