



UNIVERSIDAD NACIONAL DE COLOMBIA

**CARACTERIZACIÓN MOLECULAR Y ANÁLISIS EVOLUTIVO DE RECEPTORES
DE CÉLULAS NATURAL KILLER TIPO LECTINA EN PRIMATES
NEOTROPICALES**

John Alexander Galindo Puentes

Universidad Nacional de Colombia

Facultad de Ciencias, Departamento de Biología

Bogotá, Colombia

2012

**CARACTERIZACIÓN MOLECULAR Y ANÁLISIS EVOLUTIVO DE
RECEPTORES DE CÉLULAS NATURAL KILLER TIPO LECTINA EN
PRIMATES NEOTROPICALES**

John Alexander Galindo Puentes

Tesis presentada como requisito parcial para optar por el título de:

Magister en Ciencias Biológicas

Director:

MD Ph. D. Luis Fernando Cadavid

Línea de investigación:

Genética

Grupo de Investigación:

Inmunología Evolutiva e Inmunogenética

Universidad Nacional de Colombia

Facultad de Ciencias, Departamento de Biología

Bogotá, Colombia

2012

Resumen

La mayoría de las especies de primates neotropicales en Colombia se encuentran amenazadas o en peligro de extinción, debido principalmente a la destrucción de su hábitat y al aumento de la morbi-mortalidad como consecuencia de enfermedades infecciosas y malignas. En la interfase entre el sistema inmune adaptativo y el innato de los mamíferos, las células Natural Killer (NK) juegan un papel fundamental en el reconocimiento y eliminación de células malignas e infectadas por virus. La función de las células NK está regulada por receptores activadores e inhibidores específicos para moléculas del Complejo Mayor de Histocompatibilidad clase I presentes en las células blanco. Los receptores de células NK se clasifican en dos familias estructurales, la familia de las inmunoglobulinas y la de las lectinas. En el presente trabajo se caracterizaron dos receptores de células NK tipo lectina específicos para MHC clase I: CD94 y NKG2D, en nueve especies de primates de Colombia. RNA total se extrajo a partir de sangre periférica y fue amplificado por retrotranscriptasa reversa (RT)-PCR con iniciadores específicos. Los productos de amplificación fueron clonados y al menos 10 clones por cada individuo se secuenciaron por el método de Sanger. Las secuencias obtenidas mostraron un alto grado de conservación con aquellas de humanos y otros primates. Análisis filogenéticos con estas secuencias produjeron árboles que reflejan la filogenia de las especies. A pesar de su conservación, los genes que codifican para estos dos receptores tienen la capacidad de generar niveles altos de variabilidad mediante mecanismos de procesamiento alternativo del RNA. Catorce isoformas de CD94 y seis de NKG2D se identificaron, algunas de ellas presentes en todas las especies estudiadas mientras que otras están restringidas a solo un grupo de

VI CARACTERIZACIÓN MOLECULAR Y ANÁLISIS EVOLUTIVO DE RECEPTORES DE CÉLULAS NATURAL KILLER TIPO LECTINA EN PRIMATES NEOTROPICALES

especies. Las variantes comunes a varias especies han sido probablemente seleccionadas para llevar a cabo funciones distintas a la de la forma completa.

El modelamiento por homología de las estructuras mostró que la mayoría de las variantes son incapaces de interactuar con las moléculas del MHC clase I en la célula blanco, pero conservan regiones que podrían interactuar con otros ligandos. Los resultados de este estudio indican que los receptores de células NK tipo lectina CD94 y NKG2D son altamente conservados para funciones de reconocimiento del MHC, pero muestran un alto grado de diversificación por mecanismos de procesamiento alternativo del ARN que probablemente amplían su rango de especificidades.

Palabras clave: Receptores Natural Killer. CD94. NKG2D. Primates neotropicales. Procesamiento alternativo del mARN

Abstract

Most species of New World primates in Colombia are threatened or endangered, mainly due to habitat destruction and increased morbidity and mortality from infectious and malignant diseases. At the interface between the adaptive and the innate immune system of mammals, Natural Killer (NK) cells play a critical role in the recognition and elimination of malignant and virally infected cells. The role of NK cells is regulated by activators and inhibitors receptors that are specific for molecules of the Major Histocompatibility Complex class I present in the target cells. NK cell receptors are classified into two structural families, the immunoglobulins and lectins classes. In this thesis in nine Colombian New World Primate species, the CD94 and NKG2D C type lectine receptors specific to MHC class I molecules were characterized. Total RNA was extracted from peripheral blood, it was amplified by reverse transcriptase (RT)-PCR using specific primers. The amplification products were cloned and at least 10 clones per individual were sequenced by the Sanger method. The gene sequences are highly conserved with humans and other primates. Phylogenetic analysis with this results produced trees that reflect the species phylogeny. Although conservation, the genes coding for these two receptors are capable to produce high levels of variability by alternative splicing mechanisms. Fourteen isoforms of CD94 and six of NKG2D were identified, some of which are present in all species studied while others are restricted only one group of species. In several species common variants have probably been evolutionary selected to perform different functions of the complete protein. Homology modeling of structures showed that most of the variants are unable to interact with MHC class I molecules on the target cell, however, these retain regions that could be interact with other ligands.

VIII CARACTERIZACIÓN MOLECULAR Y ANÁLISIS EVOLUTIVO DE RECEPTORES DE CÉLULAS NATURAL KILLER TIPO LECTINA EN PRIMATES NEOTROPICALES

The results of this study indicate that CD94 and NKG2D NK lectine receptors are hihhly conserved for MHC recognition functions, but show a high degree of diversification led by alternative splicing mechanisms that probably expand their range of specificities.

Keywords: Natural killer cell receptors. CD94. NKG2D. New World primates. Alternative splicing.

Contenido

Resumen.....	V
Abstract	VII
Lista de figuras.....	XII
Lista de Tablas	XIII
1. Introducción.....	1
2. Objetivos.....	5
2.1 Objetivo general	5
2.2 Objetivos Específicos.....	5
3. Marco Teórico.....	7
3.1. LOS PRIMATES EN COLOMBIA.....	7
3.2 LAS CÉLULAS NATURAL KILLER (NK)	12
3.3 LOS RECEPTORES DE LAS CÉLULAS NATURAL KILLER	13
3.4 RECEPTORES DE CÉLULAS NK TIPO LECTINA.....	17
3.5 EL COMPLEJO NKC EN PRIMATES DEL NUEVO MUNDO	25
4. METODOLOGÍA	27
4.1 MUESTRAS DE PRIMATES Y EXTRACCIÓN DE ARN	27

X **CARACTERIZACIÓN MOLECULAR Y ANÁLISIS EVOLUTIVO DE RECEPTORES DE CÉLULAS NATURAL KILLER TIPO LECTINA EN PRIMATES NEOTROPICALES**

4.2 AMPLIFICACIÓN DE LOS GENES CD94 Y NKG2	28
4.3 CLONACIÓN Y SECUENCIACIÓN DE PRODUCTOS DE PCR	29
4.4. BÚSQUEDA DE SECUENCIAS EN BASES DE DATOS	30
4.5. ANÁLISIS DE SECUENCIAS	31
4.6. ANÁLISIS DE SELECCIÓN	31
4.7. RECONSTRUCCIÓN FILOGENÉTICA.....	32
4.8 ESTIMACIÓN DE TIEMPOS DE DIVERGENCIA CON INFERENCIA BAYESIANA	34
4.9 PREDICCIÓN Y MODELAMIENTO DE PROTEÍNAS.....	34
5. RESULTADOS	35
5.1 ALTA CONSERVACIÓN DE CD94 EN PRIMATES DEL NUEVO MUNDO	35
5.2. SELECCIÓN PURIFICADORA EXPLICA LA ALTA CONSERVACIÓN DE CD94 EN PRIMATES DEL NUEVO MUNDO.....	39
5.3 ANÁLISIS FILOGENÉTICO DE CD94.....	40
5.3.1. La filogenia de CD94 refleja la filogenia de las especies	40
5.3.2 Estimación de tiempos de divergencia de CD94 en primates.	41
5.3.3 Diversificación de CD94 por procesamiento alternativo del ARN	45
5.3.4. Predicción y modelamiento de la proteína CD94 y sus isoformas ..	50
5.4 CARACTERIZACIÓN MOLECULAR DEL GEN NKG2D	61
5.4.1 Alta conservación de NKG2D en primates del Nuevo Mundo.....	61
5.5 NKG2D ESTÁ SUJETO A UNA FUERTE PRESIÓN PURIFICADORA.....	65
5.6 ANÁLISIS FILOGENÉTICO DEL GEN NKG2D	66
5.6.1 Filogenia de NKG2D en los primates del Nuevo Mundo	66

5.6.2 Estimación de tiempos de divergencia por inferencia Bayesiana.	68
5.6.3 Isoformas de NKG2D generadas por procesamiento alternativo del ARN.....	70
5.6.4 La estructura de NKG2D es conservada en el linaje de los primates del Nuevo Mundo.....	72
6. Discusión de Resultados.....	77
6.1 CD94 ES UN GEN ALTAMENTE CONSERVADO EN EL LINAJE DE LOS PRIMATES.....	77
6.2. ASPECTOS FUNCIONALES DE LAS ISOFORMAS DE CD94	78
6.3. PROCESAMIENTO ALTERNATIVO DEL ARN COMO MECANISMO DE DIVERSIDAD	82
6.4. NKG2D ES UN GEN ALTAMENTE CONSERVADO EN EL LINAJE DE LOS PRIMATES.....	85
6.5 PROCESAMIENTO ALTERNATIVO DEL ARN EN NKG2D	86
6.6 IMPLICACIONES FUNCIONALES DE NKG2D	87
7. Conclusiones y Recomendaciones	89
8. Bibliografía	91

Lista de figuras

Figura 3-1. Filogenia de los Platyrrhini.	8
Figura 3-2. Mapa de la distribución de los primates en Colombia	10
Figura 3-3. Receptores de células NK y sus ligandos del MHC clase I	15
Figura 3-4. Organización del complejo NKC y del complejo LRC en humano y ratón	16
Figura 3-5. Organización del complejo NKC en el genoma humano.....	18
Figura 3-6. Estructura del gen CD94 donde se indica el tamaño y dominio del exón.	20
Figura 3-7. Sitios de interacción del MHC E con CD94 y NKG2	
Figura 5-8. Alineamiento de CD94 en primates del Nuevo Mundo.....	36
Figura 5-9. Árbol filogenético del gen CD94 en primates.....	43
Figura 5-10. Tiempos de divergencia del gen CD94 estimados por inferencia Bayesiana en primates.....	44
Figura 5-11. Isoformas de CD94 en primates del nuevo mundo.....	49
Figura 5-12. Estructura de CD94.....	
Figura 5-13. Estructuras predichas de CD94 en primates del Nuevo Mundo	51
Figura 5-14. Sobreposición de las estructuras de CD94 en primates del Nuevo Mundo	52
Figura 5-15. Análisis por modelamiento de la isoforma CD94B	55
Figura 5-16. Análisis estructural de la isoforma SV1 [^]	56
Figura 5-17. Estructura de la Isoforma SV6	57
Figura 5-18. Análisis estructural de la Isoforma SV12	58
Figura 5-19. Análisis estructural de la Isoforma SV13	59
Figura 5-20. Alineamiento de NKG2D en primates del Nuevo Mundo.....	62
Figura 5-21. Árbol filogenético de NKG2D.....	67
Figura 5-22. Tiempos de divergencia del gen NKG2D estimados por inferencia Bayesiana.....	69
Figura 5-23. Isoformas de NKG2D en primates del Nuevo Mundo.....	71

Lista de Tablas

	Pág.
Tabla 5-1. Porcentaje de identidad nucleotídica y de proteínas de CD94 en primates del Nuevo Mundo.	37
Tabla 5-2. Distancias evolutivas de CD94 en primates del Nuevo Mundo.....	38
Tabla 5-3. Codones seleccionados negativamente en al menos dos pruebas de selección para CD94.	40
Tabla 5-4. Porcentaje de identidad nucleotídica y de proteínas en NKG2D de primates del Nuevo Mundo.....	63
Tabla 5-5. Estimación de distancias evolutivas para NKG2D.....	64
Tabla 5-6. Codones seleccionados negativamente en al menos dos pruebas de selección para NKG2D.	65

1. Introducción

En Colombia existen entre 27 y 31 especies de primates del Nuevo Mundo, lo cual es tan solo una muestra de la gran biodiversidad que caracteriza en nuestro territorio. Desafortunadamente, 19 de estas especies están reportadas en la lista de especies amenazadas de la Unión Internacional para la Conservación de la Naturaleza (UICN) como en estado crítico, de amenaza o vulnerabilidad (1). Aún más preocupante es que las especies endémicas presentan un grado máximo de amenaza (1). Dentro de las causas de esta situación se encuentra la fragmentación del bosque primario, la presión antrópica, la ampliación de la frontera agrícola, la explotación del bosque, la ganadería, la caza y el tráfico ilegal (1,2). Sumado al problema ambiental, los primates neotropicales en su estado silvestre son susceptibles a enfermedades infecciosas de tipo viral como la fiebre amarilla, el dengue, el ébola, la influenza, el herpes y los virus de inmunodeficiencia (3,4). Igualmente, las enfermedades parasitarias como la malaria y las generadas por parásitos sanguíneos e intestinales como microfiliarias y nemátodos son también un problema de salud en estas especies (5,6). El cáncer es otra enfermedad que afecta la salud de los primates neotropicales y como se ha demostrado en otras especies, su etiología está directamente relacionada con infecciones virales (7). El deterioro en las condiciones medioambientales y la consecuente disminución del rango de distribución de estas poblaciones, ha facilitado la transmisión de infecciones y ha aumentado la morbilidad de especies vulnerables. Adicionalmente, la salud

humana puede verse afectada debido a un aumento en el contacto entre humanos y poblaciones de primates silvestres con individuos enfermos(3,4).

Las células Natural Killer (NK) son uno de los agentes inmunológicos más importantes contra parásitos intracelulares y células malignas (8). Las células NK son linfocitos citotóxicos que censan la integridad de las células propias mediante el reconocimiento de moléculas del Complejo Mayor de Histocompatibilidad (MHC) clase I en su superficie (9). A diferencia de otras poblaciones de linfocitos, estas no tienen receptores específicos contra el antígeno. Una disminución de la expresión de las moléculas del MHC clase I en la célula blanco activa la célula NK estimulando su actividad citotoxicidad, mientras que niveles de expresión adecuados del MHC clase I inhiben su citotoxicidad (10). Las células NK también tienen funciones reguladoras de la respuesta inmune celular mediante la producción de citoquinas que activan la capacidad fagocítica y bactericida de los macrófagos, regulan la activación de linfocitos T y B y favorecen las funciones de presentación de antígeno en las células dendríticas (8–10). Las células NK están en la interfase entre la inmunidad innata y la adaptativa, lo que hace que su estudio sea importante desde el punto de vista funcional y evolutivo.

El reconocimiento de las moléculas del MHC clase I en las células blanco por parte de las células NK se lleva a cabo mediante receptores específicos que pertenecen a dos familias estructurales, la familia de las inmunoglobulinas (Killer cell Ig-like receptor, KIRs) codificadas en el complejo de receptores leucocitarios (Leukocyte Receptor Complex, LRC) y la familia de las lectinas, codificadas en el complejo Natural Killer (Natural Killer Complex NKC) (8,10). Ambas familias tienen receptores activadores e inhibitorios, aunque la señalización mediada por estos últimos

predomina sobre la señalización activadora, protegiendo así a las células normales de su destrucción (8,11). Los receptores de las células NK se caracterizan por su alta variabilidad estructural, genética y haplotípica, y los complejos LRC y NKC muestran una plasticidad evolutiva similar a aquella del MHC con constantes ciclos de duplicación y eliminación de genes en distintos linajes de mamíferos (12).

La estructura, función y evolución de los receptores de células NK tipo lectina en los primates neotropicales hasta ahora se está empezando a conocer. Los primeros estudios realizados en estos primates se realizaron en las especies *Saimiri sciureus* (mono ardilla) y *Callithrix jacchus* (marmoset) donde se observó que los genes del complejo NKC son altamente conservados (13,14). En el presente trabajo se caracterizaron molecularmente los receptores de células NK de la familia de las lectinas CD94 y NKG2D en 10 especies de primates del neotrópico.

4 **CARACTERIZACIÓN MOLECULAR Y ANÁLISIS EVOLUTIVO DE
RECEPTORES DE CÉLULAS NATURAL KILLER TIPO LECTINA EN PRIMATES
NEOTROPICALES**

2. Objetivos

2.1 Objetivo general

Caracterizar molecularmente los receptores del complejo NKC CD94 y NKG2 de las células Natural Killer y evaluar sus relaciones evolutivas e interacciones moleculares en diez especies de primates neotropicales *Cebus albifrons*, *Cebus apella*, *Cebus capucinus*, *Saguinus leucopus*, *Ateles hybridus*, *Saimiri sciureus*, *Aotus sp*, *Lagothrix lagothricha* y *Alouatta seniculus*.

2.2 Objetivos Específicos

Obtener muestras de sangre de las especies *Cebus albifrons*, *Cebus apella*, *Cebus capucinus*, *Saguinus leucopus*, *Ateles hybridus*, *Saimiri sciureus*, *Aotus sp*, *Lagothrix lagothricha* y *Alouatta seniculus*.

Secuenciar los genes CD94 y NKG2 de las especies de primates neotropicales obtenidas.

Establecer las relaciones filogenéticas de los genes CD94 y NKG2 en primates utilizando los algoritmos de reconstrucción evolutiva basados en secuencias nucleotídicas y de aminoácidos.

Predecir las interacciones moleculares de los receptores CD94/NKG2 con las moléculas del MHC de los primates utilizando programas de interacción y predicción de proteínas.

6 **CARACTERIZACIÓN MOLECULAR Y ANÁLISIS EVOLUTIVO DE
RECEPTORES DE CÉLULAS NATURAL KILLER TIPO LECTINA EN PRIMATES
NEOTROPICALES**

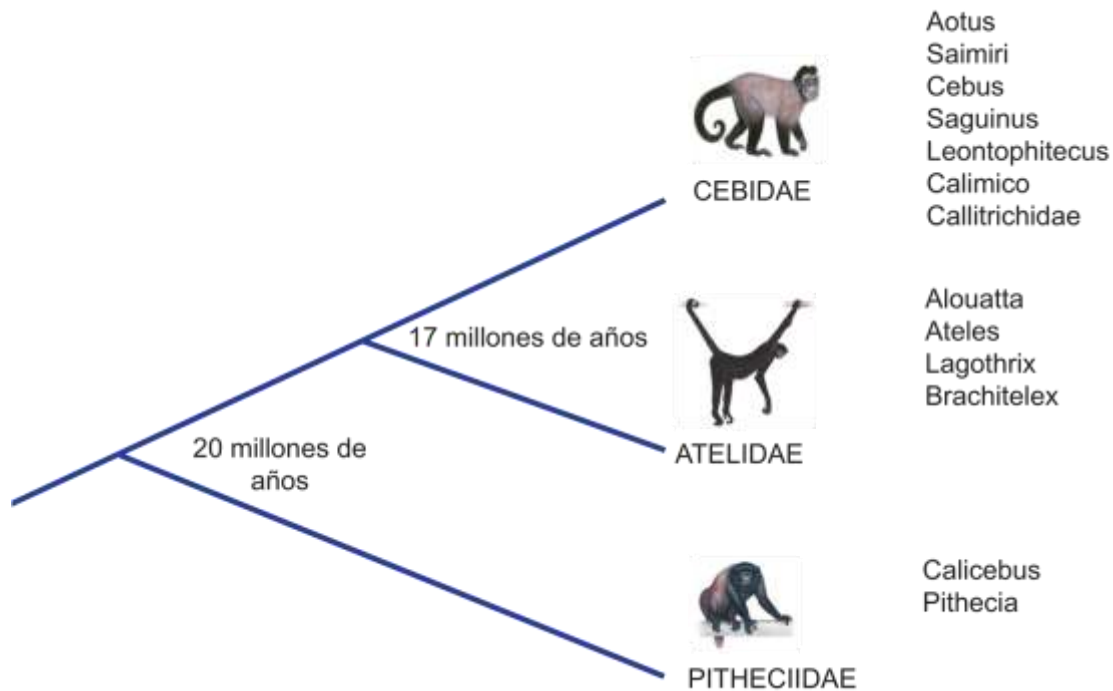
3. Marco Teórico

3.1. LOS PRIMATES EN COLOMBIA

Los primates del neotrópico se distribuyen en una extensa área que va desde México hasta el norte de Argentina y aunque el registro fósil no es contundente, se presume que se originaron hace 35-40 millones de años de un ancestro común compartido con los monos del Viejo Mundo (15,16). Análisis filogenéticos basados en datos moleculares y morfológicos muestran que los primates del Nuevo Mundo (infraorden Plathyrrhini) pueden agruparse en tres familias, Pitheciidae, Atelidae y Cebidae (15,16). Adicionalmente, estudios con las secuencias del gen de la globina ϵ , muestran que la primera separación entre los linajes de primates del Nuevo Mundo ocurrió entre las familias Pitheciidae y los Cebidae hace cerca de 20 millones de años y un evento posterior de divergencia entre Cebidae y los Atelidae ocurrió hace 17 millones de años aproximadamente (15). Los Pitheciidae ocupan una posición basal entre los Plathyrrhini siendo el grupo soporte del grupo hermano de Cebidae y Atelidae (17). La especiación y divergencia de estos primates se ha seguido dando de manera dinámica y en la actualidad se reconocen 15 géneros de primates del Nuevo Mundo (17). Dentro de los géneros de la familia Atelidae se considera *Alouatta* como el más basal mientras que *Lagothrix* y *Brachyletes* y *Ateles* y *Alouatta* forman grupos hermanos. Las relaciones entre los demás géneros no están totalmente resueltas (17). Por otro lado, aunque la familia Cebidae y la subfamilia Callithrichinae son monofiléticas, mientras las relaciones entre *Aotus*, *Saimiri* y *Cebus* no está resuelta del todo y muestra algunas contradicciones que se deben en parte a la rápida dinámica evolutiva de los primates neotropicales(17,18)

(17,18). En la Figura 3-1, se observan las tres familias de Platyrrhini: Pitheciidae, Atelidae y Cebidae con sus respectivos géneros.

Figura 3-1. Filogenia de los Platyrrhini. Se indican las tres familias de primates del Nuevo Mundo Cebidae, Atelidae y Pitheciidae y los tiempos estimados de divergencia. Para cada familia, se indican los géneros. Modificado de Osterholz, 2009. Las ilustraciones fueron tomadas del libro de DeFler, TR, "Primates de Colombia" 2004 con la autorización del autor.

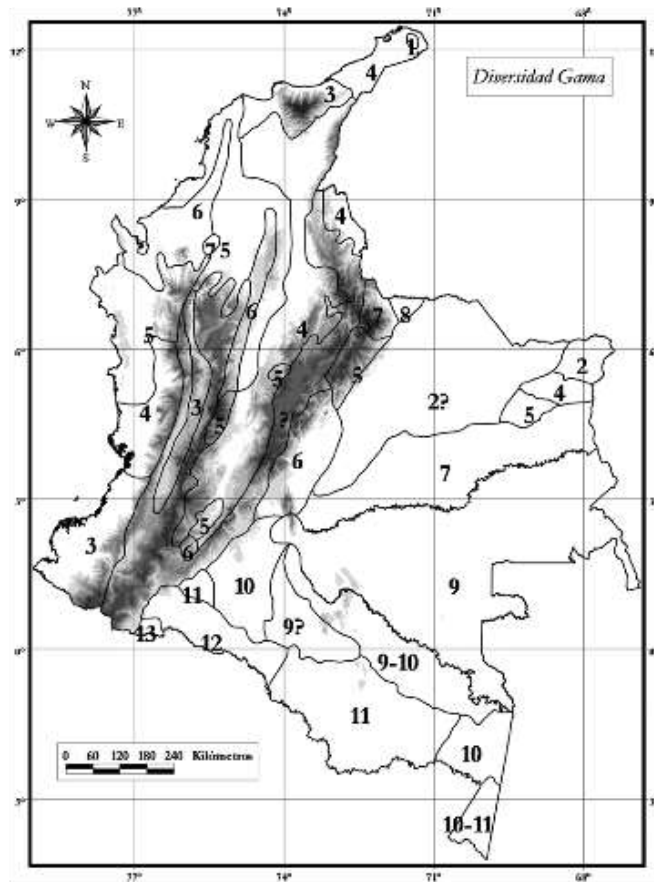


Los monos del Nuevo Mundo se caracterizan porque sus manos y pies se han especializado para coger y manipular objetos. Son principalmente arbóreos, aunque se pueden encontrar en el suelo buscando alimentos o interactuando con otras especies. Estos primates presentan organizaciones sociales complejas y sus interacciones se dan por medio de vocalizaciones, contactos físicos y señales visuales, las cuales ayudan a la cohesión del grupo. Pueden convivir en manadas

de diferentes especies en una misma zona del bosque o se pueden observar individuos solitarios. A su vez, su comportamiento sexual y reproductivo varía desde individuos monógamos a polígamos. Las crías son máximo de dos por periodo de gestación reflejando una reducción del índice de reproducción debido a un aumento en el tiempo del cuidado parental (1).

Colombia es un país con una gran variedad topográfica y un alto número de ecosistemas donde confluyen varias unidades biogeográficas y biota perteneciente a la región de América Central, la Amazonía y la región Andina. En Colombia se han reportado entre 27 y 31 especies (Figura 3-2), de los cuales 6 se encuentran en estado crítico, 5 amenazados y 11 vulnerables de acuerdo con los reportes de la Unión Internacional para la Conservación de la Naturaleza (1). Las especies *Ateles hybridus hybridus*, *Ateles hybridus brunneus*, *Ateles geoffroyi rufiventris*, *Lagothrix lagothricha lugens*, *Saguinus oedipus*, *Callicebus caquetensis* son especies endémicas o casi endémicas para Colombia y presentan un grado de máxima amenaza (1,19).

Figura 3-2. Mapa de la distribución de los primates en Colombia. Se indica el número de especies reportadas por región. Tomado de Defler, TR, 2004.



Las causas de esta situación de vulnerabilidad o riesgo de extinción son variadas y complejas, e incluyen el deterioro ambiental generado por la tala del bosque, las actividades agrícolas y ganaderas, la caza, la fragmentación del bosque primario y el tráfico ilegal, entre otros. Un caso particular de deterioro ambiental se ha dado en la planicie del Caribe, donde el establecimiento de cultivos y pastizales y las actividades madereras han deteriorado el hábitat de especies endémicas y vulnerables como *Saguinus oedipus* y *Saguinus leucopus*. El género *Ateles* es el que presenta un mayor riesgo de extinción en el país y se debe en gran parte a la cacería

que afecta principalmente a individuos en edades reproductivas. Otra presión ejercida sobre los primates es su uso como mascotas,

afectando a un gran número de especies y generando daños dentro de las manadas ya que muchas veces en la captura se sacrifican a las madres de los individuos (1).

Los primates neotropicales desarrollan enfermedades infecciosas y cancerosas que afectan gravemente la salud de los individuos y de las poblaciones (6,7). En Brasil, primates neotropicales en cautiverio han sido diagnosticados con *Leishmania chagasi* y han desarrollado Leishmania visceral con serias complicaciones clínicas (6). También se ha observado que los grupos de *Callithrix jacchus*, *Callithrix pygmea* y *Pithecia pithecia* son altamente sensibles a infecciones parasitarias por *Toxoplasma gondii* generandoles lesiones necróticas en pulmones, intestino e hígado (20). En algunas regiones en Panamá, Trinidad y Brasil, se han encontrado monos principalmente del género *Alouatta* sp., muertos a causa de infecciones por fiebre amarilla (21). Estudios histológicos y endoscópicos realizados en *Saguinus oedipus*, especie endémica de Colombia, han mostrado que estos individuos desarrollan cáncer de colon con características similares al cáncer colorrectal en humanos y, aunque no se conoce completamente la etiología de la enfermedad en los monos, se entrevé una asociación con las condiciones de cautiverio, como el estrés, la dieta y la actividad social (22). Estas y otras patologías, sumadas a la fragmentación del bosque, pueden acelerar la extinción de los primates debido al aumento de la transmisión de infecciones dentro de las poblaciones. A su vez, esto puede ser un problema de salud pública debido a la interacción entre humanos y primates silvestres, principalmente en regiones tropicales, donde se conserva una estrecha relación con los reservorios y con la transmisión de algunas enfermedades infecciosas. Se ha visto que este fenómeno de zoonosis ha venido ocurriendo a lo largo de la evolución y son ejemplos las enfermedades como la fiebre amarilla, el

dengue, la malaria, el ébola, la influenza, el herpes y más recientemente los virus de inmunodeficiencia (3,4,21).

3.2 LAS CÉLULAS NATURAL KILLER (NK)

El sistema inmune de los mamíferos presenta una notable complejidad de mecanismos que interactúan para reconocer, controlar y eliminar agentes potencialmente patógenos. Un primer nivel de protección generalista es regulado por la inmunidad innata que despliega una variada gama de mecanismos para la defensa del hospedero que incluye sistemas de reconocimiento de patrones moleculares, como los receptores Toll-like (Toll-like receptors), las barreras anatómicas y fisiológicas y las células fagocíticas (23,24). El segundo nivel de protección está representado por la inmunidad adaptativa que se caracteriza por tener células especializadas, los linfocitos B y T, que expresan receptores contra el antígeno (inmunoglobulinas y receptores de células T) con altísima diversidad y especificidad, siendo capaces de reconocer con precisión potencialmente cualquier toxina o patógeno. Las inmunoglobulinas reconocen el antígeno directamente, mientras que los receptores de células T requieren que el antígeno les sea presentado por las moléculas de MHC. El reconocimiento antigénico permite la expansión clonal de los linfocitos y la generación de una memoria inmunológica protectora a largo plazo haciendo que una respuesta secundaria contra el mismo antígeno sea mas rápida e intensa que la primera. Los linfocitos controlan la respuesta inmune humoral y celular donde intervienen otros tipos de células como los fagocitos, las células dendríticas y las células Natural Killer (23,24).

Las células Natural Killer (NK) son linfocitos efectores que tienen la capacidad de reconocer y destruir células infectadas por virus, bacterias intracelulares, parásitos y células tumorales (23). Las células NK tienen actividad citotóxica mediada por gránulos citoplasmáticos que contienen perforinas y granzimas. Estos gránulos son liberados en la interfase celular e inducen apoptosis en la célula blanco mediante la activación de la vía de las caspasas (25). Adicionalmente, las células NK tienen una función reguladora de la inmunidad innata y adaptativa mediante la producción de citoquinas como el interferón gamma (IFN- γ) el cual puede activar la capacidad fagocítica y de ataque microbiano de los macrófagos (23,26). Así mismo, una subpoblación de células NK asociadas al tejido linfoide de la mucosa, secreta interleuquina (IL)-22 la cual tiene diferentes efectos proinflamatorios y antiinflamatorios a nivel de mucosas (26). Ciertamente, estas células NK22 inducen la proliferación de células epiteliales del colón, activan rutas antiapoptóticas y secretan IL-10, lo cual es importante en la protección de la mucosa (26). Adicionalmente, las células NK desempeñan un papel importante en la regulación de la hematopoyesis y en el desarrollo placentario (27,28).

3.3 LOS RECEPTORES DE LAS CÉLULAS NATURAL KILLER

La función citotóxica en las células NK es regulada por múltiples tipos de receptores de membrana que generan señales de activación o inhibición. Estos receptores interactúan con las moléculas MHC clase I en la superficie de la célula blanco. Típicamente, la señalización inhibitoria predomina sobre la activadora, de tal modo que la interacción entre células NK y células normales conduce a anergia. En células infectadas por virus o cancerosas se disminuye o elimina la expresión de moléculas

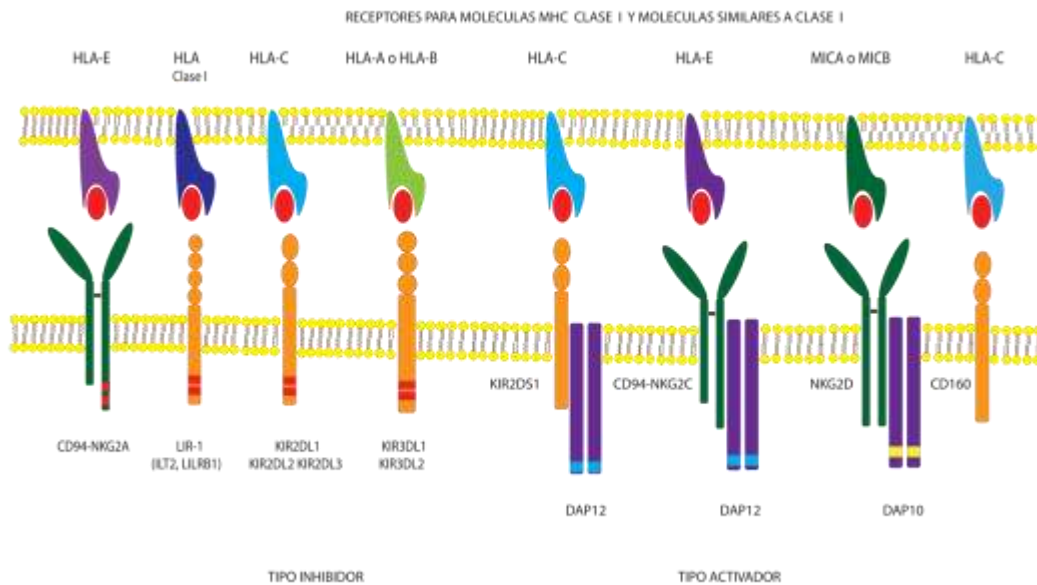
de MHC clase I, lo que conduce a una pérdida de la señalización inhibitoria dando como resultado la señalización activadora y la subsecuente inducción de apoptosis en la célula blanco (29,30).

Estructuralmente, los receptores de las células NK que reconocen a las moléculas MHC I pertenecen a dos superfamilias, la de las inmunoglobulinas (Immunoglobulin-like receptor o KIR) y la de las lectinas (C-type lectin like) (10)(Figura 3-3). Los KIRs son codificados en el complejo de receptores de leucocitos (LRC) localizado en el cromosoma 19 en humanos (12) y comprenden un número variable de proteínas con dos o tres dominios extracelulares de tipo inmunoglobulina y un dominio citoplasmático que puede ser largo o corto. Los KIRs con dominios citoplasmáticos largos son inhibitorios mientras que aquellos con dominios cortos son activadores. Los receptores de células NK tipo lectinas se codifican en el complejo Natural Killer (NKC) localizado en el brazo corto del cromosoma 12 en humanos. Estos receptores tienen dominios extracelulares de lectina tipo C encargados de reconocer a las moléculas de MHC clase I en la célula blanco, así como también dominios citoplasmáticos largos o cortos con funciones inhibitorias y activadoras, respectivamente (10).

Los receptores de células NK inhibitorios, tanto de la familia de las inmunoglobulinas como de las lectinas, tienen regiones citoplasmáticas largas debido a la presencia de motivos ITIM (Immunoreceptor Tyrosine-based Inhibitory Motif). Estos motivos fosforilan su tirosina y se asocian con fosfatas tipo Src como SHP1, las cuales defosforilan al regulador de citoesqueleto de actina Vav, que a su vez bloquea las señales de activación dependientes de actina (29,31–34). Por otro lado, los receptores activadores tienen regiones citoplasmáticas cortas debido a la ausencia

de ITIMs. Estos receptores, sin embargo, tienen residuos cargados positivamente en la región transmembrana que permiten su asociación con moléculas adaptadoras como DAP10 o DAP12 que contienen motivos activadores ITAM (Immunoreceptor Tyrosine-based Activation Motif). Estas moléculas adaptadoras fosforilan su tirosina y se unen a quinasas que interactúan con Vav y Rac1 y activan una cascada que polimeriza la actina y estimula la liberación de citoquinas y gránulos citotóxicos (29,31–34).

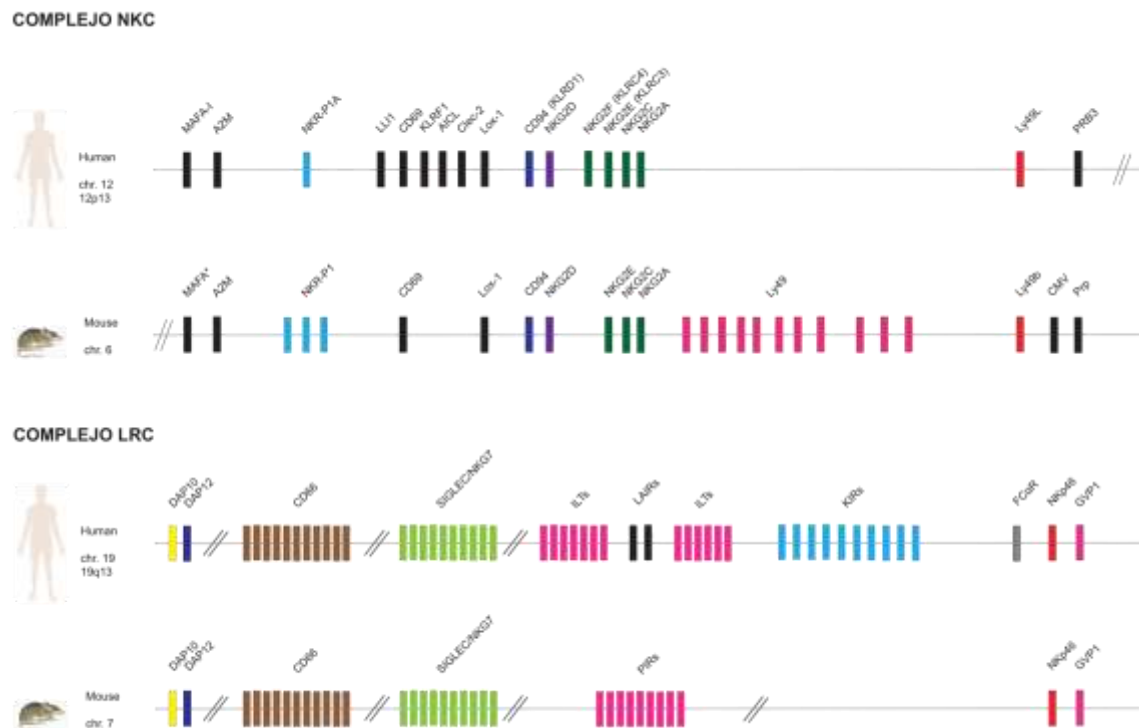
Figura 3-3. Receptores de células NK y sus ligandos del MHC clase I. En verde se muestran los receptores tipo lectinas y en naranja los tipo inmunoglobulina. HLA, MHC humanos.



Evolutivamente los genes KIR han divergido dramáticamente entre diferentes especies de primates, con una rápida expansión específica para cada especie. En primates, la familia de genes KIRs se expandió hace 31-44 millones de años y son más de 17 genes y pseudogenes que están arreglados en tándem ocupando más de 150 kb (8,35). En contraste, la región del complejo LRC humano en el ratón no

tiene genes KIR y su función se ha sustituido por receptores tipo lectina (Ly49) codificadas en el complejo NKC (Figura 3-4) (8,10,12). Los humanos no tienen genes Ly49 funcionales en el complejo NKC, lo que muestra la expansión diferencial de genes KIR en humanos y Ly49 en ratones (8,10,12). Curiosamente, orangutanes y babuinos expresan tanto Ly49 como KIRs, sugiriendo que esta expansión diferencial es linaje-específica (36,37).

Figura 3-4. Organización del complejo NKC y del complejo LRC en humano y ratón. A. Comparación del complejo NKC entre humano y el ratón. En el ratón se observa la expansión de los genes Ly49 y en humano solo se observa el pseudogen Ly49 (rectángulos naranja). Los otros genes conservan una única copia. B. Comparación del complejo LRC entre humano y ratón. En humano hay expansión de los genes KIR mientras que en ratón no existen. Modificado de Barten, 2001.

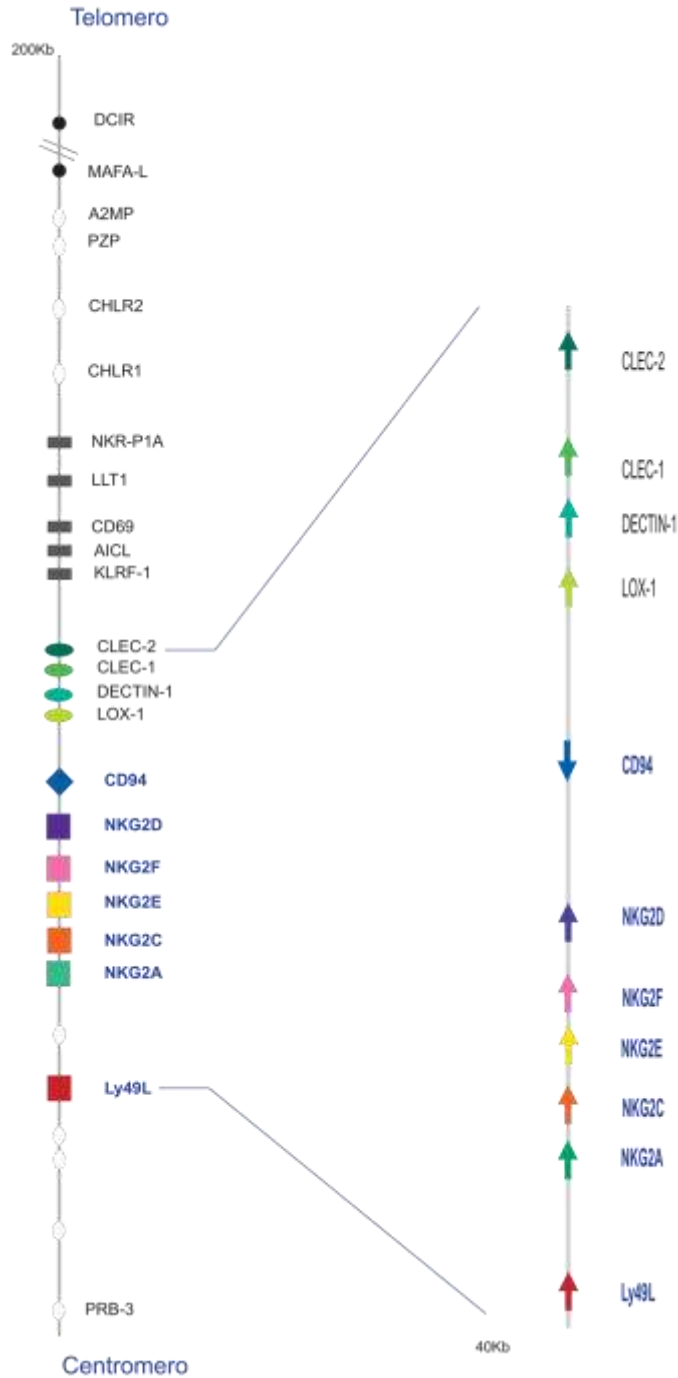


3.4 RECEPTORES DE CÉLULAS NK TIPO LECTINA

Dentro del complejo NKC se encuentran los genes que codifican para los receptores *CD94*, *NKG2*, *Ly49*, *NKR-P1*, *CD69*, *LOX-1*, *Dectin 1*, *CLEC1*, *CLEC2*, *AICL* y *LLT1* (Figura 3-5)(8,38–40). Los genes *CD94*, *NKG2* y *Ly49* son los que codifican para receptores que reconocen a las moléculas del MHC clase I. *NKR-P1* codifica para una proteína que estimula la liberación de gránulos de las células NK, mientras que el gen *CD69* codifica para una lectinas tipo C involucradas en la proliferación linfocitaria funcionando como un receptor señalizador en linfocitos, células NK y plaquetas (41,42). Por otro lado, el gen *LOX-1* codifica para un receptor para la lipoproteína oxidasa de baja densidad en monocitos y células endoteliales (41,43). Los genes *CLEC* codifican para proteínas con diversas funciones, como adhesión celular, señalización, inflamación y respuesta inmune(41,42,44,45). El gen *AICL* parece estar relacionado con la activación de los linfocitos (41,42). El gen *LLT1* se expresa en una subpoblación de células NK y en otras células del sistema inmune(40,46), es el ligando para el receptor NKR-P1A y participa en la secreción de IFN-gamma.(40,42,43,46,47).

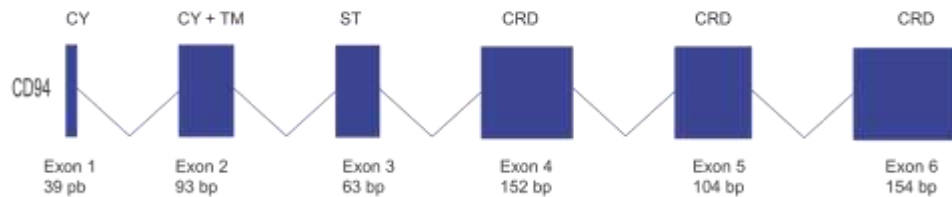
En mamíferos como los primates, los roedores, los carnívoros y los artiodáctilos se ha observado que el contenido de genes en el complejo NKC es heterogéneo con diferentes tasas de generación y muerte de linajes de genes (43,48). Sin embargo, este complejo se ha conservado de manera sinténica en las especies de mamíferos hasta ahora estudiadas presentando la misma orientación, orden y estructura en los cromosomas 12 de humanos, 6 de ratones, 4 de ratas y 1 de gallinas (8,41,42).

Figura 3-5. Organización del complejo NKC en el genoma humano. Cromosoma 12 humano



Los receptores de células NK tipo lectina que interactúan con MHC son proteínas transmembranales que se forman por la dimerización de la molécula CD94 con una de las moléculas NKG2 (A/B, C, E o F) (Figura 3-3). Estas diferentes combinaciones entre CD94 y NKG2 forman receptores tanto de tipo activador como inhibidor(13,49–51). El gen CD94 está conformado por 6 exones que codifican de la siguiente manera: el exón 1 y parte del exón 2 codifican para la región citoplasmática, la parte restante del exón 2 codifica para la región transmembranal, el exón 3 codifica para la región del tallo y los exones 4, 5 y 6 codifican el dominio de reconocimiento de carbohidratos (DRC) (Figura 3-6). Por otro lado, los genes NKG2 están conformados por 7 exones, el exón 2 codifica la región citoplasmática, el exón 3 codifica la región transmembranal, el exón 4 el tallo extracelular y los exones 5, 6 y 7 codifican el dominio DRC. Los análisis de genómica comparativa de CD94/NKG2 entre humanos (38) y otros primates como chimpancés (*Pan troglodytes*) (38), macacos (*Macaca mulatta*) (52), el tití (*Callithrix jacchus*) (14) y el mono ardilla (*Saimiri sciureus*) (13) han mostrado que el complejo CD94/NKG2 es altamente conservado. En el suborden Strepsirrhine ha ocurrido un extenso proceso de duplicación génica de los genes CD94 y NKG2. Las proteínas producidas por estos genes al combinarse producen entre 15 y 24 distintos receptores ampliando la capacidad de reconocimiento (45). La expansión de CD94 y NKG2 se ha visto en lémures (53).

Figura 3-6. Estructura del gen CD94 donde se indica el tamaño y dominio del exón. El diagrama está basado en la descripción del gen CD94 para el tití *Callithrix jacchus* descrito en Averdam A, 2007.



La mayoría de las variaciones se encuentran en el dominio de reconocimiento de carbohidratos generadas por mecanismos de inserción, deleción y sustituciones de un solo nucleótido (SNPs) (13,14,52). CD94 y NKG2 tienen múltiples isoformas generadas por procesamiento alternativo del ARN que incluye la pérdida de exones, el desplazamiento en el marco de lectura del dominio de reconocimiento a carbohidratos y la inclusión de regiones intrónicas producidas por desplazamiento en los sitios de empalme durante el procesamiento del ARN (13,14,52). Algunas isoformas pueden estar bajo presiones selectivas determinadas por alguna función de reconocimiento. Todas estas variaciones sugieren posibilidades sobre la función biológica de las isoformas. Por un lado, las que contienen deleciones transmembranales pueden funcionar a nivel intracelular regulando la expresión de otras isoformas o pueden ser secretadas y actuar asociadas con ligandos. Otra posible función asociada a estas isoformas se puede dar a nivel citoplasmático donde pueden conservar la capacidad de asociarse por interacciones no covalentes y de alguna manera modular la función de formas completas. Es también concebible que simplemente sean vestigios no funcionales del procesamiento del ARN (13,14).

Dentro del espectro de receptores NKC se encuentra el heterodímero inhibidor CD94/NKG2A que se caracteriza por la presencia de dos motivos inhibidores ITIM en el dominio citoplasmático de NKG2A. La correcta interacción del heterodímero con la molécula MHC-E protege a las células blancas de la citotoxicidad mediada por las células NK (13,14,38,54). Los análisis comparativos del gen NKG2A entre diferentes especies de primates muestran que el gen es altamente conservado (55,56). Sin embargo, mecanismos de procesamiento alternativo del ARN donde se pierde el exón 4 que codifica la región del tallo, generan la isoforma NKG2B, la cual ha sido reportada en humanos, macacos y el mono ardilla (13). Esta variante no se ha observado en el marmoset, probablemente por una baja expresión o simplemente porque no se expresa (13). Sin embargo, en esta especie se han observado dos variantes de NKG2A que se generan por cambios en el sitio aceptor del procesamiento alternativo del RNA en el exón 2 que codifica la región citoplasmática y por la pérdida del exón 3 que codifica la región transmembranal respectivamente. Así mismo, en el marmoset se han encontrado polimorfismos no sinónimos en el tallo de la región extracelular y en el dominio de reconocimiento de carbohidratos (14).

El heterodímero CD94/NKG2C es activador y contiene una lisina cargada positivamente en el dominio transmembranal que permite su interacción con la molécula adaptadora DAP12, la cual posee un motivo activador ITAM (13). La comparación de NKG2C entre humanos y chimpancés muestra que la mayoría de sustituciones nucleotídicas ocurren en el dominio DRC, pudiendo afectar la especificidad de unión al ligando y probablemente siendo mantenidas por procesos de selección ejercida por agentes patógenos (38).

En los macacos se identificaron homólogos de NKG2C de humanos y solo se observaron algunas diferencias nucleotídicas y algunas variantes de procesamiento del ARN similares a las observadas entre humanos y chimpancés. Adicionalmente, existe una estrecha relación entre los genes NKG2C y NKG2E (55). La diferencia radica en que la variante NKG2E usa exclusivamente el exón 7 (38). En orangutanes (57) y en el marmoset (14), existe un único gen NKG2CE el cual en su forma basal pudo haberse duplicado en los primates del viejo mundo y haber dado origen a estos dos genes (14,57).

Dentro del complejo NKC se encuentra el gen que codifica para el receptor NKG2D, el cual forma homodímeros de tipo activador (Figura 3-3). Al igual que otras lectinas tipo C, conserva dentro de su estructura 2 láminas β y 2 láminas α . NKG2D está presente en la superficie de las células NK, en algunas células NKT, y en las células T citotóxicas CD8+, en linfocitos T $\gamma\delta$ y bajo algunas condiciones, en las células T CD4+ (58). NKG2D se asocia con el adaptador DAP10 el cual interactúa con la molécula fosfatidilinositol 3 quinasa activando la ruta de señalización (29). NKG2D presenta una amplia diversidad de ligandos, que son moléculas homólogas de MHC clase I, como MICA y MICB, las cuales son codificadas por genes que se encuentran en el MHC y están genéticamente ligados al HLA-B (58). NKG2D también pueden interactuar con las moléculas ULBP1 y ULBP2, las cuales son glicoproteínas que se expresan en la superficie celular y están asociadas con la familia de las ULBP16 una proteína del citomegalovirus. Estos ligandos se expresan a niveles muy bajos en la superficie de las células normales y el incremento en la expresión se induce por procesos infecciosos y tumorales (59–61).

NKG2D es altamente conservado y con muy bajo polimorfismo. Tres alelos se han identificado en humanos y uno solo en chimpancés con un porcentaje de identidad nucleotídica superior al 98% (51). Al igual que en los otros miembros de la familia NKG2 existen isoformas generadas por procesamiento alternativo del ARN. En humanos se han encontrado 6 variantes, de las cuales dos son compartidas con chimpancés (38). La variabilidad se da por la inserción de segmentos intrónicos o por la pérdida de exones. Estos cambios pueden no afectar el transcrito o por el contrario pueden generar un codón prematuro de parada (38). Los ligandos de NKG2D presentan variabilidad tanto a nivel de aminoácidos como de su estructura (62) y se unen con distinta afinidad a su receptor, permitiendo un ajuste gradual en la activación de las células NK (63).

A diferencia de la mayoría de receptores de las células NK, los receptores CD94/NKG2 en humanos interactúan exclusivamente con la molécula HLA-E (38,51,55). Esta asociación molecular entre estos receptores se basa principalmente en la interacción entre CD94 con MHC-E y el péptido (64,65). NKG2 tiene un contacto limitado con una pequeña región de la hélice 2 de MHC-E y con el péptido por medio de una sola interacción de Van Der Waals. Al haber varias isoformas de CD94 y varios receptores de NKG2 (A, B, C y E) se sugiere que puede haber una mayor diversidad de reconocimiento del péptido y una mayor funcionalidad (64,65).

El MHC-E hace parte de los genes MHC clase I no clásicos y se caracteriza por tener un bajo polimorfismo y ser altamente conservado en primates (66,67). Se ha visto que las infecciones virales utilizan mecanismos que interfieren con la expresión de moléculas de MHC clase I generando una baja regulación e interfiriendo con la carga del péptido dependiente de TAP. Sin embargo, se ha observado que los

niveles de expresión de MHC-E no se reducen, sugiriendo que el péptido que presenta el MHC-E proviene de TAP y pudo haber sido seleccionado evolutivamente como una ventaja frente a una infección viral (67).

La molécula del MHC-E, fue previamente caracterizada en algunas especies de primates del Nuevo Mundo (68). En el alineamiento de la Figura 3-7 se observa los sitios de interacción del MHC-E con el complejo CD94/NKG2. Las posiciones pertenecientes a la región alfa 1 interactúan exclusivamente con el dominio DRC de CD94. De esta manera las posiciones 69D, 71Q y 72I de la región alfa 1 del MHC-E interactúan con la posición 114F de CD94. A su vez, la posición 76V del MHC-E interactúa con la posición 160N de CD94. Otras posiciones como 75R y 72I interactúan con 164E de CD94 y finalmente, la posición 72I con la posición 170Q de CD94. MHC-E es muy conservada en sus sitios de interacción con CD94 y los pocos polimorfismos que se encuentran en esta región son vecinos a las posiciones que interactúan en el reconocimiento de MHC-E por CD94. La región alfa 2 interactúa con NKG2 y hay un mayor polimorfismo tanto en el receptor como en el ligando. (14, 36, 58).

Figura 3-7. Sitios de interacción del MHC E con CD94 y NKG2. Las secuencias y el alineamiento fue modificado del trabajo de Knap, 1998. En la parte superior se muestra la región alfa1 que interactúa con CD94; las posiciones marcadas con + rojo son los sitios de interacción con CD94. En la parte inferior se muestra un fragmento de la región alfa2; los asteriscos (*) rojos indican los sitios de interacción con NKG2.

	Alfa 1
	+ ++ +
H. sapiens	ARDTAQIFRVNLR*TLRGYYNQSEA
S. oedipusS.....LR.....
C. jacchusS.....LR.....
A. trivigatusS.....LR.....
P. phiteciaT.....
A. belzebuthS.....LR.....
A. fuscicepsS.....LR.....



3.5 EL COMPLEJO NKC EN PRIMATES DEL NUEVO MUNDO

El tamaño genómico del complejo NKC en el marmoset (*C. jacchus*) es 1.5 veces menor comparado con el humano, es colinear y se mantienen las relaciones de los genes ortólogos CD94, NKG2 y Ly49L. Estos genes son conservados y presentan diferentes transcritos generados por mecanismos de procesamiento alternativo del ARN principalmente en el gen CD94 donde 11 isoformas han sido reportadas (14). En el mono ardilla (*S. sciureus*) también se caracterizaron los genes CD94 y NKG2 identificándose cuatro isoformas para CD94. Una característica particular encontrada en los primates del nuevo mundo es la presencia del gen NKG2CE, el cual puede ser el ancestro de los genes activadores NKG2C y NKG2E encontrados en humanos y los grandes simios (14,57).

4. METODOLOGÍA

4.1 MUESTRAS DE PRIMATES Y EXTRACCIÓN DE ARN

Se recolectaron muestras de sangre periférica de un individuo de las siguientes especies de primates: *Cebus albifrons*, *Cebus apella*, *Cebus capucinus*, *Saguinus leucopus*, *Ateles belzebuth*, *Ateles hybridus*, *Saimiri sciureus*, *Aotus sp*, *Lagothrix lagothricha* y *Alouatta seniculus*. Estas especies han sido mantenidas en la Unidad de Rescate y Rehabilitación de Fauna Silvestre de la Universidad Nacional de Colombia, sede Bogotá, en el Centro de rescate y rehabilitación de la Corporación Autónoma Regional de Norte de Santander Corponor, ubicado en el municipio de Zulia y en el centro de rescate y rehabilitación de la Corporación Autónoma Regional del Alto Magdalena- CAM. Las muestras se tomaron por el método de venipunción de la vena femoral en tubos con EDTA, siguiendo los protocolos de manejo de animales de cada centro. El ARN se obtuvo a partir de sangre total utilizando el reactivo de Trizol (Invitrogen) siguiendo las especificaciones del fabricante. El ARN se almacenó en alícuotas y se ha mantenido a -70 °C.

4.2 AMPLIFICACIÓN DE LOS GENES CD94 Y NKG2

Se sintetizó cDNA a partir del ARN total utilizando el kit de retrotranscripción de Fermentas, siguiendo las especificaciones del fabricante. El cDNA se usó para amplificar por PCR la secuencia completa de los genes CD94 y NKG2D. Para el gen CD94 se amplificaron muestras de las especies *S. sciureus*, *C. albifrons*, *C. apella*, *L. lagothericha*, *A. hybridus* y *Aotus sp.* usando los iniciadores 5'-GTGCCTTCTCTACTTCGCTCTTG-3' y 5'-GTCTTTACTCTCCACCTTCTCTGCC-3' sentido y antisentido respectivamente (13, 14). La mezcla maestra de PCR se preparó con 400 µM de dNTP, 0,5 µM de iniciadores CD94 sentido y CD94 antisentido, 2 mM de MgCl₂ y 0,5 U de Taq DNA polimerasa (Invitrogen) en un volumen de 25 µl. El programa de amplificación consistió en un ciclo de denaturación inicial a 94°C durante 5 minutos, seguido por 35 ciclos a 94°C durante 30 segundos, a 57°C durante 30 segundos y a 72°C durante un minuto y una extensión final a 72°C durante 10 minutos.

El gen NKG2D se amplificó en las especies *C. albifrons*, *C. apella*, *C. capucinus*, *L. Lagothericha*, *Aotus sp.*, *S. leucopus*, *A. seniculus*. Se utilizaron los iniciadores 5'-GTGG/CATTGAAGACTTCAGATTCTTTTCTG-3' y 5'-CAGTGTTTCCGCTGTATAGTCCCTTCATTGTT-3' sentido y antisentido respectivamente (13, 14). La mezcla maestra de PCR se preparó con 200 µM de dNTP, 0,5 µM de los iniciadores, 1,5 mM de MgCl₂ y 1 U de Taq DNA polimerasa (Invitrogen). El programa de amplificación consistió en un ciclo de denaturación inicial a 95°C durante 5 minutos, seguido por 35 ciclos a 94°C durante 30 segundos, a 54°C durante 45 segundos y a 72°C 45 segundos y una extensión final a 72°C durante 10 minutos.

Los productos de PCR para CD94 y NKG2D fueron separados por electroforesis en gel de agarosa al 1.5%, teñidos con Syber Safe a una concentración de 10X y visualizados sobre luz UV. Los productos fueron purificados en una solución al 0.3M Acetato de Sodio (CH_3COONa) y 2.5 volúmenes de etanol absoluto grado biología molecular frío, incubándose por 15 minutos a temperatura ambiente. Luego se centrifugó a 12.000 rpm durante 15 minutos, se lavó tres veces con etanol al 70% y luego se centrifugaron por 3 minutos en centrifuga evaporadora, para finalmente resuspenderse en Agua Ultra pura en un volumen de 15 μl .

4.3 CLONACIÓN Y SECUENCIACIÓN DE PRODUCTOS DE PCR

Los productos de amplificación fueron clonados utilizando el kit TOPO pCR topo 4 (invitrogen) o el sistema pGEM T Easy Vector System (Promega). La transformación se realizó utilizando células quimiocompetentes Top 10 o JM109. Las colonias positivas se seleccionaron creciéndolas en medio LB con antibiótico (ampicilina o kanamicina) y suplementado con Xgal e IPTG. Todos los procedimientos fueron realizados siguiendo las recomendaciones de los fabricantes. Los clones seleccionados fueron repicados en LB con antibiótico y se hizo PCR de colonia con los iniciadores y las condiciones de PCR descritas en los protocolos de amplificación. Los productos fueron separados en geles de agarosa al 1,5% y purificados con el método descrito anteriormente. Las concentraciones se ajustaron de acuerdo a las condiciones solicitadas por el servicio de secuenciación. Al menos 10 clones por individuo fueron secuenciados en Macrogen (Corea), utilizando el método estándar de secuenciación de Sanger.

4.4. BÚSQUEDA DE SECUENCIAS EN BASES DE DATOS

Se realizó una búsqueda de las secuencias disponibles en las bases de datos del GenBank para el gen CD94 en el grupo de primates. La estrategia de búsqueda consistió en hacer Blastn y Blastp utilizando la secuencia de CD94 humana como referencia de búsqueda. Las secuencias encontradas fueron de las siguientes especies: Ratón (*Mus Musculus*), Mumu AF039025 (grupo externo); Lémur Naranja (*Microcebus Murinus*), Mimu FJ869057; Lémur Rufo (*Varecia variegata*), Vava FJ869082.1; Gibón (*Nomascus leucogenys*), Nole XM_003265484.1; humano (*Homo sapiens*), Hosa U30610.1; Chimpancé (*Pan troglodytes*), Patr GU256361.1; Orangután de Sumatra (*Pongo abelli*), Poab XM_002822907.1; Orangután de Borneo (*Pongo pygmaeus*), Popy AF470385.1; Mono Rhesus (*Macaca mulatta*) Mamu, NM_001032828.1; Macaco cangrejero (*Macaca fascicularis*), Mafu AJ585526.1; y marmoset (*Callithrix. jacchus*), Caja EF050438.1.

Se realizó una búsqueda de las secuencias disponibles en las bases de datos del GenBank para el gen NKG2D en el grupo de primates. Las secuencias encontradas fueron de las siguientes especies: Ratón (*M. Musculus*), Mumu 134023659| (grupo externo); Lémur Naranja (*M. Murinus*), Mimu FJ869077.1; Lémur Rufo (*V. variegata*), Vava FJ869096.1; Gibón (*N. leucogenys*), Nole XM_003265487.1; Humano (*H. sapiens*) Hosa AF461811.1, Chimpancé (*P. troglodytes*), Patr AM113545.1; Orangután de Borneo (*P. pygmaeus*), Popy AF470404.1; Orangután de Sumatra (*P. abelli*), Poab XM_002822925.1; Papión de Anubis (*Papio Anubis*), Paan NM_001112645.1; Mono Rhesus (*M. mulatta*), Mamu NM_001032889.1; Macaco cangrejero (*M. fascicularis*), Mafu AJ426429.1; y marmoset (*C. jachuss*), Caja EF050435.1.

4.5. ANÁLISIS DE SECUENCIAS

Las secuencias obtenidas se alinearon utilizando el algoritmo ClustalW (69) y se editaron manualmente mediante el programa BioEdit (70). Se calculó el porcentaje de identidad nucleotídica y de proteínas utilizando el programa Chimera (71). La estimación de divergencia evolutiva para las secuencias nucleotídicas se realizó utilizando el modelo de Kimura 2 parámetros (72), asumiendo una distribución Gamma con un valor de alfa igual a 2. La divergencia evolutiva en las secuencias de aminoácidos se realizó estimando las distancias PAM (73). Estos análisis se hicieron en el programa Mega 5.05 (74).

4.6. ANÁLISIS DE SELECCIÓN

Para el análisis de selección se realizaron los siguientes análisis: a) la prueba de *APARtitioning approach for Robust Inference of Selection* (PARRIS) (75) para selección positiva, la cual por medio de máxima verosimilitud a partir de un modelo nucleotídico obtiene la longitud de las ramas y las tasas de sustitución. Luego calcula la tasa de distribución independiente para ω (dN/dS), donde dN es el número de sustituciones no sinónimas y dS es el número de sustituciones sinónimas y dS permitiendo heterogeneidad tanto en las tasas de sustitución sinónima como no sinónima. b) La prueba de single-likelihood ancestral counting (SLAC) (76), la cual a través de máxima verosimilitud, reconstruye la secuencia del codón ancestral sitio por sitio aumentando la posibilidad de obtener los caracteres de los estados ancestrales. SLAC estima el dN y el dS e indica si el codón es negativamente o positivamente seleccionado. c) La prueba de fixed effects likelihood (FEL) (76) la

cual consiste en estimar inicialmente las tasas independientemente con dos parámetros α (tasa por sitio sinónimo instantáneo) y β (tasa por sitio No sinónimo instantáneo), con la hipótesis nula de $\alpha=\beta$. Para inferir si α es diferente de β utiliza el test de la razón de la verosimilitud y si el valor de p es significativo, el sitio se considera negativo $\alpha>\beta$ o positivo si $\alpha<\beta$. d) Finalmente, se realizó la prueba de random effects likelihood (**REL**) (76) donde se computan dos factores Bayes, el primero donde $\{dN<dS\}$ indica selección negativa, y el otro $\{dN>dS\}$ indica selección positiva. El factor 1/Bayes es análogo al valor p. Estas pruebas se llevaron a cabo usando el servidor datamonkey.org (77)

4.7. RECONSTRUCCIÓN FILOGENÉTICA

Para estimar el modelo de sustitución nucleotídica para cada gen, se utilizó el programa ModelTest (78), evaluando 56 modelos. Para cada uno de estos modelos se determinó un valor de likelihood (Likelihood $-\ln L$). La estimación y selección del modelo de sustitución se realizó usando los criterios LRT ("Likelihood Ratio Test"), AIC (Akaike Information Criterion) y BIC (Bayesian Information criterion).

Se realizaron tres aproximaciones para reconstruir la filogenia de cada uno de los genes, "Neighbor Joining", "Máxima Verosimilitud" e "Inferencia Bayesiana". La reconstrucción filogenética Neighbor Joining (79) se basó en distancias obtenidas con el método de Máxima Verosimilitud Compuesta y asumiendo una distribución gamma con un parametro alfa=2. La topología del árbol se evaluó por bootstrap utilizando 1000 repeticiones (79). Estos parámetros fueron utilizados tanto para CD94 como para NKG2D. Los análisis se realizaron en el programa MEGA 5.05 (74).

Para la reconstrucción filogenética mediante el método de máxima verosimilitud se utilizó para CD94 el modelo de sustitución GTR y se evaluó la topología del árbol mediante bootstrap con 400 repeticiones. Para el gen NKG2D el modelo de sustitución utilizado fue HKY y se realizaron 400 repeticiones de bootstrap. Los análisis se corrieron en el programa Garli Genetic Algorithm for Rapid Likelihood inference (80). En el programa PAUP, se recuperaron los árboles generados y se obtuvo un árbol consenso, se utilizó como regla que la mayoría fuera por encima del 50% para obtener el árbol más confiable (81).

La reconstrucción filogenética por inferencia Bayesiana se llevó a cabo usando el programa MrBayes (82), el cual emplea el método de Monte Carlo de Cadena de Markov (MCMC) para calcular la probabilidad *a posteriori* de la distribución de árboles. Para CD94 se utilizó el modelo de sustitución GTR + G y para NKG2D el modelo de sustitución fue HKY + G. Se realizaron 100.000 iteraciones, haciendo muestreos cada 100 iteraciones. Las primeras 1000 muestras del árbol fueron descartadas como "Burn in".

4.8 ESTIMACIÓN DE TIEMPOS DE DIVERGENCIA CON INFERENCIA BAYESIANA

Los tiempos de divergencia de los genes fueron calculados haciendo análisis bayesianos usando un reloj molecular relajado con el programa BEAST (83). Se asumieron tasas de sustituciones independientes en las diferentes ramas del árbol. El reloj molecular se calibró utilizando los tiempos de divergencia entre *H. sapiens* y *P. troglodytes* (6 millones de años) como el grupo con ancestro común más reciente y otros puntos de calibración fueron utilizados como *H. sapiens* y *V. variegata* (74 millones de años) y *H. sapiens* y *C. albifrons* (42 millones de años). Los tiempos de divergencia fueron tomados en la página de "The TimeTree of Life" <http://www.timetree.org/book.php>. un servicio público de información sobre tiempos evolutivos de la vida que realiza una búsqueda de los tiempos de divergencia de los organismos en la literatura publicada (84).

4.9 PREDICCIÓN Y MODELAMIENTO DE PROTEÍNAS

La estructura de las proteínas se modeló por homología utilizando como plantilla la estructura de CD94 (1B6E) y NKG2D (1MPU) humanos utilizando la Suite de Swiss model (85). La estimación de la calidad del modelo se realizó con la misma suite mediante el algoritmo de QMEAN score (86) y a su vez se validaron mediante el test de Ramachandran. El análisis de plegamiento de las proteínas de las estructuras obtenidas se validó mediante el servidor Phyre (87). Los análisis de sobreposición y edición de las imágenes se realizaron el programa Chimera (71) y en VMD (88).

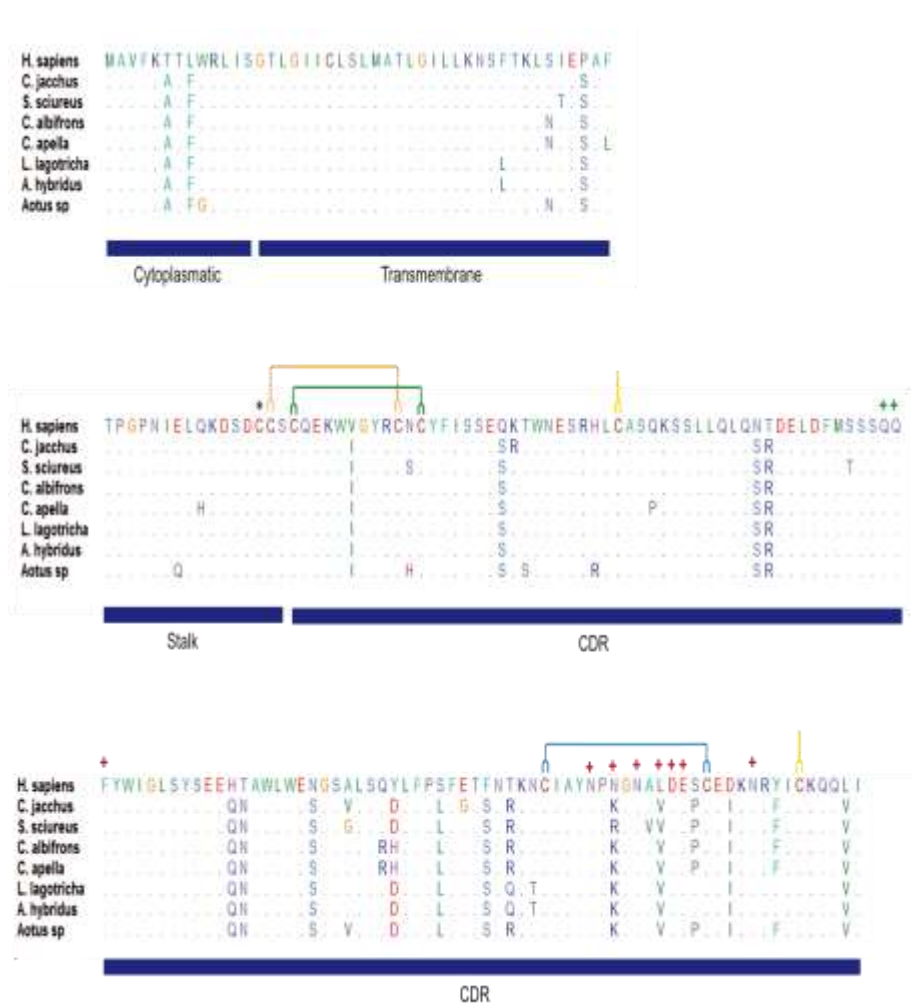
5. RESULTADOS

5.1 ALTA CONSERVACIÓN DE CD94 EN PRIMATES DEL NUEVO MUNDO

El gen CD94 se clonó y se secuenció en las especies de primates del nuevo mundo *S. sciureus*, *C. albifrons*, *C. apella*, *L. lagothericha*, *A. hybridus* y *Aotus sp.* En las especies analizadas se encontró que el gen CD94 es de copia única. La organización génica y el tamaño de los exones es igual a aquellas de las especies de primates previamente estudiadas. Las secuencias de cDNA fueron traducidas generando una proteína predicha de 179 aminoácidos, correspondientes a la región citoplasmática y transmembranal de 44 aminoácidos, el tallo de 21 aminoácidos y el dominio DRC de 114 aminoácidos. Este último corresponde a un dominio lectinas tipo C (CLEC), según el análisis con la base de datos Pfam (89). Como se observa en la Figura 8, en el dominio DRC se conservan las cisteínas que forman enlaces disulfuro característicos de las lectinas tipo C (Cys 59-Cys70, Cys61-Cys72, Cys89-Cys174 y Cys152-Cys176). Adicionalmente, los CD94 de primates neotropicales conservan la cisteína del tallo que permite la dimerización con NKG2 mediante el establecimiento de un enlace disulfuro. Finalmente, en el dominio DRC se encuentran conservadas todas las posiciones que interactúan con MHC-E (Figura 5-8).

El gen CD94 está altamente conservado en primates del Nuevo Mundo. En promedio, la identidad nucleotídica de CD94 de las especies estudiadas con su homólogo humano fue 90,66% mientras que la identidad a nivel de proteínas fue 87,07% (Tabla 5-1).

Figura 5-8. Alineamiento de CD94 en primates del Nuevo Mundo. Las regiones citoplasmática, transmembranal, el tallo y el dominio de reconocimiento a carbohidrato CDR se señalan con una barra de color azul en la parte inferior. Los puntos indican identidad de secuencia. En el tallo se indica con el asterisco (*) la cisteína implicada en la interacción con las proteínas NKG2. Los números sobre las cisteínas indican los residuos involucrados en un puente disulfuro dado. Las cruces (+) rojas señalan los residuos que interactúan con MHC directamente con la molécula del MHC E. Las cruces verdes indican los residuos que interactúan con el péptido.



Adicionalmente la identidad nucleotídica promedio de CD94 de los primates del Nuevo Mundo con homólogos de otras especies de primates fue de 90,66% con el del chimpancé, 91.03% con las dos especies de orangutanes y el Gibón, 90,20% con las dos especies de macacos y 76,88% con los lémures. Como es de esperarse, las distancias genéticas basadas en secuencias de CD94 son muy cercanas entre los primates, siendo en muchos casos muy cercanos a cero, como aquella entre las especies de *L. lagothericha* y *A. hybridus* (Tabla 5-2). Aunque las distancias son similares en comparaciones por dominios, el dominio DRC tiende a tener valores más altos que aquellos de los otros dominios (Tabla 5-2).

Tabla 5-1. Porcentaje de identidad nucleotídica y de proteínas de CD94 en primates del Nuevo Mundo.

	<i>Caja</i> ¹		<i>Sasc</i>		<i>Ceal</i>		<i>Ceap</i>		<i>Lala</i>		<i>Athy</i>		<i>Aosp</i>	
	<i>Nt</i> ²	<i>aa</i> ³	<i>nt</i>	<i>aa</i>	<i>Nt</i>	<i>aa</i>	<i>nt</i>	<i>aa</i>	<i>nt</i>	<i>aa</i>	<i>nt</i>	<i>aa</i>	<i>Nt</i>	<i>aa</i>
<i>Hosa</i>	90.13	87.15	90.13	86.03	90.69	87.71	89.18	86.03	91.99	88.83	92.18	88.83	90.32	84.92
<i>Caja</i>			96.09	95.93	96.28	96.65	95.12	94.97	96.09	95.53	96.28	95.53	96.83	95.53
<i>Sasc</i>					95.90	94.97	94.78	93.30	95.52	93.85	95.90	93.85	96.28	93.85
<i>Ceal</i>							98.51	98.32	95.72	95.53	95.90	95.53	96.46	95.53
<i>Ceap</i>									94.59	93.85	94.78	93.85	95.34	93.85
<i>Lala</i>											98.81	100	96.28	93.30
<i>Athy</i>													96.46	93.30
<i>Aosp</i>														

1. *Hosa*, *H. sapiens*; *Caja*, *C. jachuss*; *Sasc*, *S. sciureus*; *Ceal*, *C. albifrons*; *Ceap*, *C. apella*; *Lala*, *L. lagothericha*; *Athy*, *A. hybridus* y *Aosp* *Aotus* sp.
2. Nucleótidos. 3. Aminoácidos.

Tabla 5-2. Distancias evolutivas de CD94 en primates del Nuevo Mundo

	Caja ²			Sasc			Ceal			Ceap			Lala			Athy			Aosp		
	Com ³	Tm ⁴	DRC ⁵	Com	Tm	DRC	Com	Tm	DRC	Com	Tm	DRC	Com	Tm	DRC	Com	Tm	DRC	Com	Tm	DRC
Caja				0.045	0.017	0.060	0.034	0.017	0.043	0.051	0.051	0.052	0.045	0.017	0.059	0.045	0.017	0.059	0.050	0.068	0.042
Sasc	0.041	0.047	0.038				0.051	0.034	0.060	0.069	0.069	0.069	0.062	0.034	0.077	0.062	0.034	0.077	0.069	0.086	0.061
Ceal	0.039	0.041	0.038	0.043	0.041	0.044				0.017	0.034	0.008	0.045	0.034	0.051	0.045	0.034	0.051	0.051	0.050	0.051
Ceap	0.052	0.073	0.041	0.055	0.073	0.047	0.015	0.029	0.008				0.063	0.069	0.060	0.063	0.069	0.060	0.069	0.086	0.060
Lala	0.041	0.029	0.048	0.045	0.041	0.047	0.045	0.035	0.050	0.058	0.066	0.053				0.000	0.000	0.000	0.074	0.086	0.068
Athy	0.039	0.029	0.044	0.043	0.041	0.044	0.043	0.035	0.047	0.055	0.066	0.050	0.002	0.000	0.003				0.074	0.086	0.068
Aosp	0.033	0.041	0.029	0.039	0.053	0.032	0.037	0.035	0.038	0.049	0.066	0.041	0.039	0.035	0.041	0.037	0.035	0.038			

1. Debajo de la diagonal en azul se muestran la tasa de sustituciones nucleotídicas basadas en el método de Kimura 2 parámetros (72). En la parte superior en verde se muestran las tasas de sustituciones aminoacídicas con matrices (73).
2. Secuencia completa, 3. Com: regiones citoplasmáticas, transmembranal y el tallo, 4. TM; dominio de reconocimiento a carbohidratos. 5. DRC.

5.2. SELECCIÓN PURIFICADORA EXPLICA LA ALTA CONSERVACIÓN DE CD94 EN PRIMATES DEL NUEVO MUNDO

Los análisis para identificar selección positiva por máxima verosimilitud utilizando las pruebas de PARRIS, SLAC, y FEL y la prueba de REL por inferencia Bayesiana no mostraron evidencia de selección positiva para ninguna posición. Sin embargo, los análisis mostraron que 6 posiciones son seleccionadas negativamente en al menos dos pruebas de selección (Tabla 5-3). El codón 11 codifica una leucina que se ubica en la región citoplasmática y los codones 26 y 31 codifican una treonina y una leucina respectivamente que se encuentran en la región transmembranal. Aunque no se puede hacer una predicción precisa de la función de estas posiciones, es posible que sean importantes en el mantenimiento de la estructura de la molécula. En el dominio DRC tres posiciones son seleccionadas negativamente, Ser121, Tyr122 y Tyr155. Para determinar si estas posiciones interactúan con MHC-E se hizo predicción de interacción proteína-proteína en la Suite del Meta-PPISP (90) Este análisis mostró que los sitios no hacen parte de la interacción con MHC-E. Sin embargo, el residuo el residuo Tyr155 hace parte del loop 6 que va de la posición 156 a la 159 y que se encuentra en el medio de las láminas $\beta 5$ y $\beta 6$. El loop 6 entra en contacto con el loop 3, gracias a la formación de puentes de hidrógeno donde participan la Ala156, la Ans158 y con la Pro157. Debido a la estructura química de la tirosina, esta puede ser importante en la estructura del esqueleto del loop 6 y de esta forma facilitar la interacción de las posiciones vecinas.

Tabla 5-3. Codones seleccionados negativamente en al menos dos pruebas de selección para CD94.

Codón	Símbolo	Aminoácido	Región
11	L	Leucina	Citoplasma
26	T	Treonina	Transmembrana
31	L	Leucina	Transmembrana
121	Y	Tirosina	DRC
122	S	Serina	DRC
155	Y	Tirosina	DRC

5.3 ANÁLISIS FILOGENÉTICO DE CD94

5.3.1. La filogenia de CD94 refleja la filogenia de las especies

La reconstrucción filogenética del gen CD94 se hizo por el método de distancias Neighbor Joining (NJ), Máxima Verosimilitud (ML) e Inferencia Bayesiana. La topología reconstruida por los tres métodos es convergente y muestra dos clados que representan los subordenes de los primates. Por un lado están los Strepsirrhini que son primates que habitan principalmente en Madagascar y por el otro lado los Haplorrhini donde se encuentran los monos del Viejo y del Nuevo Mundo, los grandes simios, los gibones y los humanos (Figura 9). La probabilidad *a posteriori* que soporta estos clados es 1. El grupo Haplorrhini se bifurca formando los Catarrhini y los Platyrrhini. En los Catarrhini los valores *a posteriori* de soporte para

todas las ramas es 1. En el nodo donde se agrupa el Gibón (Nole) con el chimpancé (Patr) y el hombre (Hosa) los valores de bootstrap para NJ y ML son de 37% y 57%, respectivamente, indicando un bajo soporte para su agrupación (Figura 5-9).

En el clado de los Platyrrhini se identifican claramente dos agrupaciones con alto soporte. Por un lado están las especies de la familia Atelidae, *L. laghotricha* y *A. hybridus*, y por el otro lado las cinco especies de la familia Cebidae y Aotidae. La topología con los tres métodos es la misma aunque los valores de soporte de bootstrap y probabilidad *a posteriori* son relativamente bajos a excepción de aquellos que soportan la agrupación de las dos especies de *Cebus* (Figura 5-9). La reconstrucción filogenética del gen CD94 para los primates refleja en general la filogenia de las especies aunque no hay una completa resolución de la agrupación del gibón con la familia Hominidae.

5.3.2 Estimación de tiempos de divergencia de CD94 en primates.

Se calcularon los tiempos de divergencia de CD94 mediante el método de inferencia bayesiana (Figura 5-10). La filogenia del gen refleja la filogenia de las especies y los tiempos de divergencia del gen concuerdan en general con los tiempos de divergencia de las especies estimados por métodos moleculares y morfológicos (91), Por ejemplo, de acuerdo con el registro fósil Strepsirrhini y Haplorrhini divergieron hace 78 millones de años y el tiempo de divergencia de CD94 entre estos grupos es de 73.73 millones de años. Al igual ocurre entre Catarrhini y Platyrrhini que divergieron hace 43.5 millones de años y las estimaciones basadas en CD94 resultaron en 37.7 millones de años. La divergencia entre Atelidae y

Cebidae ocurrió aproximadamente hace 20 millones de años (91) y la divergencia de CD94 entre estas dos familias se estimó en 18.03 millones de años. Así mismo, el registro fósil muestra que la divergencia entre Cebinae y Saimirinae ocurrió hace unos 17 millones de años, pero la estimación basada en CD94 es de tan solo 12 millones de años. En general, los tiempos de divergencia basados en CD94 tienden a subestimar los tiempos inferidos con otras moléculas o por el registro fósil. Esta diferencia probablemente se debe a que el gen CD94 está sujeto a una fuerte presión selectiva purificadora.

Figura 5-9. Árbol filogenético del gen CD94 en primates. La reconstrucción filogenética se hizo con NJ, ML e inferencia bayesiana con una topología convergente con los tres métodos. Los porcentajes de bootstrap se muestran en rojo para NJ y verde para ML. La probabilidad a posteriori se muestra en azul. La zona sombreada muestra el suborden Platyrrhini. *M. musculus* Ratón, *M. murinus* Lémur Naranja, *V. variegata* Lémur Rufo, *N. leucogenys* Gibón, *H. sapiens* Humano, *P. troglodytes* Chimpancé, *P. pygmaeus* Orangután de Borneo, *P. abelli* Orangután de Sumatra, *M. mulatta* Mono Rhesus, *M. fascicularis* Macaco cangrejero, *C. jacchus*, *S. sciureus*, *C. albifrons*, Ceap *C. apella*, Lala *L. lagotrucha*, Athy *A. hybridus*, Aosp *Aotus sp.*

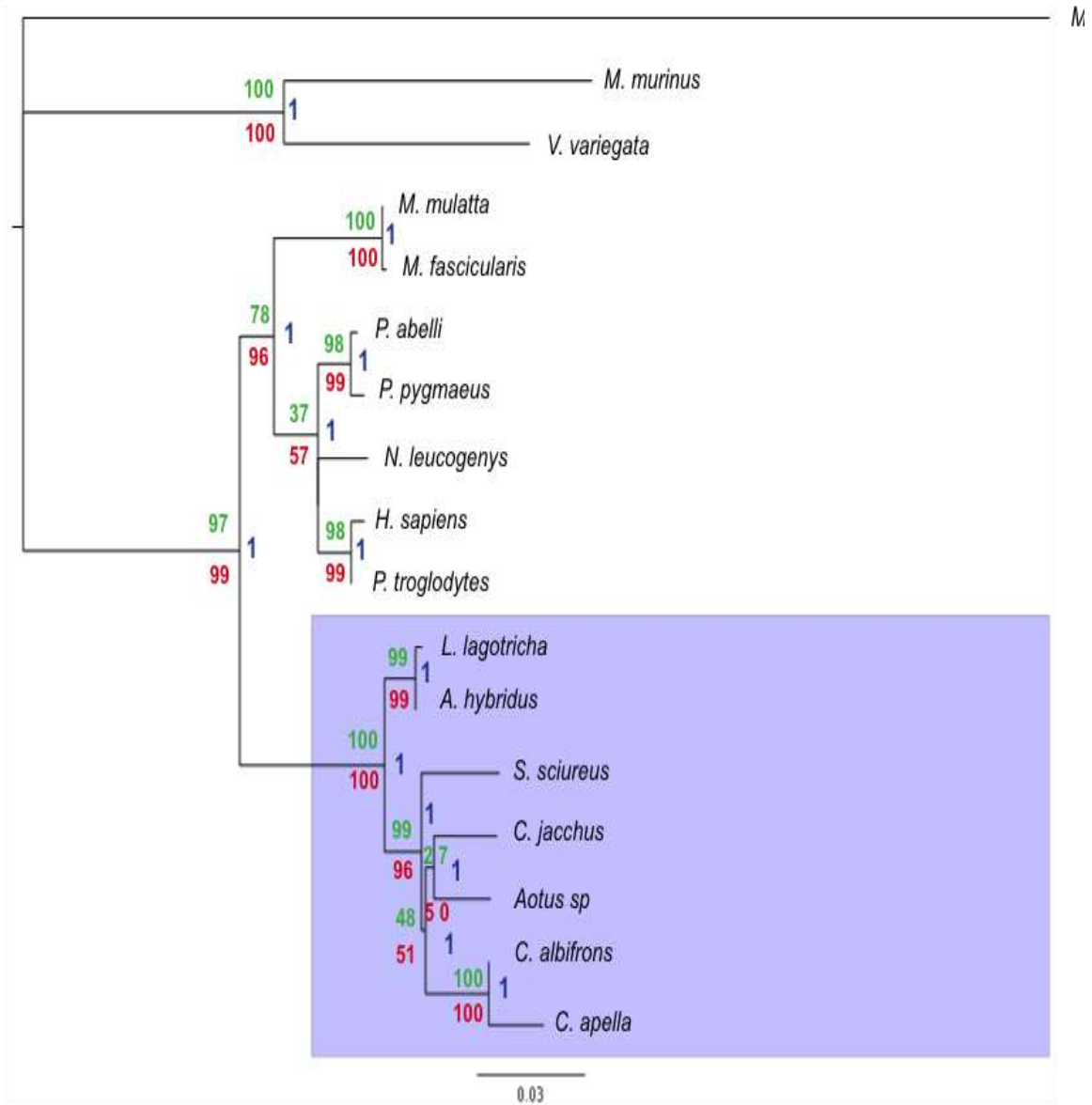
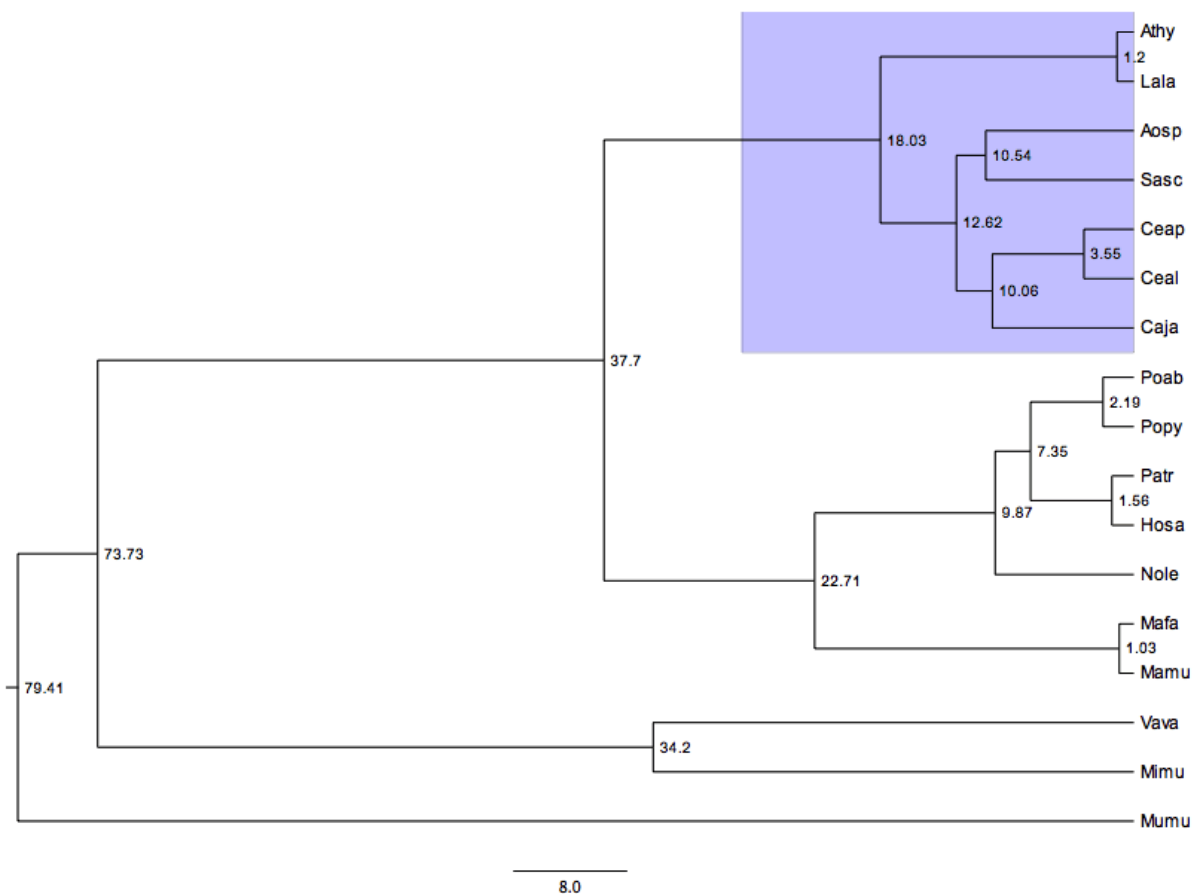


Figura 5-10. Tiempos de divergencia del gen CD94 estimados por inferencia Bayesiana en primates. El reloj molecular se calibró el reloj con los tiempos de divergencia entre humanos y chimpancé (6 millones de años), humanos y el lémur (74 millones de años) y entre humanos y *C. apella* (42 millones de años). El área resaltada en lila corresponde Platyrrhini. Los valores en los nodos corresponden a millones de años. *Mumu M. musculus* Ratón, *Mimu M. murinus* Lémur Naranja, *Vava V. variegata* Lémur Rufo Nole *N. leucogenys* Gibón, *Hosa H. sapiens* Humano, *Patr P. troglodytes* Chimpancé, *Popy P. pygmeus* Orangután de Borneo, *Poab P. abelli* Orangután de Sumatra Paan Papión de Anubis, *Mamu M. mulatta* Mono Rhesus, *Mafa M. fascicularis* Macaco cangrejero Caja *C. jacchus*. *Sasc S. sciureus*, *Ceal C. albifrons*, *Ceap C. apella*, *Lala L. lagotherica*, *Athy A. hybridus*, *Aosp Aotus sp.*



5.3.3 Diversificación de CD94 por procesamiento alternativo del ARN

Catorce isoformas distintas de CD94 generadas por procesamiento alternativo del ARN se encontraron en los primates del Nuevo Mundo estudiados (Figura 5-11). Nueve de estas variantes ya fueron reportadas en *C. jacchus*, *S. sciureus* y *H. Sapiens*. La isoforma CD94B se produce por el cambio de sitio aceptor de procesamiento del ARN en el extremo 5' del exón 5 generando la inclusión de tres pares de bases del extremo 3' del intrón 4. Esta inserción genera un nuevo codón (CAG) en el transcrito incorporando una glutamina en el dominio RCD en una región distante al sitio de interacción con el MHC-E, probablemente sin alterar su función de esta variable. Esta isoforma se encontró en todas las especies estudiadas, excepto en *Aotus sp* (Figura 5-11).

La isoforma SV1[^] fue encontrada previamente en *S. sciureus* y se caracteriza por la inserción del tetranucleotido GCAT en el final 3' del exón 4 proveniente del final 5' del intrón 4. La inserción cambia el marco de lectura generando un codón de parada prematuro en el exón 5 y por tanto un dominio DRC truncado. La isoforma SV2 pierde el exón 3 que codifica para el tallo y conserva el marco de lectura, mientras que la isoforma SV2[^] presenta la inserción GCAT en el exón 4 vista en SV1[^] y además la inserción CAG en el exón 5 observada en la isoforma CD94B. Esta isoforma tiene un codón de parada prematuro en el exón 5 debido a un corrimiento en el marco de lectura. La isoforma SV3 pierde el exón 3 que codifica el tallo y adicionalmente inserta el codón CAG (glutamina) en el extremo 5' del exón 5, conservando intacto el marco de lectura.

Las isoformas SV4 y SV5 pierden el exón 4 el cual codifica el primer segmento del dominio RCD. La única diferencia entre estas dos isoformas es la inclusión del codón CAG en el exón 5 de SV5. Estas variantes tienen un corrimiento en el marco de lectura y generan un codón de parada prematuro en el exón 5, dejando un dominio DRC de tan solo 47 aminoácidos. Además, estas variantes substituyen el residuo Cys58 el cual está involucrado en la formación del enlace disulfuro entre CD94 y los receptores NKG2. En las isoformas SV6 y SV7, el exón 4 pierde 61 pb debido a la presencia de un sitio críptico de procesamiento del ARN dentro del exón. Estas dos isoformas difieren en la inclusión del codón CAG en el exón 5' de SV7 pero ambas tienen un codón prematuro de parada en el exón 5 debido a un corrimiento en el marco de lectura.

Cinco nuevas isoformas de CD94 fueron encontradas en este estudio. La isoforma NSV12 fue hallada en *S. sciureus*, se caracteriza por el desplazamiento del sitio aceptor de procesamiento del ARN dentro del intrón 5 (CTAG), el desplazamiento genera la adición de 302 pb en la región 5' del exón 6. Esta inserción genera un cambio en el marco de lectura y crea un codón de parada prematuro, acortando el dominio RCD de la proteína hipotética. La isoforma NSV13 fue encontrada en *L. lagothericha* y se caracteriza por un cambio en el sitio críptico aceptor de procesamiento del ARN en el exón 4, generando una inserción de un segmento de 117 pb del intrón 5 y un segundo cambio en el sitio aceptor que adiciona un codón (CAG) en el extremo 5' del exón 5. Estas inserciones producen un transcrito de mayor longitud con un codón de parada prematuro en el exón 6. Otra nueva isoforma fue encontrada en esta especie, NSV14, la cual pierde el exón 4, se inserta el codón CAG en el extremo 5' del exón 5 y muestra un sitio críptico de procesamiento del ARN que acorta el exón 5 dejándolo de 36 pb.

Las isoformas NSV15 y NSV16 fueron encontradas únicamente en *A. hybridus*. NSV15 pierde los exones 1 y 2 e inserta el codón CAG en el extremo 5' del exón 5. El marco de lectura se conserva y las características de este transcrito son similares a las reportadas para la isoforma CD94 Δ tm encontrada en humanos y en *S. sciureus* donde la región transmembranal se pierde y se conserva el tallo y el dominio RCD (36, 43). La isoforma NSV16, es muy similar a las dos isoformas mencionadas anteriormente ya que conserva el primer exón al igual que CD94 Δ tm pero pierde la mayoría del exón 2 conservando solo las últimas 39 pb.

El procesamiento alternativo del ARN en CD94 es particularmente interesante por dos razones. La primera, es el alto número de isoformas que se pueden encontrar en una especie e incluso en un mismo individuo. Como se observa en la Figura 5-11, especies como *C. jacchus*, y *L. lagothericha* presentan hasta 7 isoformas, aunque en las demás especies el número es menor. La única excepción a la diversidad de isoformas es en *Aotus* sp donde solo se encontró la forma completa de CD94. La segunda razón, es que algunas de estas isoformas se expresan en diferentes especies que divergieron hace decenas de millones de años. La isoforma CD94B es el ejemplo más claro, ya que se ha identificado en seis de las siete especies de primates del Nuevo Mundo estudiadas hasta la fecha. Otras isoformas como SV4, SV5 y SV7 se encuentran en tres o más especies de primates del Nuevo Mundo.

La amplia gama de isoformas de CD94 en los primates del Nuevo Mundo y su conservación entre distintas especies sugiere un proceso de diversificación funcional de este gen por mecanismos epigenéticos. Este gen que es de copia única y tiene limitada variabilidad, podría estar creando diversas formas proteicas con distintas

especificidades de unión a ligandos a través del procesamiento alternativo del ARN. La presencia de las mismas isoformas en varias especies sugiere presiones selectivas para su conservación, favoreciendo la posibilidad de que algunas de estas isoformas sean funcionales. Alternativamente, algunas de estas isoformas pueden ser vestigios evolutivos de un mecanismo de procesamiento alternativo del ARN sin que necesariamente cumplan una función.

5.3.4. Predicción y modelamiento de la proteína CD94 y sus isoformas

La estructura de las proteínas CD94 predichas para los primates del Nuevo Mundo se modelaron utilizando como plantilla la estructura cristalografiada del dominio DRC de la proteína humana (92). La predicción se hizo por homología utilizando la Suite de Swiss model (85) y se observa que las estructuras obtenidas conservan el mismo plegamiento con las 7 láminas β antiparalelas y la hélice α , características del dominio tipo C (Figuras 5-12 y 5-13).

Figura 5-11. Isoformas de CD94 en primates del nuevo mundo. En azul se observa la estructura del gen. En Gris oscuro se encuentran las isoformas previamente reportadas encotradas también es este estudio. Las variantes en gris claro son variantes reportadas previamente no encontradas en éste estudio. Las nuevas variantes están en negro. A la derecha se observa la presencia de las variantes en los primates estudiados

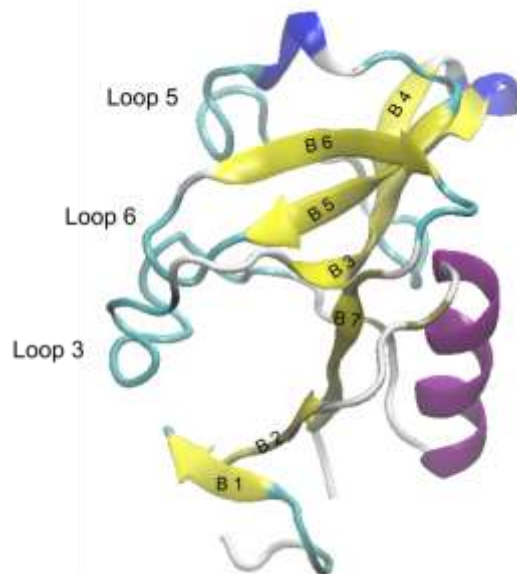
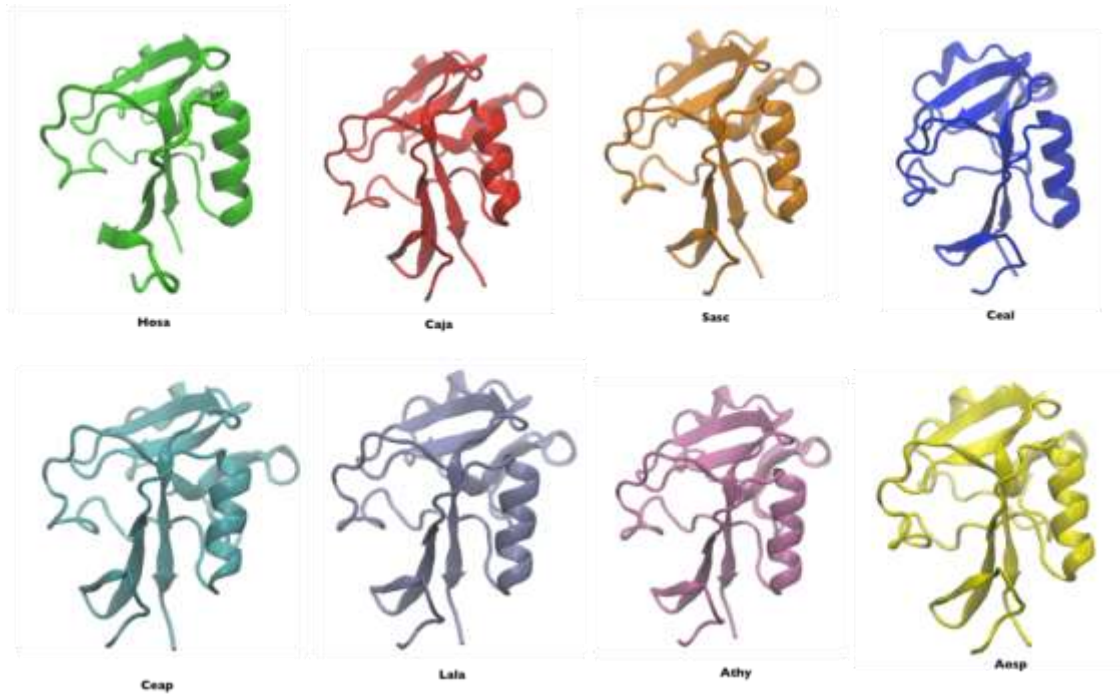
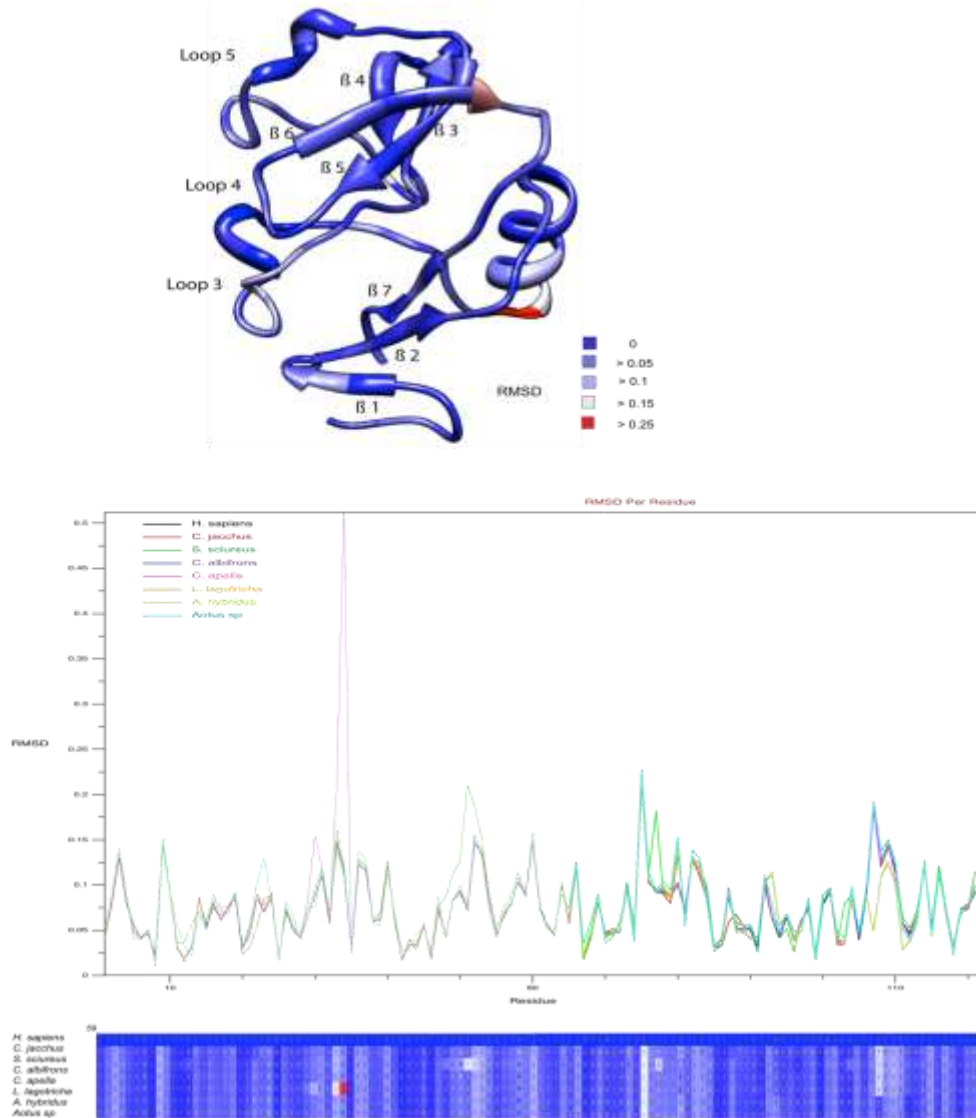


Figura 5-13. Estructuras predichas de CD94 en primates del Nuevo Mundo. Las estructuras se obtuvieron a partir de la estructura con cristalografía de CD94 humano (PDB 1B6E). Utilizando la suite de *Swiss Model*. Caja *C. jacchus*, Sasc *S. sciureus*, Ceal *C. albifrons*, Ceap *C. apella*, Lala *L. lagotherica*, Athy *A. hybridus*, Aosp *Aotus sp.*



Para comparar las estructuras de CD94 de las especies estudiadas se hizo sobreposición de las mismas y se evaluaron las regiones conservadas y variables. También se calculó el *root-mean square deviation* (RMSD) para cada estructura. El análisis indica que las proteínas se conservan tanto a nivel de secuencia como a nivel de plegamientos y de estructura. Se destaca que en *C. apella* hay un pico cerca al residuo 90 del dominio RCD que indica que hay un cambio conformacional con respecto a las otras proteínas homólogas. En esta región no hay sitios de interacción con MHC-E (Figura 5-14).

Figura 5-14. Sobreposición de las estructuras de CD94 en primates del Nuevo Mundo. Arriba Estructuras sobrepuestas de CD94 de humanos con las 7 especies de primates del Nuevo Mundo. Abajo. Plot de cálculo de RMSD por residuos del dominio DRC y el alineamiento estructural generado a partir de la sobreposición de la estructura; el azul fuerte representa las regiones más conservadas; el azul más claro muestra regiones conservadas con pequeños cambios en algunas especies y el rojo muestra diferencias a nivel de secuencia y de estructura.



La isoforma CD94B de *S. sciureus* (inserción Q106) fue sobrepuesta sobre la proteína CD94 humana (Figura 5-15). La glutamina que caracteriza a CD94B genera un plegamiento y un cambio conformacional del loop 3. Este loop es una característica única de CD94 que lo diferencia estructuralmente de otras lectinas tipo C (Figura 5-15A). Para tratar de establecer las implicaciones de esta inserción en las interacciones de la molécula se comparó el complejo CD94/NKG2A con MHC-E (figura 5-15B) con la molécula CD94B/NKG2A con MHC-E (utilizando la estructura CD94B de *S. sciureus*) (Figura 5-15C). La sobreposición indica que la inserción de la glutamina no está dentro de la región de interacción con MHC-E ni con NKG2A. Esta interacción en el loop 3 afecta el centro hidrofóbico de la molécula del cual hace parte el loop 3, el cual está estabilizado por varios residuos no polares (Leu105, Phe107, Met108 y Tyr115), dos puentes básicos (Glu104-Lys175 y Asp106-Arg69) y varios puentes de hidrógeno. Adicionalmente, el loop 3 entra en contacto con las láminas $\beta 5$ y $\beta 6$ de la misma molécula, haciendo contacto a través de puentes de hidrógeno entre las cadenas Asn156 y Asn158 e interacciones no polares con la Pro157.

La isoforma SV1[^] también fue modelada por homología utilizando como plantilla la estructura completa de CD94 humano utilizando la Suite de Swiss model (85). Luego se comparó el modelo obtenido con el complejo CD94/NKG2A MHC-E mediante sobreposición de la estructura. La isoforma SV1[^] conserva las láminas $\beta 1$ y $\beta 2$ y la hélice α , mientras que las otras laminas son remplazadas por un loop corto. Se observa, que se pierde completamente el loop que interactúa con MHC-E. Igualmente, la interacción con todo el complejo no se ve favorecida lo que sugiere que posiblemente esta isoforma no sea funcional en este complejo molecular (Figura 5-16).

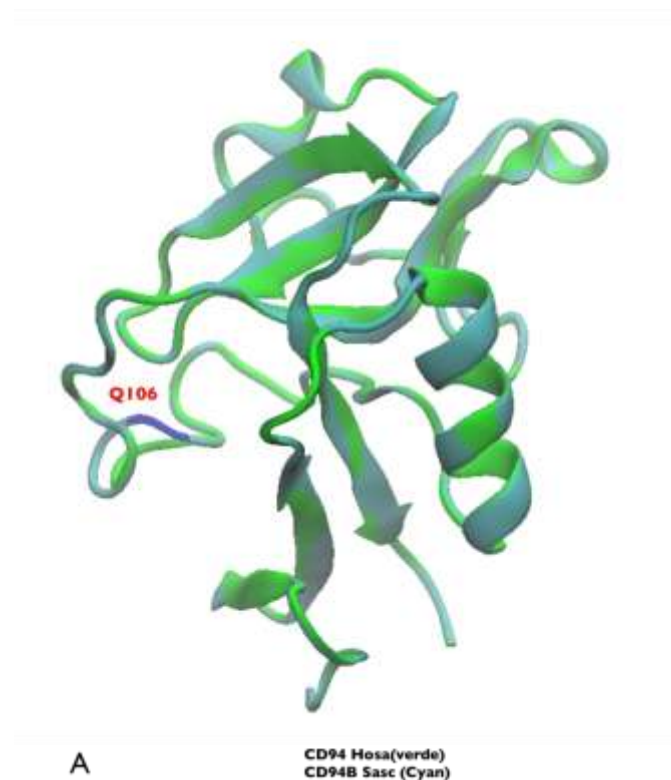
Con la isoforma SV1[^] se realizó un análisis en el servidor Phyre, cuyo algoritmo permite reconocer los posibles plegamientos y la estructura de la proteína (8). El plegamiento de la isoforma dio como resultado un dominio lectina tipo C con precisión estimada del 95%. Las moléculas con las cuales muestra una similitud a nivel estructural son LOX-1, una lipoproteína oxidasa de baja densidad y los receptores DC-SIGN, CLEC-2, y Ly49A de ratón. Estas proteínas hacen parte del complejo NKC y pueden estar relacionadas evolutivamente. También hay similitud con el receptor FC epsilon de baja afinidad de inmunoglobulinas y con la cadena beta de la coaglutinina en diferentes especies de serpientes. Estos resultados sugieren un posible rol de esta isoforma en otros contextos celulares. Igualmente, este análisis sugiere que la isoforma SV2[^] puede tener funciones iguales a las propuestas para SV1[^].

La isoforma SV2 pierde la región del tallo donde se encuentra la cisteína que forma el enlace disulfuro que permite su interacción con las moléculas NKG2A o NKG2C en humanos. Sin embargo, conserva las regiones citoplasmática y transmembranal, lo que sugiere que la proteína se puede anclar a la membrana. La cristalografía de CD94 corresponde a la región extracelular, por tal razón el modelo obtenido para SV2 es el mismo que para CD94. La isoforma SV3 presenta las mismas características de SV2 con la excepción de la inserción de la tripleta CAG, la cual genera la modificación observada en CD94B.

Las isoformas SV4, y SV5 pierden el primer exón que codifica para el dominio DRC, esta delección genera un codón de parada prematuro que produce un péptido con un dominio extracelular muy corto, por lo cual no puede ser modelado. Lo más probable es que estas isoformas no sean funcionales. Las isoformas SV6 y SV7

pierden parte del exón 4 que codifica la primera parte del dominio DRC. Esta pérdida cambia el marco de lectura del transcrito y el péptido que se genera es corto (Figura 5-17).

Figura 5-15. Análisis por modelamiento de la isoforma CD94B. A. Sobreposición de CD94 de humanos (Hosa) versus CD94B de *S. sciureus* Sasc. Se observa que en la región correspondiente al loop 3, la inserción de una Glutamina (Q106) generando una torsión del loop 3. B. Interacción de CD94 con todo el complejo (NKG2 y MHCE). C Superposición de CD94B de Sasc en el complejo CD94B/NKG2A y MHC E.



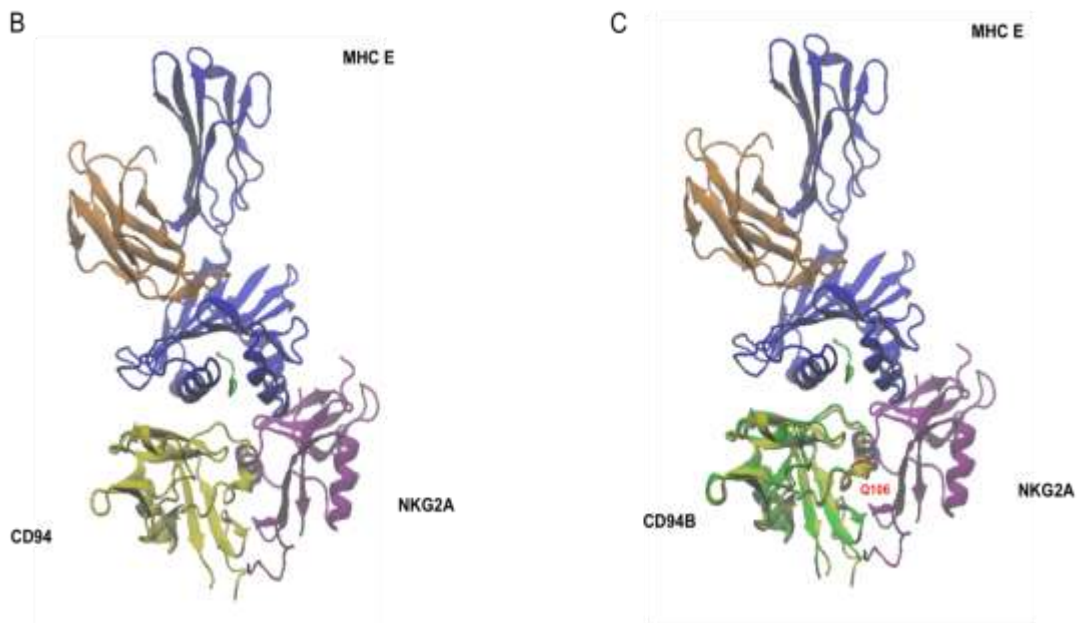


Figura 5-16. Análisis estructural de la isoforma SV1[^]. A. Estructura 3D de la isoforma SV1[^]. B. Sobreposición de SV1[^] en el complejo CD94/NKG2A-MHC E. C. Reemplazo de CD94 por SV1[^] en el complejo proteico.

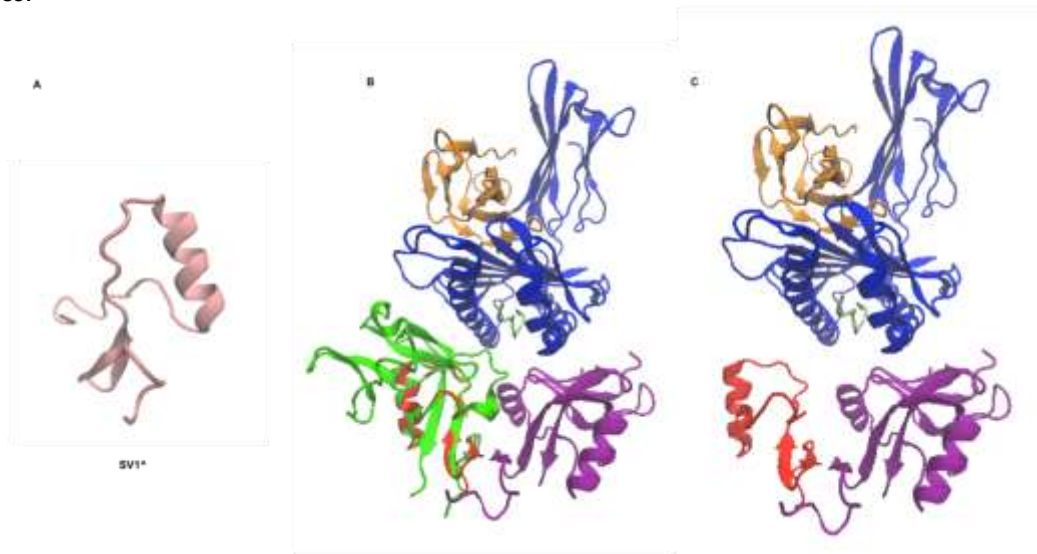
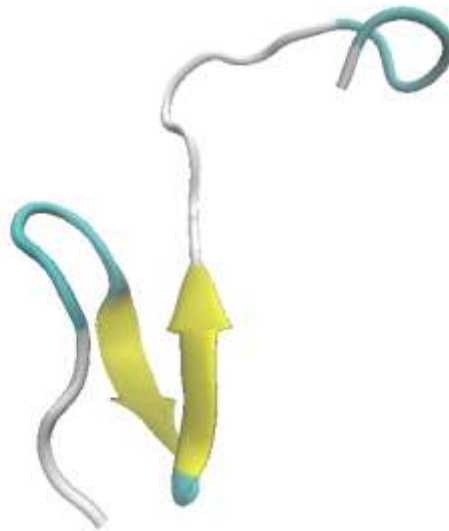
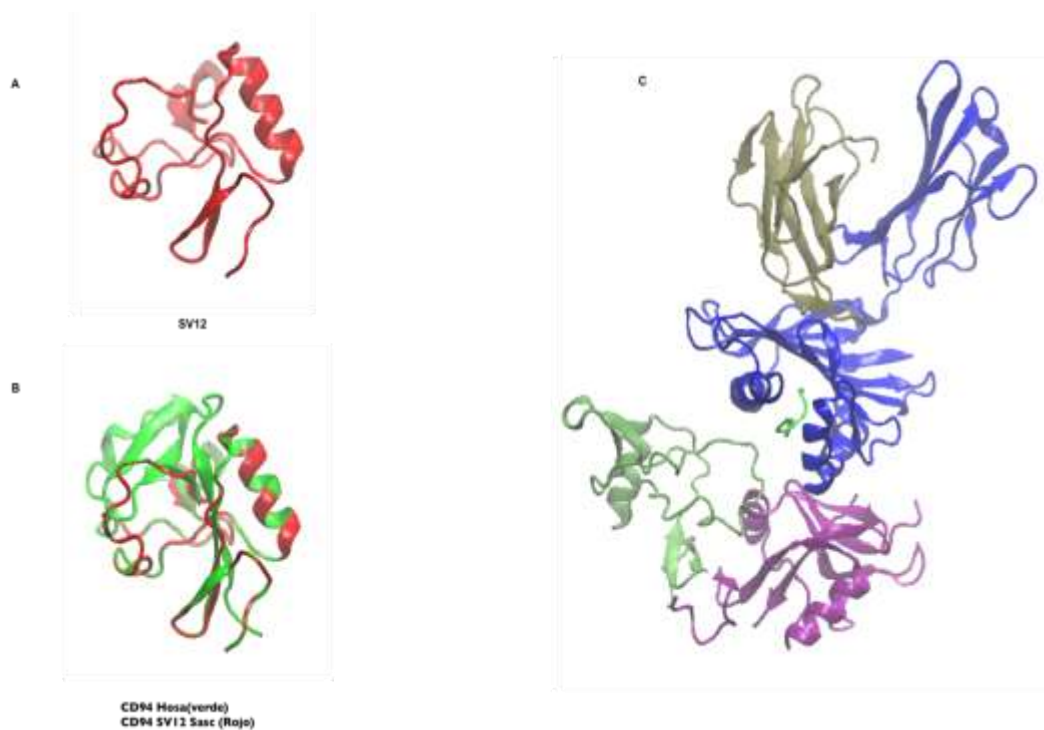


Figura 5-17. Estructura de la Isoforma SV6. Las láminas se observan en amarillo y el loop en blanco. Este péptido no muestra una estructura electroquímicamente estable. Los posibles sitios de interacción se mantienen en la región que va desde la posición 62 al 69, donde se mantienen algunas de las cisteínas que permiten los enlaces disulfuro, el resto de la proteína no muestra regiones de interacción



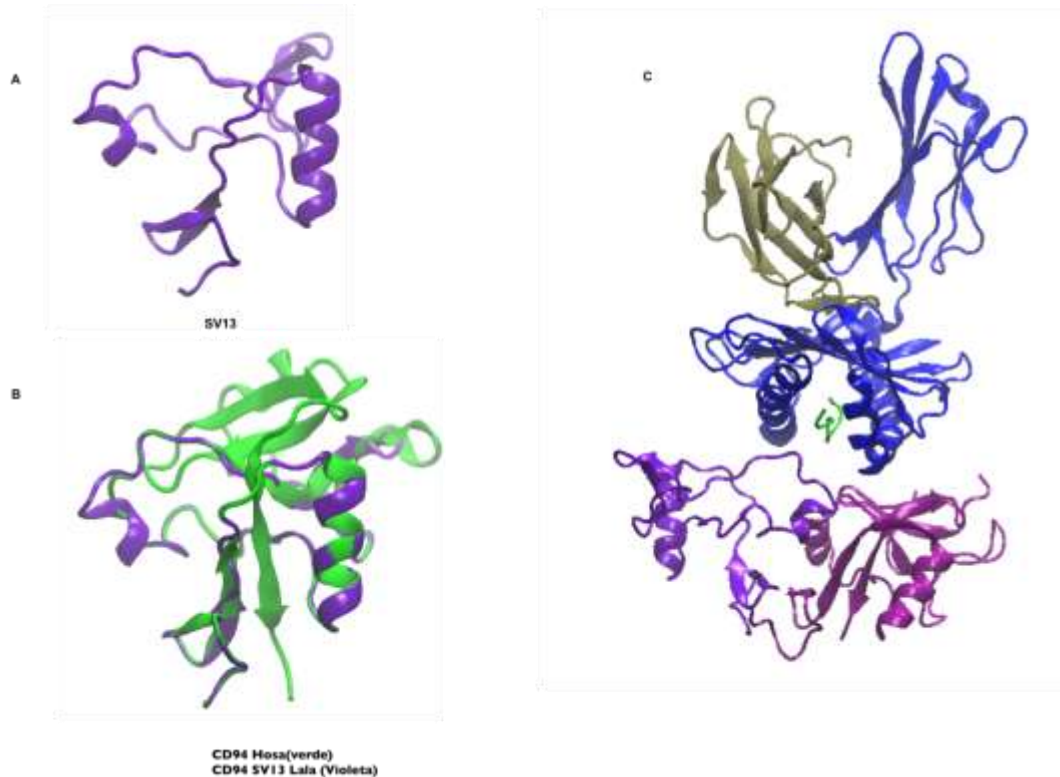
La nueva isoforma NSV12 es encontrada en este estudio *S. sciureus*. La proteína predicha conserva la identidad estructural con CD94, sin embargo, la nueva conformación estructural hace que la proteína ocupe un mayor volumen (Figura 5-18). Adicionalmente, el acortamiento de parte del dominio DRC y la pérdida de un segmento que interactúa con MHC-E, hacen que haya una interacción no estable con el complejo NKG2/MHC-E. Sin embargo, esta proteína conserva la mayoría de su superficie de interacción y probablemente es funcional.

Figura 5-18. Análisis estructural de la Isoforma SV12. A. Estructura de la isoforma SV12. B. Superposición de CD94 humana (verde) con la isoforma SV12 (roja); el dominio DRC se acorta perdiendo dos láminas B y un loop. Las regiones que se pierden interactúan con MHC-E en la proteína nativa. C La isoforma SV12 en el complejo molecular. Obsérvese que parte de la región de interacción con MHC-E se pierde, el plegamiento de la proteína ocupa un mayor volumen, dejando expuestas ciertas regiones a posibles nuevas interacciones.



La isoforma NSV13, aunque conserva parte de la similitud estructural con CD94 y su conformación estructural es muy parecida a la descrita para SV12 a pesar de ser isoformas diferentes. NSV13 presenta un mayor volumen espacial y su dominio RCD se acorta perdiendo interacción con MHC-E, pero conserva gran parte de su superficie de interacción.

Figura 5-19. Análisis estructural de la Isoforma SV13. A. Estructura de la isoforma SV13. B. Superposición de CD94 humana (verde) con la isoforma SV13 (morada); el dominio DRC se acorta perdiendo dos láminas B y un loop. Las regiones que se pierden interactúan con MHC-E en la proteína nativa. C La isoforma SV13 en el complejo molecular. Obsérvese que parte de la región de interacción con MHC-E se pierde, el plegamiento de la proteína ocupa un mayor volumen, dejando expuestas ciertas regiones a posibles nuevas interacciones



La isoforma SV14 presenta la región citoplasmática y transmembranal y 12 aminoácidos del dominio DRC, pero en el transcrito no se detectó el codón de parada. Igualmente, al perder el dominio DRC esta proteína no se puede plegar y esta isoforma seguramente es degradada intracelularmente. Algo parecido ocurre con la isoforma NSV15 que no tiene los dos primeros exones por lo tanto no se puede saber si es una isoforma funcional o si es un error en el procesamiento del ARN.

El análisis de las secuencias por modelamiento estructural del dominio DRC para las isoformas de CD94 permite inferir la función y el posible rol de las proteínas predichas para estas variables. Las isoformas CD94B, SV2, SV3 y SV16 conservan la estructura y el plegamiento del dominio DRC lo cual sugiere que conservan su función. La isoformas SV1[^] pierde parte del dominio DRC, sin embargo, su plegamiento conserva el dominio de lectinas tipo C y aunque teóricamente no es capaz de unirse a la molécula del MHC-E, puede estar actuando en otro contexto molecular. Las isoformas SV12 y SV14 aunque son productos diferentes del procesamiento del ARN, producen la misma proteína, que conserva gran parte de la estructura de CD94 y aunque pierde una región importante de la interacción con MHC-E, no se descarta que pueda tener una función.

La isoforma SV14 es un transcrito incompleto que no tiene codón de parada y que muy seguramente debe ser degradado, similar a lo que ocurre con la isoforma SV15, en la cual no aparecen en el transcrito los dos primeros exones. Con esta isoforma queda la duda si es capaz de generar una proteína funcional o si es un subproducto del mecanismo de procesamiento del ARN. Por otro lado, para las isoformas SV4, SV5, SV6 y SV7 no se pudo obtener un modelo estructural porque el fragmento que conservan del dominio DRC es muy corto y no se puede modelar, lo que sugiere que estas isoformas no son funcionales y se degradan a nivel celular. Sin embargo es interesante el hecho de que se conserven en el linaje de los primates neotropicales, lo que hace que se generen nuevos interrogantes evolutivos sobre las razones por las cuales las isoformas se conservan ampliamente en estas especies desde hace millones de años.

5.4 CARACTERIZACIÓN MOLECULAR DEL GEN NKG2D

5.4.1 Alta conservación de NKG2D en primates del Nuevo Mundo

El gen NKG2D se clonó y se secuenció en las especies *C. albifrons*, *C. apella*, *C. capucinus*, *L. lagothericha*, *Aotus sp.*, *S. leucopus*, *A. seniculus*. En las especies analizadas se encontró que NKG2D presenta una única copia. Los transcritos presentan el mismo tamaño salvo en *C. apella* donde hay una delección de 6 pb en el exón 7. Las secuencias fueron traducidas a proteínas y se compararon con las secuencias reportadas para otros primates. NKG2D consta de 281 aminoácidos, de los cuales los primeros 51 aminoácidos corresponden a la región citoplasmática, luego está la región transmembranal donde se encuentra un arginina que es la que permite la interacción con la molécula adaptadora DAP10. En el tallo se observan dos cisteínas importantes en la formación del homodímero y por último está el dominio DRC el cual es un dominio lectina tipo C, el cual fue predicho por Pfam (89) (Figura 5-20).

En los primates del Nuevo Mundo el gen NKG2D está altamente conservado mostrando un porcentaje de identidad nucleotídica promedio de 90% con su homólogo humano. Así mismo, la identidad promedio con NKG2D de lémures es de 84.5%, con el gibón de 93.66%, con el chimpancé 93.74%, con las dos especies de orangutanes 94.1%, con *Papio* un 94.17% y con las dos especies de macacos un 94.32%. Estos datos muestran que NKG2D se conserva dentro de los primates y que hay muy poca divergencia interespecífica. (Tabla 5-4).

Tabla 5-4. Porcentaje de identidad nucleotídica y de proteínas en NKG2D de primates del Nuevo Mundo

	<i>Caja</i> ¹		<i>Ceal</i>		<i>Ceap</i>		<i>Ceca</i>		<i>Lala</i>		<i>Aosp</i>		<i>Sale</i>		<i>Alse</i>	
	<i>nt</i> ²	<i>aa</i> ³	<i>Nt</i>	<i>aa</i>	<i>nt</i>	<i>aa</i>	<i>nt</i>	<i>aa</i>	<i>nt</i>	<i>aa</i>	<i>nt</i>	<i>aa</i>	<i>Nt</i>	<i>aa</i>	<i>nt</i>	<i>aa</i>
<i>Hosa</i>	91.38	91.38	93.45	93.45	93.06	93.06	93.67	93.67	93.21	93.21	93.52	93.52	92.75	92.75	93.02	93.02
<i>Caja</i>			94.67	94.67	95.45	95.45	96.08	96.08	95.61	95.61	96.71	96.71	97.65	97.65	95.14	95.14
<i>Ceal</i>					99.53	99.53	99.59	99.59	99.69	99.69	97.04	97.04	96.88	96.88	96.26	96.26
<i>Ceap</i>							99.38	99.38	99.85	99.85	96.91	96.91	96.76	96.76	96.28	96.28
<i>Ceca</i>									99.54	99.54	97.53	97.53	97.38	97.38	96.90	96.90
<i>Lala</i>											97.07	97.07	96.91	96.91	96.43	96.43
<i>Aosp</i>													97.69	97.69	97.21	97.21
<i>Sale</i>															96.74	96.74
<i>Alse</i>																

1. *Hosa*, *H. sapiens*; *Caja*, *C. Jacchus*; *Ceal*, *C. albifrons*; *Ceap*, *C. apella*; *Ceca*, *C. capucinus*; *Lala*, *L. lagotricha*; *Aosp*, *Aotus sp*, *Sale*, *S. leucopus* y *Alse*, *A. seniculus*.

2. Nucleótidos. 3. Aminoácidos

Tabla 5-5. Estimación de distancias evolutivas para NKG2D

	Caja			Ceal			Ceap			Ceca			Lala			Aosp			Sale			Alse		
	Com	Tm	DRC	Com	Tm	DRC	Com	Tm	DRC	Com	Tm	DRC	Com	Tm	DRC	Com	Tm	DRC	Com	Tm	DRC	Com	Tm	DRC
Caja				0.104	0.126	0.086	0.070	0.126	0.026	0.064	0.114	0.026	0.070	0.126	0.26	0.049	0.088	0.017	0.049	0.054	0.035	0.066	0.133	0.017
Ceal	0.047	0.066	0.033				0.039	0.000	0.076	0.045	0.010	0.076	0.039	0.000	0.076	0.082	0.077	0.086	0.098	0.089	0.016	0.099	0.115	0.086
Ceap	0.049	0.062	0.035	0.005	0.003	0.006				0.005	0.010	0.000	0.000	0.000	0.000	0.039	0.077	0.009	0.054	0.089	0.026	0.055	0.115	0.009
Ceca	0.042	0.054	0.028	0.005	0.010	0.004	0.006	0.007	0.006				0.005	0.010	0.000	0.034	0.065	0.009	0.049	0.077	0.026	0.050	0.103	0.009
Lala	0.047	0.062	0.033	0.003	0.003	0.004	0.002	0.000	0.002	0.005	0.007	0.004				0.039	0.077	0.009	0.054	0.089	0.026	0.055	0.115	0.009
Aosp	0.035	0.055	0.024	0.031	0.047	0.026	0.033	0.043	0.028	0.026	0.036	0.021	0.031	0.043	0.026				0.034	0.054	0.017	0.044	0.102	0.000
Sale	0.025	0.028	0.017	0.033	0.046	0.024	0.035	0.043	0.026	0.028	0.035	0.019	0.033	0.043	0.024	0.025	0.043	0.015				0.055	0.105	0.017
Alse	0.047	0.082	0.031	0.035	0.055	0.028	0.040	0.058	0.035	0.033	0.051	0.028	0.038	0.058	0.033	0.030	0.059	0.019	0.035	0.066	0.022			

1. Debajo de la diagonal en azul se muestran la tasa de sustituciones nucleotídicas basadas en el método de Kimura 2 parámetros (61). En la parte superior en verde se muestran las tasas de sustituciones aminoacídicas con matrices PAM (62).

2. Secuencia completa, Com: regiones citoplasmáticas, transmembranal y el tallo, TM; dominio de reconocimiento a carbohidratos DRC

5.5 NKG2D ESTÁ SUJETO A UNA FUERTE PRESIÓN PURIFICADORA

Los análisis para identificar selección por máxima verosimilitud utilizando las pruebas de PARRIS, SLAC y FEL y la prueba de REL por inferencia Bayesiana no mostraron selección positiva en ninguna posición. Sin embargo, tres posiciones son seleccionadas negativamente por al menos dos de los diferentes métodos (Tabla 5-6). Las posiciones corresponden al codón 16 que codifica una serina en la región citoplasmática, el codón 53 que codifica una fenilalanina en la región transmembranal y en el codón 157 que codifica una leucina ubicada en el dominio DRC. Para las dos primeras posiciones es difícil inferir una función debido a que solo se conoce la estructura 3D de la parte extracelular que corresponde al dominio DRC. El residuo 157L se ubica en una región distante de la zona de interacción de NKG2D con el MHC aunque es cercana a la región de interacción donde se unen los homodímeros.

Tabla 5-6. Codones seleccionados negativamente en al menos dos pruebas de selección para NKG2D.

Codón	Símbolo	Aminoácido	Región
16	S	Serina	Citoplasma
53	F	Fenilalanina	Transmembrana
157	L	Leucina	DRC

5.6 ANÁLISIS FILOGENÉTICO DEL GEN NKG2D

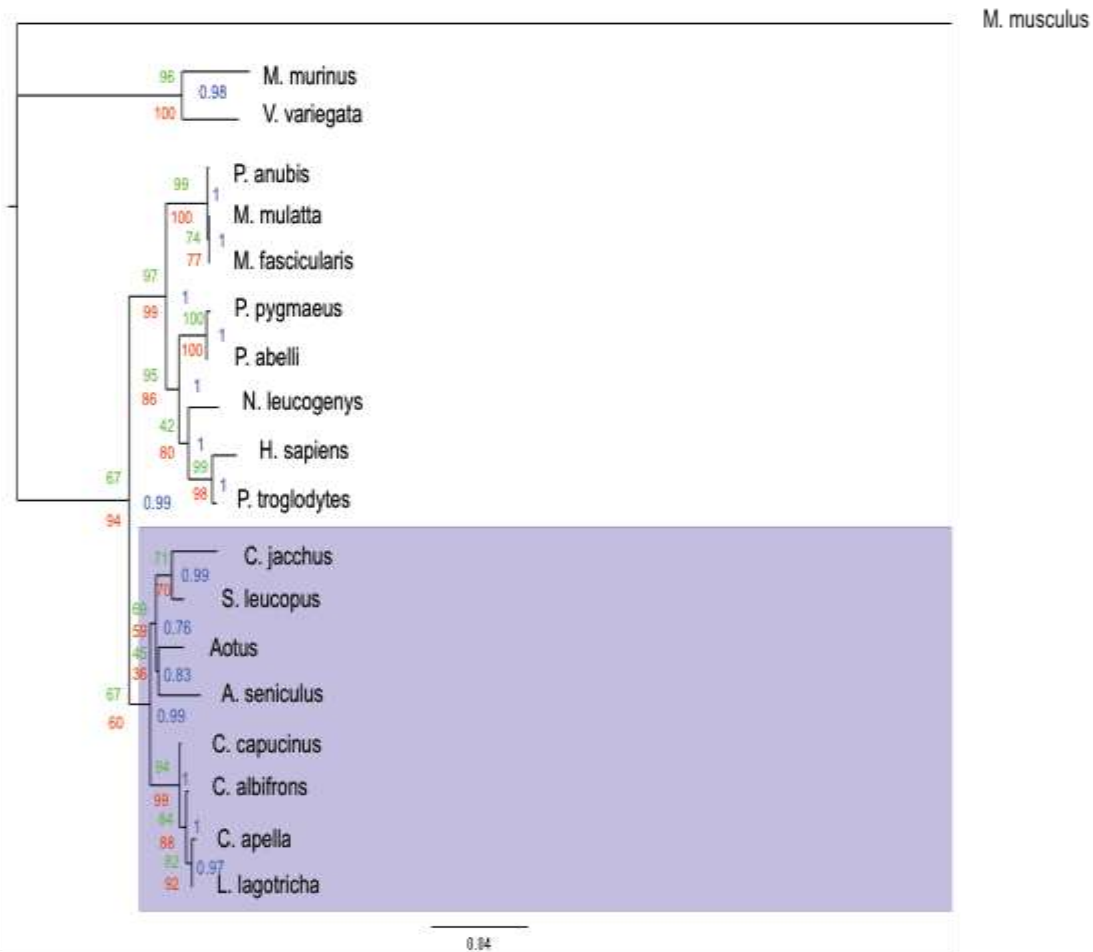
5.6.1 Filogenia de NKG2D en los primates del Nuevo Mundo

La reconstrucción filogenética del gen NKG2D se hizo por los diferentes métodos de Neighbor Joining (NJ), Máxima Verosimilitud (ML) y el de Inferencia Bayesiana. El árbol filogenético reconstruido por los tres métodos diferencia claramente los dos clados que corresponden a los subordenes Strepsirrhini y Haplorrhini (Figura 5-21). Dentro de los Haplorrhini, los valores de soporte para el clado de los Catarrhini es similar por los tres métodos y la filogenia del gen refleja la filogenia de las especies. Sin embargo, en el clado de los Platyrrhini la topología no muestra total coherencia. Los miembros de la familia Cebidae, más específicamente de la subfamilia Cebinae donde están los primates del género *Cebus*, se agrupan con un buen soporte de ramas por los tres métodos de reconstrucción filogenética utilizados. La incongruencia de este clado está en la agrupación de *C. apella* con *L. Lagothricha* un miembro de la familia Atelidae. En la otra rama se agrupan en la parte más basal *A. seniculus* un miembro de la familia Atelidae, un miembro de la familia Aotidae y los miembros de la familia Callithrichidae con las especies *C. jacchus* y *S. leucopus* (Figura 5-21).

Aunque la filogenia para las especies del Nuevo Mundo aún no está resuelta totalmente se esperaría que las familias Atelidae y Cebidae formaran clados aparte, lo cual no ocurre en la filogenia del gen NKG2D. En la topología de este árbol no se agrupan *L. lagothricha* y *A. seniculus*, ambos miembros de la familia Atelidae, estas especies se mezclan con miembros del linaje Cebidae que son especies que

divergieron tiempo después de los Atelidae. Por otro lado, en los cebidae la agrupación para el linaje Cebinae es sólida tanto en el nodo externo como en los nodos internos. La agrupación de la familia Callithrichidae es consistente en la reconstrucción para el gen NKG2D con los tres métodos. Una explicación robusta de esta inusual agrupación requeriría secuenciar NKG2D de otro individuo de la especie *L. lagotricha*.

Figura 5-21. Árbol filogenético de NKG2D. La reconstrucción filogenética se hizo con NJ, ML e inferencia bayesiana con una topología similar con los tres métodos. Los porcentajes de bootstrap se muestran en rojo para NJ y verde para ML. La probabilidad a posteriori se muestra en azul. La zona sombread muestra el suborden Platyrrhini



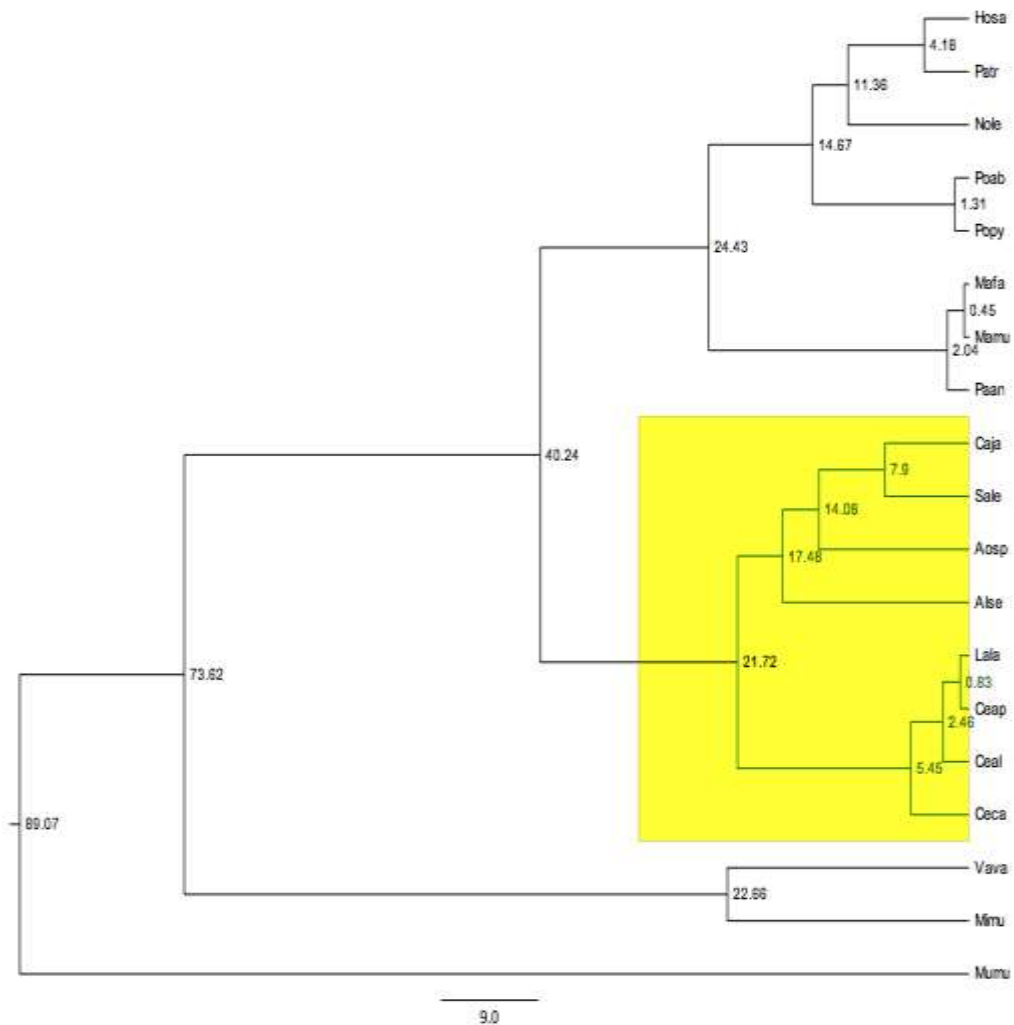
5.6.2 Estimación de tiempos de divergencia por inferencia Bayesiana.

Los tiempos de divergencia de NKG2D entre los dos subordenes Strepsirrhini y Haplorrhini es de 73.62 millones de años, lo cual es coherente con la filogenia de estos grupos que divergieron aproximadamente hace 78 millones de años en la transición del Mesozoico al Cenozoico. Dentro de los Haplorrhini, la divergencia entre Catarrhini y Platyrrhini ocurrió hace aproximadamente 43.5 Millones de años durante el Eoceno; para el gen NKG2D el tiempo de divergencia de estos grupos es de 40.24 millones de años, lo cual es consistente con el tiempo de divergencia de los grupos (Figura 5-22) (91).

En el linaje de los Platyrrhini la divergencia del ancestro común de los Phitecidae ocurrió aproximadamente hace 20.2 millones de años, y posteriormente viene la radiación de los Atelidae, siendo uno de los linajes más basales el género *Alouatta*. En el mismo espacio temporal, se inicia la radiación de taxones hermanos como los Cebinae y Saimirinae. La rápida evolución y diversificación de los primates del nuevo mundo hace que la reconstrucción filogenética sea compleja. En nuestra reconstrucción filogenética del gen NKG2D, podemos identificar el momento de separación de los Cebidae con los otros grupos y como se da un evento de radiación (Figura 5-21). Sin embargo, en los demás primates la separación de los linajes no es clara. Se esperaría que los Atelidae, un grupo basal formara un linaje propio y con un tiempo de divergencia de 20 millones de años aproximadamente. La divergencia entre NKG2D de *Alouatta* y *Aotus* es coherente con la separación de estas dos familias. Un aspecto interesante de ver es que al final de los nodos el tiempo de divergencia del gen versus el de las especies deja de ser consistente y

esto se puede deber en parte a la baja tasa de mutación del gen y a las cortas distancias de divergencia, lo que redundaría en baja resolución.

Figura 5-22. Tiempos de divergencia del gen NKG2D estimados por inferencia Bayesiana. El reloj molecular se calibró con los tiempos de divergencia entre humanos y chimpancés (6 millones de años), humanos y el lémur (74 millones de años) y entre humanos y C. apella (42 millones de años) El área resaltada en lila corresponde a Platyrrhini. Los valores en los nodos corresponden a millones de años. La zona sombreada muestra el suborden *Platyrrhini*. *Mumu Ratón*, *Mimu Lémur Naranja*, *Vava Lémur Rufo*, *Nole Gibón*, *Hosa Hombre*, *Patr Chimpancé*, *Popy Orangután*, *Poab Orangután de Sumatra*, *Paan Papión de Anubis*, *Mamu Mono Rhesus*, *Mafu Macaco cangrejero*, *Caja C. Jachus*, *Ceal C. albifrons*, *Ceap C. apella*, *Ceca C. capucinus*, *Lala L. lagotricha*, *Aosp Aotus sp*, *Sale S. leucopus* y *Alse A. seniculus*.



5.6.3 Isoformas de NKG2D generadas por procesamiento alternativo del ARN

En el marmoset (*C. jacchus*) se reportaron dos isoformas de NKG2D generadas por procesamiento alternativo del ARN (14). La primera isoforma pierde el exón que codifica para la región transmembranal y el marco de lectura se conserva. Dos características pueden afectar el funcionamiento de esta isoforma. La primera, es que probablemente esta proteína no se ancle a la membrana y, la segunda, es que en esta región se encuentra la arginina que le permite interactuar con DAP10. La otra isoforma conserva únicamente el primer y el último exón que codifican para una parte de la región citoplasmática y la parte final del dominio DRC respectivamente. Esta isoforma probablemente no es funcional.

En este estudio encontramos cuatro nuevas isoformas de NKG2D en primates del Nuevo Mundo (Figura 5-23). La isoforma SV3 se caracteriza porque ocurre un cambio en el sitio aceptor de procesamiento del ARN en el exón 2, el cual se ve reducido a 62 pb. En esta isoforma la maquinaria de procesamiento no incluye los exones 3, 4 y 5 e incluye únicamente las últimas 20 pb del exón 6 y el exón 7 se encuentra en su totalidad. La isoforma SV4 presenta todos los exones pero un cambio en el sitio aceptor de procesamiento del ARN en el último exón genera la pérdida de 88 pb. Esta isoforma mantiene abierto el marco de lectura y no se presenta un codón de parada.

La isoforma SV5 pierde el exón que codifican para el tallo y los dos primeros exones del dominio DRC, produciendo un cambio en el marco de lectura. Por esta razón la

5.6.4 La estructura de NKG2D es conservada en el linaje de los primates del Nuevo Mundo

La estructura de las proteínas NKG2D predichas para los primates del Nuevo Mundo se modelaron utilizando como plantilla la estructura cristalografiada del dominio DRC de la proteína humana (92). Las estructuras se muestran en la Figura 5-24. Se destaca en la estructura de las proteínas las dos láminas seguidas por las 2 hélices alfa que son características en el plegamiento de la familia de las lectinas tipo C. La sobreposición de las proteínas confirmó que NKG2D se conserva a nivel de secuencia y de estructura (Figura 25). Las proteínas NKG2D de primates del Nuevo Mundo muestran una región de aproximadamente 15 aminoácidos variable con respecto a la proteína humana, que se observa con azul más claro en el alineamiento estructural de la Figura 5-25. En esta región se encuentra la delección observada en *C. albifrons*. En la sobreposición no se observan cambios estructurales en esta región. Finalmente, en la Figura 5-26 se muestra las estructuras de NKG2D dimerizado interactuando con sus ligandos con su ligando ULBP16 y MICA. Se puede observar la disposición estructural del homodímero y se destaca la región de interacción la cual se caracteriza por ser espacialmente amplia lo que le permite interactuar con diferentes moléculas (Figura 5-26).

Figura 5-24. Estructuras predichas de NKG2D en primates del Nuevo Mundo. Las estructuras se obtuvieron a partir de la estructura con cristalografía de NKG2D humano (PDB 1MPU). Utilizando la suite de Swiss Model. *Caja C. jacchus*, *Ceal C. albifrons* *Ceap C. apella*, *Ceca C. capucinus*, *Lala L. lagotricha*, *Aosp Aotus sp*, *Sale S. leucopus* y *Alse A. seniculus*.

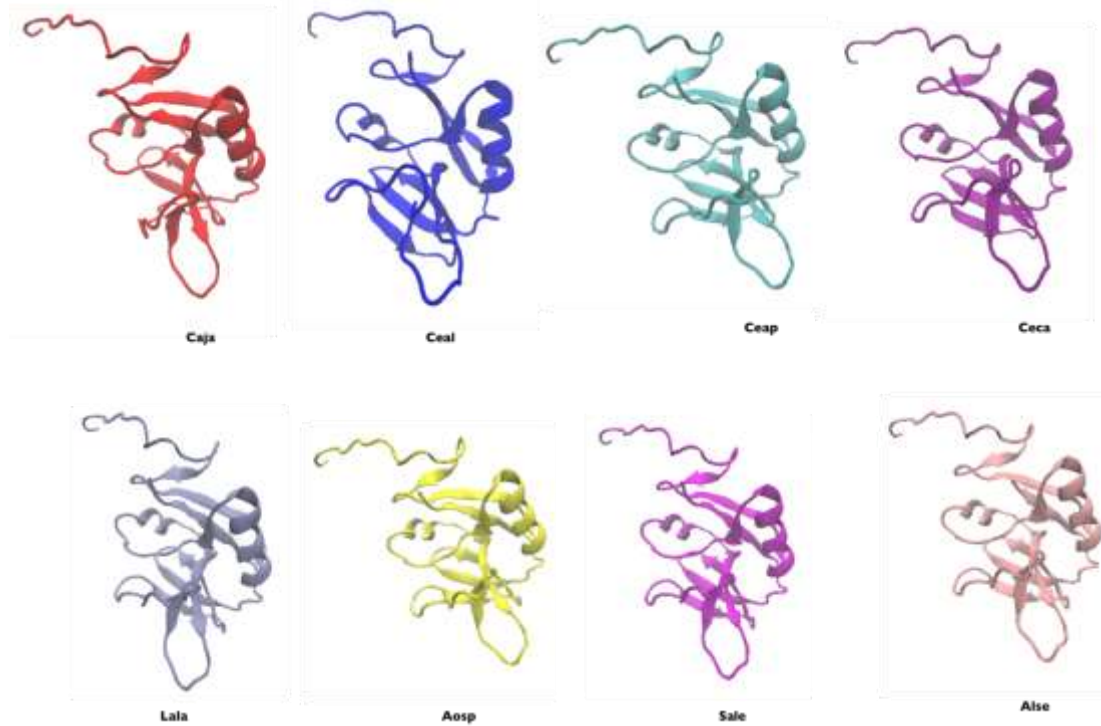


Figura 5-25. Sobreposición de las estructuras de NKG2D en primates del Nuevo Mundo. Arriba. Estructuras sobrepuestas de NKG2D de humanos con las 7 especies de primates del Nuevo Mundo. Abajo. Alineamiento de las secuencias generado a partir de la sobreposición de la estructura en la región variante; el azul fuerte representa las regiones más conservadas; el azul más claro muestra regiones conservadas con pequeños cambios en algunas especies y el rojo muestra diferencias a nivel de secuencia y de estructura.

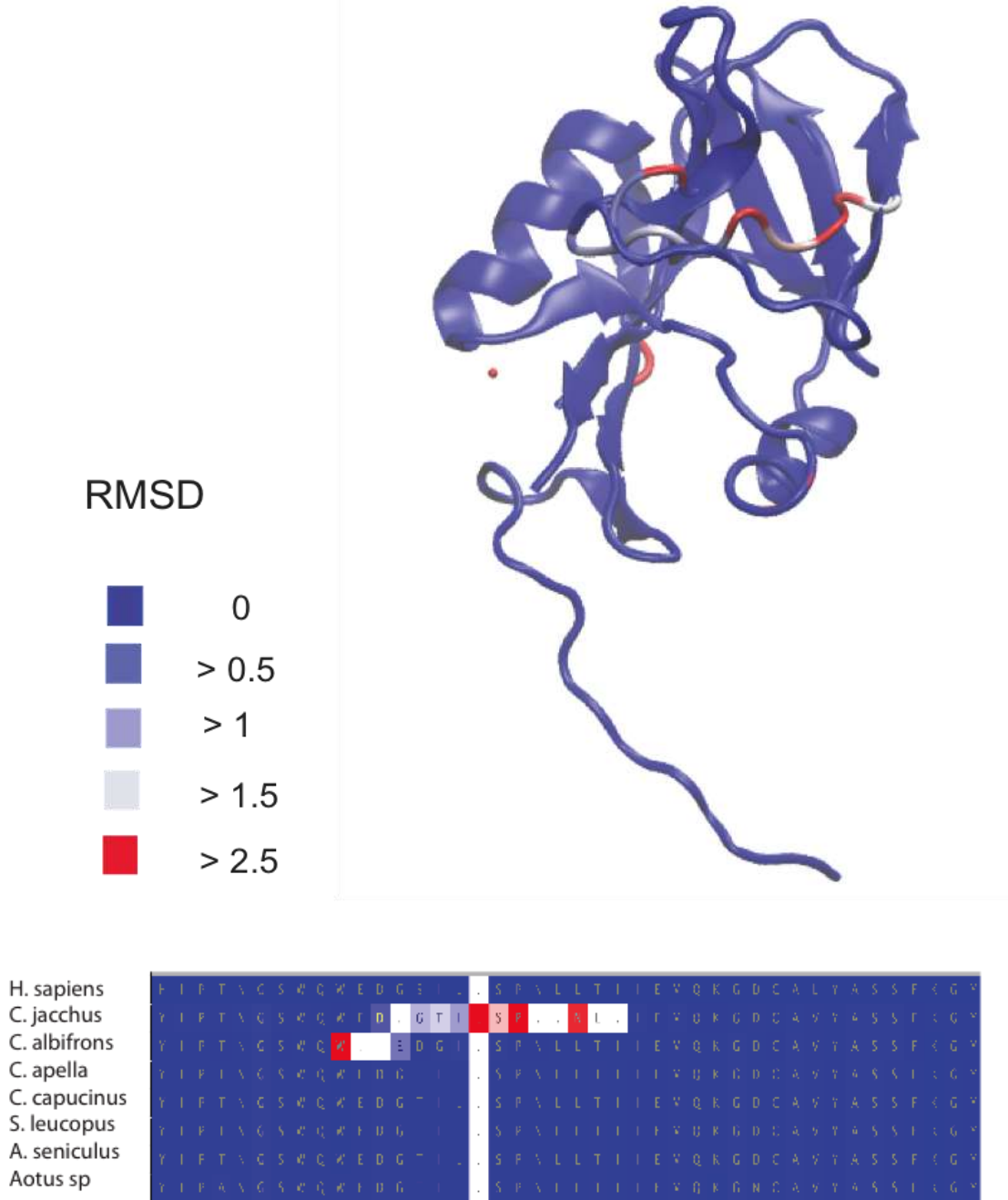
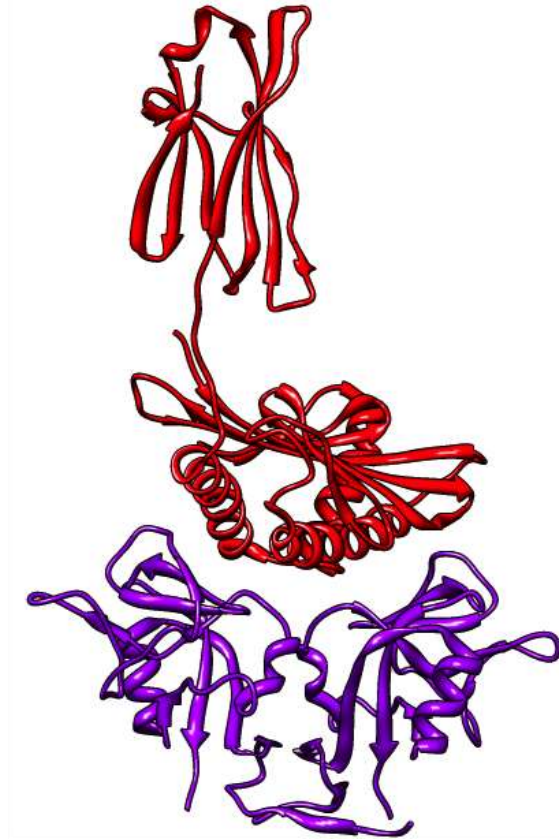
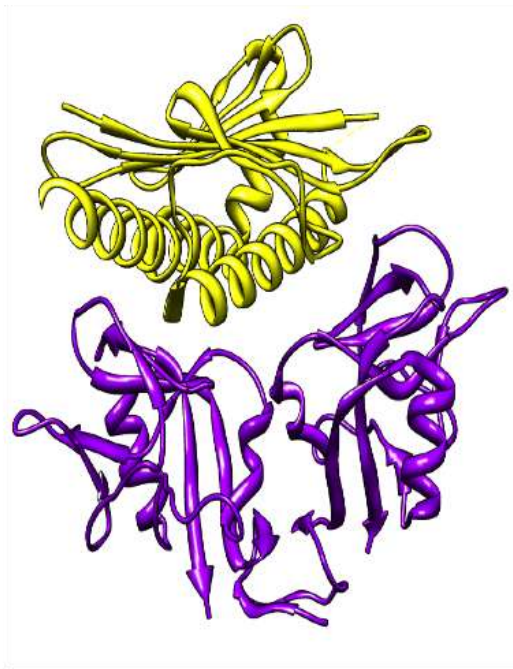


Figura 5-26. Conformación del homodímero NKG2D e interacción con sus ligandos. A la izquierda en amarillo esta la proteína ULBP 16 interactuando con NKG2D y a la derecha en rojo la molécula MICA interactuando con NKG2D. El homodímero se muestra en color morado la conformación de NKG2D permite reconocer una amplia gama de moléculas que activan la respuesta inmune.



6. Discusión de Resultados

6.1 CD94 ES UN GEN ALTAMENTE CONSERVADO EN EL LINAJE DE LOS PRIMATES

En este estudio se analizó el gen CD94 en seis especies de primates del Nuevo Mundo: *S. sciureus*, *C. albifrons*, *C. apella*, *L. lagothericha*, *A. hybridus* y *Aotus sp.* En estas especies este gen es de copia única y conserva todos sus exones. El gen se caracteriza por su alta conservación con un porcentaje de identidad superior al 90% entre Catarrhini y Platyrrhini y mayor al 95% entre Platyrrhini. El gen es monomórfico y está sujeto a un fuerte proceso de selección negativa o purificadora. La filogenia y tiempos de divergencia de CD94 reflejan la evolución de los primates y muestran que el gen se ha conservado casi intacto por más de 90 millones de años. Desviaciones de este escenario de calma evolutiva de CD94 se han reportado en otros primates. Por un lado, en el orangután *P. pygmaeus* se han identificado 6 alelos para CD94 (38), mientras que en especies del suborden Strepsirrhine ha ocurrido un proceso de duplicación génica del gen CD94 y de los genes *NKG2* que mediante un mecanismo combinatorio generan entre 15 y 24 moléculas receptoras de células NK ampliando la capacidad de reconocimiento de estos receptores (53). En los primates del Nuevo Mundo, la variabilidad parece estar impulsada por mecanismos de procesamiento alternativo del ARN, donde con un número amplio de isoformas de los genes CD94 y *NKG2* se pueden generar combinaciones de receptores tanto activadores como inhibidores y que pueden estar actuando en diferentes contextos inmunológicos(13,14,38,52).

El heterodímero CD94/NKG2 interactúa con la molécula MHC-E de forma casi exclusiva. En este estudio se pudo establecer cuáles son las posiciones involucradas en la interacción entre el dominio DRC de CD94 con la región alfa del MHC-E para los primates del Nuevo Mundo. También se determinó que las posiciones se conservan tanto en el receptor como en el ligando y que la única modificación que ocurre entre humanos y la secuencia consenso de los primates del Nuevo Mundo es en la posición 72 donde hay un cambio de una Serina por una Isoleucina. Los otros cambios ocurren en las regiones cercanas a los sitios de interacción pero no alteran el plegamiento de la proteína como lo demuestran los análisis estructurales de sobreposición. Estos hallazgos suman evidencia a que la mayor interacción electroquímica (enlaces, interacciones de Van Der Waals, etc) del complejo ocurre entre CD94 y MHC-E. Sumado a esto, el bajo polimorfismo observado en CD94 y MHC-E a lo largo del linaje de los Haplorrhini indica que un proceso co evolutivo y de selección purificante ha ocurrido tanto en el dominio de unión a MHC-E de CD94 como en la región alfa del MHC-E (13,14,38,68,93).

6.2. ASPECTOS FUNCIONALES DE LAS ISOFORMAS DE CD94

En los primates del Nuevo Mundo encontramos un repertorio de catorce isoformas de CD94 generadas por procesamiento alternativo del ARN. Estos transcritos se pueden clasificar en cuatro grupos: el primero, del cual hacen parte las isoformas CD94B, SV2, SV3 y SV16 son isoformas que conservan el plegamiento del dominio DRC. El segundo, donde están las isoformas SV1[^], SV2[^], SV12 y SV13 se caracterizan por no presentar el dominio DRC completo, sin embargo, conservan la

estructura de las lectinas tipo C. Un tercer grupo está conformado por las isoformas SV14 y SV15 donde el codón de inicio o el codón de parada no se detectan. El cuarto grupo, lo conforman las isoformas SV4, SV5, SV6 y SV7 y presentan dominios DRC muy cortos que no permiten el plegamiento de la proteína.

La isoforma CD94B que presenta la inserción Q106 genera un cambio en el loop 3 el cual no impide que se forme el heterodímero con los receptores NKG2 ni impide la interacción con MHC-E. Esta inserción cambia la conformación del centro hidrofóbico del loop3 al desplazar las posiciones Leu105, Phe107 y Met108 las cuales son estabilizadas por las posiciones Glu104-Lys175 y Asp106-Arg69 (92). Los análisis de predicción de interacción proteína-proteína no generan cambios de los residuos cargados con respecto a CD94. Esta región hidrofóbica está involucrada en la formación de un dímero, donde este centro se aplanan y se rodea con una cavidad de residuos hidrofílicos que generan un área de superficie comparable con la interfase entre los dominios V y V de los receptores de las células T (TCR). El loop 3 participa en la formación del dímero CD94-CD94 a través de las interacciones hidrofóbicas (92). Por otro lado se ha observado que en roedores el receptor CD94 se puede asociar directamente con las moléculas DAP10 y DAP12 gracias a una lisina cargada en su región transmembranal lo que muestra la plasticidad de estos receptores a pesar de ser altamente conservados en los mamíferos (94).

Las isoformas CD94 SV2 y SV3 pierden la región del tallo pero conservan el plegamiento del dominio DRC, similar a lo observado en NKG2B. Estudios funcionales han demostrado que se pueden formar diferentes tipos de heterodímeros combinando CD94 y las isoformas CD94 que no tienen el tallo (SV2 y SV3) con los receptores NKG2 y con la isoforma NKG2B. Se ha observado que las

isoformas CD94 sin tallo tienen mayor afinidad a unirse con la isoforma NKG2B la cual también ha perdido el tallo. Estas combinaciones entre las isoformas CD94 y NKG2 amplían el número de receptores tanto activadores como inhibidores con la capacidad de reconocer moléculas del MHC-E (14,38,52,55,95).

La isoforma CD94 Δ M pierde la región transmembranal y ha sido encontrada en el humano, chimpancé, macaco y el mono ardilla. En nuestro estudio encontramos en *A. hybridus* una isoforma, la SV16, que pierde casi la totalidad de la región transmembranal, por lo cual asumimos que funcionalmente puede ser igual a CD94 Δ M (14,38,52). Esta isoforma no pasa por el retículo endoplasmático y por lo cual no puede ser transportada a la membrana celular permaneciendo en el citosol, pero su función no ha sido dilucidada. Sin embargo, puede estar interactuando con isoformas que también pierden la región transmembranal como NKG2CE que ha sido encontrada en el mono ardilla y NKG2 Δ M encontrada en el macaco. Estas isoformas conservan el tallo donde se encuentra la Cys58 y el dominio DRC conserva el plegamiento (52).

El grupo de isoformas compuesto por SV1[^], SV2[^], SV12 y SV13 estructuralmente pierden la parte final del dominio DRC donde se encuentran las posiciones involucradas en la interacción con la molécula MHC-E. Sin embargo, estas moléculas tienen un plegamiento que energéticamente es favorable, conservan el dominio de lectinas tipo C y la Cys58 en el tallo que les permite la interacción con otras proteínas de membrana. Las moléculas SV1[^] y SV2[^] estructuralmente se parecen a otras lectinas tipo C del complejo NKC como la proteína LOX-1 y LY49, a otras lectinas como DC-SIGN y a la subunidad beta de la enzima coaglutinina. Estos análisis sugieren que estas isoformas pueden funcionar en otros contextos diferentes al

reconocimiento de la molécula del MHC-E, tanto reconociendo e interactuando con otras moléculas o regulando las otras isoformas de CD94.

Se ha demostrado que los receptores CD94, NKG2A, NKG2C y NKG2D son capaces de reconocer carbohidratos con estructuras sialyl Lewis, glicoproteínas con moléculas α 2,3-NeuAC, heparina y azúcares sulfatados. La expresión de estos azúcares está asociada con procesos de adhesión celular y progresión de la metástasis en células tumorales (96–99). Las interacciones moleculares en el reconocimiento de estos azúcares aún no son claras, sin embargo, se sabe que las posiciones 114 y 160 del dominio DRC de CD94 son importantes tanto en la interacción con MHC-E como en el reconocimiento de heparina (96). Lo mismo sucede con el dominio DRC de NKG2A donde algunas posiciones son importantes para la unión con ambas moléculas (98). A su vez, otras posiciones que interactúan con los azúcares no están directamente relacionadas con el reconocimiento de MHC-E. Basados en estos hallazgos se puede sugerir que tanto los receptores CD94 y NKG2 como sus isoformas pueden estar jugando un papel importante en el reconocimiento de estas moléculas y que este puede ser uno de los mecanismos utilizados por las células NK para controlar la proliferación de células malignas.

El análisis de las isoformas SV4, SV5, SV6, SV7, SV14 y SV15 no permite inferir la función de estos transcritos. Probablemente este ARN no es traducido o produce proteínas con plegamientos anormales que se degradan intracelularmente (100,101). El papel que juegan estas isoformas no es claro, aunque el hecho de que se encuentran en diferentes especies sugiere que presiones selectivas han mantenido su función hipotética. Alternativamente pueden representar vestigios evolutivos del mecanismo de procesamiento alternativo del ARN (100).

6.3. PROCESAMIENTO ALTERNATIVO DEL ARN COMO MECANISMO DE DIVERSIDAD

El mecanismo de procesamiento alternativo del ARN es una forma de generación de diversidad de proteínas y está involucrado en una variedad de procesos biológicos como el cáncer, la apoptosis, el desarrollo embrionario, el funcionamiento de las redes sinápticas, la señalización celular y la inmunidad entre otros (102–104). El mayor número de isoformas conocido que se genera por este mecanismo es el del gen Dscam en *Drosophila* donde aproximadamente existen 38.000 variantes (105). En humanos, más del 98% de los pre ARNs multi exónicos son procesados alternativamente (106). Es tal la importancia de este mecanismo que puede estar implicado en direccionar el destino de un sistema celular como son las células madre hematopoyéticas (107). Ciertamente en este sistema se ha observado que las isoformas con frecuencia generan proteínas regulatorias de las redes de diferenciación celular (107). En el sistema inmune el procesamiento alternativo del ARN tiene implicaciones en la regulación y función de la presentación antigénica por moléculas de MHC clase I, donde isoformas generadas en el gen de la cadena pesada pueden producir moléculas solubles de HLA cuyas funciones aún no han sido dilucidadas(108). También, se ha visto que este mecanismo ocurre en las proteínas de señalización de los factores de transcripción NF- κ B (109) y en la regulación de la activación y función de las células T (110).

Se conoce que hay expresión diferencial de CD94 en subpoblaciones de células NK de ratón (111,112). Un grupo con alta capacidad proliferativa y capacidad citotóxica de células NK coexpresa la molécula de CD94 y las de NKG2A/C/E. Adicionalmente, un grupo con células menos maduras presenta bajos niveles de expresión de CD94 y NKG2A/C/E está casi ausente sugiriendo que la formación del complejo CD94/NKG2/MHC-E induce la maduración funcional de este grupo celular (113,114).

En humanos la presencia de CD94 en la superficie celular también está asociada con la maduración y con la capacidad citotóxica de las células NK y pueden ser moléculas importantes en el licenciamiento que ocurre en las células NK (113–115). El repertorio de isoformas de CD94 y NKG2 en los primates puede tener implicaciones en la regulación de las células NK en procesos como la maduración, la activación o la inhibición. La amplia gama de isoformas pueden tener expresión diferencial y algunas isoformas pueden estar exclusivamente localizadas en alguna de las subpoblaciones de las células NK o pueden expresarse diferencialmente dentro del proceso de diferenciación de distintos linajes o en contextos inmunológico mediando la respuesta inmune frente a patógenos o células tumorales.

Una posibilidad que no se puede descartar es que estas isoformas alteren el funcionamiento de la repuesta inmune donde están involucrados los receptores NKC. Una hipótesis es que las isoformas que tienen reducido el dominio DRC no interactúan con MHC y pueden estar compitiendo con la proteína normal y de esta manera trunquen la formación de heterodímeros funcionales. Aunque el mecanismo de procesamiento alternativo del ARN muestra un amplio espectro de ventajas también se ha estudiado ampliamente sus implicaciones en el desarrollo de diferentes tipos de cáncer (116) y en enfermedades autoinmunes como la esclerosis múltiple (117).

Las isoformas generadas por procesamiento alternativo del ARN tienden a perder algunos dominios, lo que genera cambios a nivel estructural que pueden alterar la función de la proteína. Sin embargo, se ha observado que estas nuevas isoformas tienen un alto grado de plasticidad y son capaces de tolerar los re-arreglos a nivel estructural y generan proteínas estables, por lo tanto pueden adquirir diferentes conformaciones y nuevas funciones (100,118). Gracias a la plasticidad estructural

de las proteínas, el procesamiento alternativo del ARN puede tener un rol evolutivo ya que produce diversidad estructural y funcional (100,118). En este estudio encontramos que las estructuras de la mayoría de isoformas son estables energéticamente en su plegamiento, no muestran zonas difíciles de modelar y las estructuras que se conservan muestran una similitud con el dominio lectina tipo C y con la misma CD94. Por lo tanto se puede esperar que sean proteínas estables y posiblemente funcionales. Esta noción es favorecida por el hecho que muchas de estas isoformas se conservan a lo largo del linaje de los primates tanto en Catarrhini como en Platyrrhini.

En CD94 observamos patrones frecuentes de procesamiento del ARN como la pérdida del tallo o del exón 5 que codifica el primer fragmento del dominio CDR. También, son frecuentes los cambios en los sitios aceptores de procesamiento del ARN generando inserciones de fragmentos intrónicos, como es el caso de tres pares de base que insertan una glutamina al loop 3. Estos patrones de procesamiento alternativo del ARN pueden estar siendo favorecidos por procesos de selección. Es el caso de la inserción Q106 que está presente en siete isoformas puede estar favoreciendo el plegamiento de la estructura del loop 3, el cual juega un papel importante en la conformación del centro hidrofóbico de la proteína, además interactúa con las láminas $\beta 5$ y $\beta 6$ y es pieza clave en la formación del homodímero (CD94-CD94) (92). Esto sugiere que un mecanismo de selección natural sobre el mecanismo de procesamiento del ARN puede estar ocurriendo y que el procesamiento alternativo del ARN puede jugar un rol importante en la evolución de distintos procesos biológicos (119,120).

El gen CD94 se conserva a lo largo del linaje de los primates y ha co evolucionado tanto con sus co receptores como son los genes NKG2 y con su ligando la molécula MHC-E. Todo parece indicar que los procesos coevolutivos han ocurrido también en las isoformas generadas por procesamiento alternativo del ARN y esto se refleja en que isoformas de CD94 que pierden el tallo o la región transmembranal tienen su contraparte en los receptores NKG2, lo que puede facilitar su dimerización y su función. En los primates neotropicales una expansión de isoformas de CD94 por procesamiento alternativo del ARN ha ocurrido, posiblemente ampliando el espectro de funciones e interacciones. Probablemente estos receptores tienen la capacidad de pensar otro tipo de moléculas como son carbohidratos glicosilados y glicoproteínas y pueden ser importantes en entender procesos de inmunovigilancia de células tumorales o infectadas por virus. En este estudio se hace evidente que el procesamiento alternativo del ARN es un mecanismo de generación de diversidad y que en los primates neotropicales en el gen CD94 ciertas isoformas del gen CD94 se expresan regularmente.

6.4. NKG2D ES UN GEN ALTAMENTE CONSERVADO EN EL LINAJE DE LOS PRIMATES

Analizamos el gen NKG2D en las especies de primates neotrópicales *C. albifrons*, *C. apella*, *C. capucinus*, *L. lagothric*, *Aotus sp*, *S. leucopus*, *A. seniculus*. El gen NKG2D es de copia única, es altamente conservado en su estructura y en su secuencia y dentro del linaje de los primates del Nuevo Mundo el porcentaje de identidad nucleotídica es de 97.24%. Este gen es monomorfo en Platyrrhini y tiene un único motivo ITAM el cual le permite interactuar exclusivamente con la molécula adaptadora DAP10. *C. apella* presenta una deleción de 6 pb en el exón 7 que hace parte del dominio DRC. Los análisis de selección no muestran selección positiva y la selección negativa se limita a tres posiciones que no tienen participación con su ligando.

En la zarigüeya, un marsupial basal en la evolución de los mamíferos se identificó el ortólogo de NKG2D pero no los ortólogos de CD94/NKG2 (121). En el ornitorrinco un mamífero antiguo, no se han encontrado genes homólogos a CD94 y NKG2 (122). Aunque la caracterización del gen NKG2D no se ha realizado en la mayoría de los mamíferos se sabe que está presente con una única copia en los genomas de carnívoros (48,123), artiodáctilos y roedores (48,123). En estos últimos una duplicación del motivo ITAM ha ocurrido lo que le permite interactuar con moléculas DAP10 y DAP12. En los dos subordenes de primates, Strepsirrhine y Haplorrhini el gen NKG2D está presente en una única copia (124) y el análisis filogenético de este estudio con los tiempos de divergencia reconstruye la historia evolutiva del orden con precisión. Sin embargo, en los primates del nuevo mundo el linaje Atelidae se confunde con Cebidae generando incongruencias en la filogenia de este grupo.

6.5 PROCESAMIENTO ALTERNATIVO DEL ARN EN NKG2D

Para los primates del Nuevo Mundo en el gen NKG2D se ha encontrado seis isoformas, cinco de las cuales probablemente sean degradadas intracelularmente. La predicción funcional por métodos bioinformáticos no arrojaron ningún resultado. Además, no hay isoformas compartidas entre las diferentes especies sugiriendo que son subproductos no funcionales de la maduración del ARN (125,126). La isoforma SV1 que fue reportada previamente en el marmoset, pierde la región transmembranal donde se encuentra el motivo ITAM. Al perder la región transmembranal esta isoforma no puede anclarse en la membrana y probablemente esta proteína se retiene en el citoplasma, y podría unirse a la proteína NKG2D y actuar como regulador (127).

6.6 IMPLICACIONES FUNCIONALES DE NKG2D

NKG2D es un homodímero activador que se expresa en células NK, en células NKT y linfocitos T citotóxicos. En primates se asocia exclusivamente con la molécula adaptadora DAP10 que activa la ruta de señalización activadora mediante la interacción con la molécula fosfatidil inositol 3 quinasa (127). NKG2D mantiene una estructura tan conservada posiblemente porque esta le permite sin necesidad de cambiar su conformación estructural unirse a una variedad de ligandos con estructuras comunes y de esta manera puede actuar en diferentes contextos inmunológicos. Por lo tanto, conservar su plegamiento es lo que le da la afinidad molecular con sus ligandos y por esta razón conservar su estructura es una ventaja evolutiva (63,127). En humanos se ha reportado siete ligandos dentro de los cuales están las moléculas MICA y MICB que son homologas de MHC clase I y otras glicoproteínas como son la UBPL que también se relacionan con MHC clase I (62).

NKG2D es también capaz de reconocer carbohidratos con estructuras sialyl Lewis, moléculas $\alpha 2,3$ -NeuAC, y otras glicoproteínas (96,97,99). Por lo tanto, su funcionalidad y su interacción es dada por la expresión y la afinidad por el ligando. Como es el caso del ligando RAE-1 que no se expresa en tejidos sanos sino en células tumorales e infectadas con virus. De esta manera al ocurrir la interacción entre el receptor y ligando las células NK pueden desempeñar su papel y activar una respuesta inmune frente al tejido dañado. Con este tipo de interacciones las células NK son capaces de reconocer un tejido sano de uno enfermo (127).

7. Conclusiones y Recomendaciones

Los genes CD94 y NKG2D son altamente conservados en el linaje de los primates y su filogenia refleja la de sus especies. La baja variabilidad se debe a que están bajo selección purificante para mantener la interacción con sus ligandos.

En los primates neotropicales el procesamiento alternativo del ARN en el gen CD94 juega un papel importante en la generación de variabilidad, produciendo isoformas posiblemente funcionales que pueden estar jugando un papel importante en las células Natural Killer y en los procesos inmunológicos de control de células infectadas por virus y células tumorales.

Un proceso evolutivo y de selección puede estar ocurriendo en algunas isoformas de CD94, como ocurre en CD94B, la cual puede estar favoreciendo el plegamiento de la estructura y sus interacciones con otros receptores. Adicionalmente, otras isoformas pueden estar ampliando su rango de interacción molecular, probablemente uniendo azúcares sulfatados.

NKG2D tiene un amplio rango de interacción molecular y se debe en parte a que conserva su plegamiento estructural lo cual le permite interactuar con un amplio número de ligandos con estructuras conservadas y de esta manera jugar un papel importante en la activación de la respuesta inmune.

8. Bibliografía

1. Defler TR. Historia natural de los primates colombianos. Conservación Internacional Colombia. 2004.
2. Stevenson PR, Guzmán DC DT. Conservation of Colombian primates: an analysis of published research. *Tropical Conservation Science*. 2010;3:45-62.
3. Wolfe ND, Dunavan CP, Diamond J. Origins of major human infectious diseases. *Nature* 2007;447:279-83.
4. Wolfe ND, Escalante a a, Karesh WB, et al. Wild primate populations in emerging infectious disease research: the missing link? *Emerging infectious diseases* 1998;4:149-58.
5. Chinchilla M, Guerrero OM, Gutierrez-Espeleta GA, Sánchez R VCI. Parásitos en monos carablanca *Cebus capucinus* (primates: Cebidae) de Costa Rica. *Parasitol Latinoam*. 2007;62:170-175.
6. Malta MCC, Tinoco HP, Xavier MN, et al. Naturally acquired visceral leishmaniasis in non-human primates in Brazil. *Veterinary parasitology* 2010;169:193-7.
7. Wood JD, Peck OC, Tefend KS, et al. Colitis and colon cancer in cotton-top tamarins (*Saguinus oedipus oedipus*) living wild in their natural habitat. *Digestive diseases and sciences*. 1998;43:1443-53.
8. Kelley J, Walter L, Trowsdale J. Comparative genomics of natural killer cell receptor gene clusters. *PLoS genetics*. 2005;1:129-39.
9. Dimasi, Nazzareno BR. Review Article Structural and functional aspects of the Ly49 natural killer cell receptors. *Immunology and cell biology*. 2005;83:1-8.

10. Parham P. MHC class I molecules and KIRs in human history, health and survival. *Nature reviews. Immunology*. 2005;5:201-14.
11. Yokoyama WM. Commentary HLA class I specificity for natural killer cell receptor CD94NKG2A: Two for one in more ways than one. *Rheumatology*. 1998;95:4791-4794.
12. Barten R, Torkar M, Haude a, et al. Divergent and convergent evolution of NK-cell receptors. *Trends in immunology* 2001;22:52-7.
13. LaBonte ML, Russo J, Freitas S, et al. Variation in the ligand binding domains of the CD94/NKG2 family of receptors in the squirrel monkey. *Immunogenetics*. 2007;59:799-811.
14. Averdam A, Kuhl H, Sontag M, et al. Genomics and diversity of the common marmoset monkey NK complex. *Journal of immunology*. 2007;178:7151-61.
15. Canavez FC, Moreira M a, Ladasky JJ, et al. Molecular phylogeny of new world primates (Platyrrhini) based on beta2-microglobulin DNA sequences. *Molecular phylogenetics and evolution*. 1999;12:74-82.
16. Schneider H, Schneider MPC, Sampaio I, Harada ML, Stanhope M, Czelusniak GM. Molecular phylogeny of the New World Monkeys (Platyrrhini, primates). *Molecular phylogenetics and evolution*. 1993;2:225-242.
17. Osterholz M, Walter L, Roos C. Retropositional events consolidate the branching order among New World monkey genera. *Molecular phylogenetics and evolution*. 2009;50:507-13.
18. Schrago CG, Russo C a M. Timing the origin of New World monkeys. *Molecular biology and evolution*. 2003;20:1620-5.
19. Defler TR, Bueno ML, García J. Callicebus caquetensis: A New and Critically Endangered Titi Monkey from Southern Caquetá, Colombia. *Primate conservation*. 2010;25:1-9.

20. Carme B, Ajzenberg D, Demar M, et al. Veterinary Parasitology Outbreaks of toxoplasmosis in a captive breeding colony of squirrel monkeys. 2009;163:132-135.
21. Bryant J, Wang H, Cabezas C, et al. Enzootic transmission of yellow fever virus in Peru. *Emerging infectious diseases*. 2003;9:926-33.
22. Brack M. Lectin Histochemistry and Carcinoembryonic Antigen in Spontaneous Colonic Cancers of Cotton-top Tamarins (*Saguinus oedipus*). *Veterinary Pathology* 1995;32:668-673.
23. Cooper MD, Herrin BR. How did our complex immune system evolve? *Nature reviews. Immunology* 2010;10:2-3.
24. Litman GW, Rast JP, Fugmann SD. The origins of vertebrate adaptive immunity. *Nature reviews. Immunology*. 2010;10:543-53.
25. Andoniou CE, Andrews DM, Degli-Esposti M a. Natural killer cells in viral infection: more than just killers. *Immunological reviews*. 2006;214:239-50.
26. Cella M, Fuchs A, Vermi W, et al. A human natural killer cell subset provides an innate source of IL-22 for mucosal immunity. *Nature* 2009;457:722-5.
27. Hiyama M, Kusakabe KT, Kuwahara A, et al. Differentiation of uterine natural killer cells in pregnant SCID (scid/scid) mice. *The Journal of veterinary medical science* 2011;73:1337-40.
28. Spanholtz J, Tordoier M, Eissens D, et al. High log-scale expansion of functional human natural killer cells from umbilical cord blood CD34-positive cells for adoptive cancer immunotherapy. *PloS one* 2010;5:e9221.
29. Cooper M a, Colonna M, Yokoyama WM. Hidden talents of natural killers: NK cells in innate and adaptive immunity. *EMBO reports*. 2009;10:1103-10.
30. Vivier E, Ugolini S, Blaise D, et al. Targeting natural killer cells and natural killer T cells in cancer. *Nature reviews. Immunology* 2012;12:239-52.

31. Smyth MJ, Cretney E, Kelly JM, et al. Activation of NK cell cytotoxicity. *Molecular immunology*. 2005;42:501-10.
32. Bryceson YT, Chiang SCC, Darmanin S, et al. Molecular mechanisms of natural killer cell activation. *Journal of innate immunity*. 2011;3:216-26.
33. Bryceson YT, March ME, Ljunggren H-gustaf, et al. Activation, coactivation, and costimulation of resting human natural killer cells. *Immunological Reviews*. 2006;214:73-91.
34. Long EO. Regulation of immune responses through inhibitory receptors. *Annual review of immunology* 1999;17:875-904.
35. Hershberger KL, Shyam R, Miura a, et al. Diversity of the killer cell Ig-like receptors of rhesus monkeys. *Journal of immunology* 2001;166:4380-90.
36. Sambrook JG, Bashirova A, Palmer S, et al. Single haplotype analysis demonstrates rapid evolution of the killer immunoglobulin-like receptor (KIR) loci in primates. *Genome research*. 2005;15:25-35.
37. Mager DL, McQueen KL, Wee V, et al. Evolution of natural killer cell receptors: coexistence of functional Ly49 and KIR genes in baboons. *Current biology: CB*. 2001;11:626-30.
38. Shum BP, Flodin LR, Muir DG, et al. Conservation and variation in human and common chimpanzee CD94 and NKG2 genes. *Journal of immunology*. 2002;168:240-52.
39. Chiang H-I, Zhou H, Raudsepp T, et al. Chicken CD69 and CD94/NKG2-like genes in a chromosomal region syntenic to mammalian natural killer gene complex. *Immunogenetics*. 2007;59:603-11.
40. Nikolaidis N, Makalowska I, Chalkia D, et al. Origin and evolution of the chicken leukocyte receptor complex. *Proceedings of the National Academy of Sciences*. 2005;102:4057-62.

41. Ryan JC, Naper C, Hayashi S, et al. Physiologic functions of activating natural killer (NK) complex-encoded receptors on NK cells. *Immunological reviews*. 2001;181:126-37.
42. Hofer, E, Sobanov Y, Brostjan C, Lehrach H DM. The centromeric part of the human NK receptor complex lectin like receptor genes expressed in NK, dendritic and endothelial cells. *Immunological reviews*. 2001;181:5-19.
43. Bull C, Sobanov Y, Röhrdanz B, et al. The centromeric part of the human NK gene complex: linkage of LOX-1 and LY49L with the CD94/NKG2 region. *Genes and immunity*. 2000;1:280-7.
44. Spreu J, Kienle EC, Schrage B, et al. CLEC2A: a novel, alternatively spliced and skin-associated member of the NKC-encoded AICL-CD69-LLT1 family. *Immunogenetics*. 2007;59:903-12.
45. Batliner J, Mancarelli MM, Jenal M, et al. CLEC5A (MDL-1) is a novel PU.1 transcriptional target during myeloid differentiation. *Mol Immunol*. 2011;48:714-719.
46. Bambard ND, Mathew SO, Mathew P a. LLT1-mediated activation of IFN-gamma production in human natural killer cells involves ERK signalling pathway. *Scandinavian journal of immunology*. 2010;71:210-9.
47. Carlyle JR, Jamieson AM, Gasser S, et al. Missing self-recognition of Oclt⁻Clr-b by inhibitory NKR-P1 natural killer cell receptors. *Proceedings of the National Academy of Sciences*. 2004;101:3527-52.
48. Yoder J a, Litman GW. The phylogenetic origins of natural killer receptors and recognition: relationships, possibilities, and realities. *Immunogenetics*. 2011;63:123-41.
49. Lieto LD, Maasho K, West D, et al. The human CD94 gene encodes multiple, expressible transcripts including a new partner of NKG2A/B. *Genes and immunity*. 2006;7:36-43.
50. Ryan JC, Seaman WE. Divergent functions of lectin-like receptors on NK cells. *Immunological reviews*. 1997;155:79-89.

51. Braud VM, Allan DS, O'Callaghan C a, et al. HLA-E binds to natural killer cell receptors CD94/NKG2A, B and C. *Nature*. 1998;391:795-9.
52. LaBonte ML, Levy DB, Letvin NL. Characterization of rhesus monkey CD94/NKG2 family members and identification of novel transmembrane-deleted forms of NKG2-A, B, C, and D. *Immunogenetics*. 2000;51:496-9.
53. Averdam A, Petersen B, Rosner C, et al. A novel system of polymorphic and diverse NK cell receptors in primates. *PLoS genetics*. 2009;5:e1000688.
54. LaBonte ML, Letvin NL. Variable NKG2 expression in the peripheral blood lymphocytes of rhesus monkeys. *Clinical and experimental immunology*. 2004;138:205-12.
55. Lazetic S, Chang C, Houchins JP, et al. Human natural killer cell receptors involved in MHC class I recognition are disulfide-linked heterodimers of CD94 and NKG2 subunits. *Journal of immunology*. 1996;157:4741-5.
56. LaBonte ML, Letvin NL. Variable NKG2 expression in the peripheral blood lymphocytes of rhesus monkeys. *Clinical and experimental immunology*. 2004;138:205-12.
57. Guethlein L a, Flodin LR, Adams EJ, et al. NK cell receptors of the orangutan (*Pongo pygmaeus*): a pivotal species for tracking the coevolution of killer cell Ig-like receptors with MHC-C. *Journal of immunology*. 2002;169:220-9.
58. Champsaur M LL. Effect of NKG2D ligand expression on host immune responses. *Immunological reviews*. 2010;235:267-285.
59. Kondo M, Maruoka T, Otsuka N, et al. Comparative genomic analysis of mammalian NKG2D ligand family genes provides insights into their origin and evolution. *Immunogenetics*. 2010;62:441-50.
60. Li P, Morris DL, Willcox BE, et al. Complex structure of the activating immunoreceptor NKG2D and its MHC class I-like ligand MICA. *Nature immunology*. 2001;2:443-51.

61. Hayakawa Y, Smyth MJ. NKG2D and cytotoxic effector function in tumor immune surveillance. *Seminars in immunology*. 2006;18:176-85.
62. Eagle R a, Trowsdale J. Promiscuity and the single receptor: NKG2D. *Nature reviews. Immunology*]. 2007;7:737-44.
63. Cerwenka A. New twist on the regulation of NKG2D ligand expression. *The Journal of experimental medicine*. 2009;206:265-8.
64. Kaiser BK, Pizarro JC, Kerns J, et al. Structural basis for NKG2A/CD94 recognition of HLA-E. *Proceedings of the National Academy of Sciences*. 2008;105:6696-701.
65. Kaiser BK, Barahmand-Pour F, Paulsene W, et al. Interactions between NKG2x immunoreceptors and HLA-E ligands display overlapping affinities and thermodynamics. *Journal of immunology*. 2005;174:2878-84.
66. Petrie EJ, Clements CS, Lin J, et al. CD94-NKG2A recognition of human leukocyte antigen (HLA)-E bound to an HLA class I leader sequence. *The Journal of experimental medicine*. 2008;205:725-35.
67. Lee N, Llano M, Carretero M, et al. HLA-E is a major ligand for the natural killer inhibitory receptor CD94/NKG2A. *Proceedings of the National Academy of Sciences*. 1998;95:5199-204.
68. Knapp L a, Cadavid LF, Watkins DI. The MHC-E locus is the most well conserved of all known primate class I histocompatibility genes. *Journal of immunology*. 1998;160:189-96.
69. Larkin M a, Blackshields G, Brown NP, et al. Clustal W and Clustal X version 2.0. *Bioinformatics*. 2007;23:2947-8.
70. Hall T. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95:98:NT. *Nucleic acids symposium series*.1999;41:95-98.
71. Meng EC, Pettersen EF, Couch GS, et al. Tools for integrated sequence-structure analysis with UCSF Chimera. *BMC bioinformatics*. 2006;7:339.

72. Kimura M. A simple method for estimating evolutionary rates of substitutions through comparative studies of Nucleotide sequence. *Journal of Molecular Evolution*. 1980;16:111-120.
73. Wilbur WJ. On the PAM matrix model of protein evolution. *Molecular biology and evolution*. 1985;2:434-47.
74. Tamura K, Peterson D, Peterson N, et al. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular biology and evolution*. 2011;28:2731-9.
75. Scheffler K, Martin DP, Seoighe C. Robust inference of positive selection from recombining coding sequences. *Bioinformatics*. 2006;22:2493-9.
76. Kosakovsky Pond SL, Frost SDW. Not so different after all: a comparison of methods for detecting amino acid sites under selection. *Molecular biology and evolution*. 2005;22:1208-22.
77. Delport W, Poon AFY, Frost SDW, et al. Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. *Bioinformatics*. 2010;26:2455-7.
78. Posada D, Crandall KA. Modeltest : testing the model of DNA substitution. *Evolution*. 1998;14:817-818.
79. Felsenstein. Confidence Limits on Phylogenies : An Approach Using the Bootstrap. *Evolution*. 1985;39:783-791.
80. Zwickl DJ. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. Dissertation. *University of Texas at Austin*. 2006;
81. Swofford DL. Phylogenetic Analysis Using Parsimony *. *Sinaeur Associates, Sunderland, Massachusetts*. 2002;

-
82. Huelsenbeck JP, Ronquist F. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics*. 2001;17:754-5.
83. Drummond AJ, Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC evolutionary biology*. 2007;7:214-222.
84. Hedges SB, Dudley J, Kumar S. TimeTree: a public knowledge-base of divergence times among organisms. *Bioinformatics*. 2006;22:2971-2.
85. Arnold K, Bordoli L, Kopp J, et al. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics*. 2006;22:195-201.
86. Benkert P, Schwede T, Tosatto SC. QMEANclust: estimation of protein model quality by combining a composite scoring function with structural density information. *BMC structural biology*. 2009;9:35.
87. Kelley L a, Sternberg MJE. Protein structure prediction on the Web: a case study using the Phyre server. *Nature protocols*. 2009;4:363-71.
88. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. *Journal of molecular graphics*. 1996;14:33-8, 27-8.
89. Finn RD, Tate J, Mistry J, et al. The Pfam protein families database. *Nucleic acids research*. 2008;36:281-8.
90. Qin S, Zhou H-X. meta-PPISP: a meta web server for protein-protein interaction site prediction. *Bioinformatics*. 2007;23:3386-7.
91. Perelman P, Johnson WE, Roos C, et al. A molecular phylogeny of living primates. *PLoS genetics*. 2011;7:e1001342.
92. Boyington JC, Riaz AN, Patamawenu A, et al. Structure of CD94 Reveals a Novel C-Type Lectin Fold: Implications for the NK Cell – Associated CD94 / NKG2 Receptors. *Immunity*. 1999;10:75-82.

93. Joly E. Various hypotheses on MHC evolution suggested by the concerted evolution of CD94L and MHC class Ia molecules. *Biology direct* 2006;1:3.
94. Saether PC, Hoelsbrekken SE, Fossum S, et al. Rat and mouse CD94 associate directly with the activating transmembrane adaptor proteins DAP12 and DAP10 and activate NK cell cytotoxicity. *Journal of immunology*. 2011;187:6365-73.
95. LaBonte ML, Choi EI, Letvin NL. Molecular Determinants Regulating the Pairing of NKG2 Molecules with CD94 for Cell Surface Heterodimer Expression. *The Journal of Immunology*. 2004;172:6902-12.
96. Higai K, Imaizumi Y, Suzuki C, et al. NKG2D and CD94 bind to heparin and sulfate-containing polysaccharides. *Biochemical and biophysical research communications*. 2009;386:709-14.
97. Higai K, Suzuki C, Imaizumi Y, et al. Binding affinities of NKG2D and CD94 to sialyl Lewis X-expressing N-glycans and heparin. *Biological & pharmaceutical bulletin*. 2011;34:8-12.
98. Xin X, Higai K, Imaizumi Y, et al. Natural killer group 2A (NKG2A) and natural killer group 2C (NKG2C) bind to sulfated glycans and α 2,3-NeuAc-containing glycoproteins. *Biological & pharmaceutical bulletin*. 2011;34:480-5.
99. Higai K, Matsumoto S, Kimura M, et al. NKG2D: Binding Properties for Glycan Ligands, and Mutagenesis Analysis. *The Open Biotechnology Journal*. 2011;5:33-38.
100. Keren H, Lev-Maor G, Ast G. Alternative splicing and evolution: diversification, exon definition and function. *Nature reviews. Genetics*. 2010;11:345-55.
101. Artamonova II, Gelfand MS. Comparative genomics and evolution of alternative splicing: the pessimists' science. *Chemical reviews*. 2007;107:3407-30.
102. Ghigna C, Valacca C, Biamonti G. Alternative splicing and tumor progression. *Current genomics*. 2008;9:556-70.

103. Revil T, Gaffney D, Dias C, et al. Alternative splicing is frequent during early embryonic development in mouse. *BMC genomics*. 2010;11:399.
104. Jelen N, Ule J, Zivin M, et al. Evolution of Nova-dependent splicing regulation in the brain. *PLoS genetics*. 2007;3:1838-47.
105. Graveley B. Mutually exclusive splicing of the insect Sscam pre mRNA directed by competing intronic RNA secondary structures. *Cell*. 2005;123:65-73.
106. Wang ET, Sandberg R, Luo S, et al. Alternative Isoform Regulation in Human Tissue Transcriptomes. *Nature*. 2009;456:470-476.
107. Pritsker M, Doniger TT, Kramer LC, et al. Diversification of stem cell molecular repertoire by alternative splicing. *Proceedings of the National Academy of Sciences*. 2005;102:14290-5.
108. Belicha-Villanueva A, Blickwedehl J, McEvoy S, et al. What is the role of alternate splicing in antigen presentation by major histocompatibility complex class I molecules? *Immunologic research*. 2010;46:32-44.
109. Leeman JR, Gilmore TD. Alternative splicing in the NF- κ B signaling pathway. *Gene*. 2008;423:97-107.
110. Lynch KW. Consequences of regulated pre-mRNA splicing in the immune system. *Nature reviews. Immunology* [electronic article]. 2004;4:931-40.
111. Yu J, Wei M, Mao H, et al. CD94 Defines phenotypically and functionally distinct mouse NK cells subsets. *Journal of immunology*. 2009;183:4968-4974.
112. Hayakawa Y, Huntington ND, Nutt SL, et al. Functional subsets of mouse natural killer cells. *Immunological reviews*. 2006;214:47-55.
113. Freud AG, Yokohama A, Becknell B, et al. Evidence for discrete stages of human natural killer cell differentiation in vivo. *The Journal of experimental medicine*. 2006;203:1033-43.

114. Freud AG, Caligiuri M a. Human natural killer cell development. *Immunological reviews*. 2006;214:56-72.
115. Vivier E, Raulet DH, Moretta A, et al. Innate or adaptive immunity? The example of natural killer cells. *Science*. 2011;331:44-9.
116. Ward AJ, Cooper TA. The pathobiology of splicing. *J Phatol*. 2010;220:152-163.
117. Evsyukova I, Somarelli J a, Gregory SG, et al. Alternative splicing in multiple sclerosis and other autoimmune diseases. *RNA biology*. 2010;7:462-73.
118. Birzele F, Csaba G, Zimmer R. Alternative splicing and protein structure evolution. *Nucleic acids research*. 2008;36:550-8.
119. Xing Y, Lee C. Evidence of functional selection pressure for alternative splicing events that accelerate evolution of protein subsequences. *Proceedings of the National Academy of Sciences*. 2005;102:13526-31.
120. Calarco JA, Xing Y, Cáceres M, et al. Global analysis of alternative splicing differences between humans and chimpanzees. *Genes & Development*. 2007;29:63-2975.
121. Belov K, Sanderson CE, Deakin JE, et al. Characterization of the opossum immune genome provides insights into the evolution of the mammalian immune system. *Genome research* [electronic article]. 2007;17:982-91.
122. Wong ESW, Sanderson CE, Deakin JE, et al. Identification of natural killer cell receptor clusters in the platypus genome reveals an expansion of C-type lectin genes. *Immunogenetics*. 2009;61:565-79.
123. Gagnier L, Wilhelm BT, Mager DL. Ly49 genes in non-rodent mammals. *Immunogenetics*. 2003;55:109-15.
124. Barten R, Trowsdale J. The human Ly-49L gene. *Immunogenetics*. 1999;49:731-734.

-
125. Staley JP WJ. Assembly of ribosomes and spliceosomes: complex ribonucleoprotein machines. *Current opinion in cell biology* 2009;21:109-118.
 126. Hsu S-ning, Hertel KJ. Splicing errors and their impact on cellular function. *Rna Biology*. 2009;6(5):526-530.
 127. Müller S, Zocher G, Steinle A, et al. Structure of the HCMV UL16-MICB complex elucidates select binding of a viral immunoevasin to diverse NKG2D ligands. *PLoS pathogens*. 2010;6:e1000723.