



UNIVERSIDAD NACIONAL DE LA PLATA

FACULTAD DE CIENCIAS EXACTAS

DEPARTAMENTO CIENCIAS BIOLÓGICAS

Trabajo de Tesis Doctoral:

Estudio genómico y antigénico de una cepa virulenta de virus Junín

Tesista: Pablo Daniel Thomas

Directores: Ricardo Martín Gómez

María Florencia Ferrer

Año: 2024

Agradecimientos

A los que me brindaron el ámbito y la financiación para realizar mis estudios de posgrado, incluyendo: el Instituto de Biotecnología y Biología Molecular, el Departamento de Ciencias Biológicas, la Facultad de Ciencias Exactas, la Universidad Nacional de La Plata, y el Consejo Nacional de Investigaciones Científicas y Técnicas.

A mi director Dr. Ricardo Gómez, por dirigir este trabajo. Por la orientación a lo largo del doctorado, la disponibilidad y presencia constante, timonear el doctorado durante la tempestuosa situación económica del país y la pandemia, y su atención al detalle.

A mi directora Dr. María Florencia Ferrer, por guiarme y brindar su conocimiento en el laboratorio, su asistencia cuando la requería y traducir por Ricardo cuando fue necesario. Agradezco su paciencia, por sobre todo.

A los Dres. Nora López, Cybele García, Antonio Carrera Silva, Carla Tomatis y Mauricio Lozano, por su ayuda en la realización de este trabajo.

A las técnicas de laboratorio Silvana Tongiani y Larisa Vasquez por su asistencia y darme la oportunidad de aprender de ellas en los comienzos de mi doctorado.

A mis primeros compañeros de laboratorio, aquellos que me recibieron y me ayudaron al comienzo, como Tomás Masón y Laura Fabre. Agradezco a Paula Arrías por tutorarme en toda nueva técnica y ser alguien con quien se podía hablar de cualquier cosa, como también por la amistad que brindó. A todos ellos, le dedico muchas gracias por hacerme sentir bienvenido y el aprendizaje que me dieron.

Al Dr. Matías Pidre, al comienzo de mi trabajo otro compañero de laboratorio y quien brindó humor además de su ayuda y experticia, y la Dra. Leticia Ferrelli, quien dio su asistencia y apoyo cuando lo requería.

Con especial cariño quiero dar mi agradecimiento a mis compañeros de laboratorio con quienes finalizo el trabajo: Santiago Gómez Bergna, Abril Marchesini, Silvia Aquila y María Luján Scalise. Siempre los recordaré como amigos y excelentes compañeros de laboratorio, brindando ayuda profesional y haciendo de momentos muertos en el laboratorio más llevaderos, sin hacer mención de los buenos ratos y compañía fuera del instituto. Junto a ellos, quiero dar mi agradecimiento a Leslie Amorós, una amiga y compañera con quien hemos pasado buenos y malos ratos en el laboratorio, y quien con he contado una amistad casi constante desde el inicio de mi doctorado, y sin mencionar sus mates especiales™.

A Lucas Bacigalupo, a quien tuve el gusto de ayudar en su tesina de grado y formación académica, y quien me dio el honor de ser tutor y darme mi primera experiencia enseñando. Agradezco su amistad, y por ser alguien con quien se puede compartir cosas nerd.

Finalmente, agradezco a mi familia. A mi hermana, Amalia, quien pudo mandar su afecto y apoyo inclusive ante la distancia junto a James. Agradezco particularmente a mi madre y padre, Beatriz y Luis Thomas, por darme todo lo que necesitara y más, por preocuparse y ayudarme en los quehaceres del día a día, por siempre mantener su buena disposición y dar palabras de apoyo, y por ser sostén inquebrantable durante mi doctorado en todos los aspectos de mi vida.

Para concluir, quiero dar gracias especiales a mi hermano Giuliano, un amigo con quien hemos charlado de las locuras que más nos interesan, compartido chistes, discutido desde estupideces hasta filosofía, aventurándonos en un sinfín de mundos virtuales y con quien siempre he podido compartir mis hobbies. Agradezco su tiempo online casi todas las noches de mi doctorado y ser el mejor hermano que pude haber pedido ¡Gracias Giuliano!

Resumen

El virus Junín (JUNV), actualmente clasificado dentro del género *Mammarenavirus*, es el agente etiológico de la fiebre hemorrágica Argentina (FHA). Así como existen algunos mammarenavirus asociados enfermedad humana, también existen cepas de JUNV atenuadas o patogénicas. La cepa de JUNV atenuada mejor caracterizada es la cepa Candid 1 (C#1), de secuencia conocida, y utilizada en la vacuna para prevenir la FHA. En este trabajo se estudió la cepa patogénica P3441 de JUNV, que junto a la cepa C#1 han sido utilizadas en estudios previos donde se ha observado que suscitan distintas respuestas celulares. Con el objetivo de lograr una mayor comprensión de la base molecular de las diferencias observadas entre estas cepas y profundizar nuestro entendimiento de los mecanismos de patogenicidad del JUNV, se obtuvo la secuencia de la cepa P3441 y se realizó un análisis comparativo con otras cepas, incluyendo C#1. Adicionalmente, se clonaron las proteínas virales de la cepa P3441 en un vector bajo control de un promotor procariota para después subclonar cada una de ellas en vectores de expresión bajo un promotor eucariota de forma tal que se expresen unidas a fluoróforos reporteros. Se realizaron estudios de la expresión de las proteínas recombinantes por inmunofluorescencia y microscopía confocal. Posteriormente, con excepción de la proteína L, donde se carece de antisuero, se estudió la reactividad antigénica de las proteínas recombinantes con anticuerpos específicos. Finalmente, a manera de profundizar en la caracterización molecular de la cepa, como JUNV en general, se estudió el sesgo de uso de codones. Los resultados obtenidos y las herramientas generadas sientan la base para profundizar nuestra comprensión de las bases moleculares de la patogenicidad de JUNV.

Índice

Contenido

Agradecimientos	3
Resumen.....	5
Abreviaturas	10
Introducción	12
A. Fiebre Hemorrágica Argentina.....	12
A.1 Aspectos históricos	12
A.2 Epidemiología y transmisión	13
A.3 Historia natural	14
A.4 Diagnóstico.....	14
A.5 Tratamiento	14
A.6 Prevención	15
A.7 Modelos experimentales	15
B. Virus Junín	17
B.1 Clasificación.....	17
B.2 Estructura y composición	20
B.3 Organización genómica	21
B.4 Ciclo de replicación	22
B.5 Aislamientos y cepas	24
Hipótesis y objetivos	28
Hipótesis de trabajo	28
Objetivo General	28
Objetivos Particulares	28
Capítulo 1. Secuenciación y clonados	29
1.1 Preparación y diseños	29
1.2 Resultados	29
1.2.1. Amplificación y secuenciación.....	29
1.2.2 Clonado de las proteínas virales.....	30
1.3 Discusión	32
Capítulo 2. Proteína Z.....	34
2.1 Introducción	34

2.1.1 Estructura de la proteína Z.....	34
2.1.2 Replicación viral	35
2.1.3 Brotación y ensamblado viral.....	35
2.1.4 Regulación de funciones celulares mediada por Z.....	36
2.1.5 Evasión de la respuesta inmune innata.....	37
2.2 Resultados	37
2.2.1 Secuencia.....	37
2.2.2 Clonado en vector de expresión eucariota	39
2.2.3 Localización, interacciones y antigenicidad	40
2.3 Discusión	41
Capítulo 3. Proteína N	43
3.1 Introducción	43
3.1.1 Estructura de la proteína N	43
3.1.2 Replicación viral	43
3.1.3 Regulación de funciones celulares mediada por N	44
3.1.4 Evasión de la respuesta inmune innata.....	44
3.2 Resultados	45
3.2.1 Secuencia.....	45
3.2.2 Generación de vectores eucariotas de N	49
3.2.3. Localización, interacciones y antigenicidad	51
3.3 Discusión	52
Capítulo 4. Glicoproteínas de superficie	54
4.1 Introducción	54
4.1.1 Brotación y ensamblado viral.....	55
4.1.2 Determinantes moleculares de atenuación	55
4.2 Resultados	56
4.2.1 Secuencia.....	56
4.2.2 Generación de vectores de expresión eucariotas para GPC.....	63
4.2.3. Localización e interacciones.....	64
4.3 Discusión	65
Capítulo 5. Proteína L.....	67
5.1 Introducción	67
5.1.1 Estructura de la proteína L.....	67

5.1.2 Replicación viral	68
5.1.3 Virulencia.....	69
5.2 Resultados	69
5.2.1 Secuencia.....	69
5.2.2 Generación de un vector de expresión eucariota de L	75
5.3 Discusión	77
Capítulo 6. Sesgo de uso de codones	79
6.1 Introducción	79
6.2 Resultados	80
6.2.1 Composición.....	80
6.2.2 Análisis del uso relativo de codones sinónimos.....	81
6.2.3 Número efectivo de codones	82
6.2.4 Adaptación y deoptimización relativa de codones	84
6.2.5 Ensayo de correspondencia	89
6.2.6 Ensayo de correlación	90
6.3 Discusión	94
Discusión general y conclusiones.....	97
Materiales y Métodos	102
Células	102
Virus.....	102
Ensayo de unidades formadoras de placas	102
Transcripción reversa y reacción en cadena de la polimerasa.....	102
Secuenciación.....	103
Desfosfatación.....	103
Clonado	103
Minipreparación de plásmidos por el método de lisis alcalina.....	104
Purificación de ADN por extracción con fenol-cloroformo	104
Transfección	105
Ensayos de inmunofluorescencia	105
Microscopía confocal	105
Secuencias	105
Programas y servicios informáticos.....	105
Análisis Estadístico	106

Anexo Materiales y métodos	107
Tabla 24. Anticuerpos primarios utilizados para inmunofluorescencia.....	107
Tabla 25. Anticuerpos secundarios utilizados para inmunofluorescencia.....	107
Tabla 26. Oligonucleótidos para clonado de vectores y proteínas recombinantes.....	107
Tabla 27. Oligonucleótidos utilizados para la secuenciación.....	108
Tabla 28. Enzimas de Restricción utilizadas para el clonado de proteínas virales.....	108
Anexo RSCU	109
Tabla RSCU GPC. La frecuencia relativa de uso de codón sinónimo (RSCU) por aminoácido (AA) y cepa de JUNV en GPC.	109
Tabla RSCU N. La frecuencia relativa de uso de codón sinónimo (RSCU) por aminoácido (AA) y cepa de JUNV en N.	110
Tabla RSCU L. La frecuencia relativa de uso de codón sinónimo (RSCU) por aminoácido (AA) y cepa de JUNV en L.....	112
Tabla RSCU Z. La frecuencia relativa de uso de codón sinónimo (RSCU) por aminoácido (AA) y cepa de JUNV en Z.....	113
Bibliografía	116

Abreviaturas

µg: microgramo

µl: microlitro

ADN: ácido desoxirribonucleico

ADNc: ácido desoxirribonucleico complementario

ARN: ácido ribonucleico

ARNdc: ácido ribonucleico de doble cadena

ARNm: ácido ribonucleico mensajero

ARNt: ácido ribonucleico de transferencia

CAP: caperuza

DC-SIGN: molécula de adhesión intercelular específica de células dendríticas-3 no unidora a integrina

Dpi: días posinfección

eIF4E: factor de iniciación de traducción eucariota 4 E

ESCRT: complejos de clasificación endosómicos necesarios para el transporte

FHA: Fiebre Hemorrágica Argentina

GPC: complejo de glicoproteínas

hDC-SIGN: DC-SIGN humano

ICTV: *International Committee on Taxonomy of Viruses*

IFN: interferón

IKKε: quinasa ε del inhibidor del factor κB

JUNV: virus Junín

kDa: kilodaltons

LASV: virus de Lassa

LCMV: virus de la coriomeningitis linfocitaria

LSEctin: Lectina de células endoteliales sinusoidales del hígado

MACV: virus Machupo

MAVS: del inglés *Mitochondrial antiviral-signaling protein*

MDA5: proteína 5 asociada a la diferenciación del melanoma

ml: mililitro

NB: del inglés *Nuclear Bodies*, o cuerpos nucleares

NW: del inglés *New World*, o Nuevo Mundo

OW: del inglés *Old World*, o Viejo Mundo

PA: polimerasa ácida de influenza

PB2: polimerasa básica 2 de influenza

PCR: reacción de polimerasa en cadena

PKR: proteína quinasa R

PML: proteína de la leucemia promielocítica

PNH: primates no humanos

RE: retículo endoplasmático

RIG-I: del inglés *Retinoic acid-Inducible Gene 1*

RING: del inglés, *Really interesting new gene*, o gen nuevo muy interesante

RLR: receptores tipo RIG-I

RNP: ribonucleoproteínas
RNT: región no traducida
RTC: complejos de replicación-transcripción
RT-PCR: PCR con transcriptasa reversa
S1P: proteasa de sitio 1
SKI-1: isoenzima-1 queixina subtilisina
SNT: síndrome neurológico tardío
TBK1: quinasa unidora de TANK 1
TfR1: receptor de transferrina 1
Tm: del inglés *melting temperature*, o temperatura de fusión
TNF: factor de necrosis tumoral
Tyro3: Receptor de tirosina-proteína quinasa 3
U: unidades de actividad enzimática
VLPs: del inglés *virus-like particles*, o partículas similares a virus

Introducción

A. Fiebre Hemorrágica Argentina

A.1 Aspectos históricos

La Fiebre Hemorrágica Argentina (FHA), también conocida como mal de los rastros, es una enfermedad endemo-epidémica de tipo zoonótica y origen viral que inicialmente afectaba principalmente a trabajadores rurales en la región central de la pampa húmeda de la República Argentina [revisado en (1)]. Esta enfermedad fue reconocida por primera vez en 1955 cuando Arribalza describió los aspectos clínicos de una nueva enfermedad de etiología desconocida que había aparecido en el noroeste de la Provincia de Buenos Aires (2). La recurrencia de esta enfermedad febril hemorrágica en Junín en 1958 permitió a los investigadores realizar estudios adicionales en los que se basaron para reportar que el agente causal era un virus transmitido por artrópodos, el cual fue denominado virus Junín (JUNV) por la ciudad donde fue identificado por primera vez (3). Esas observaciones fueron confirmadas por otros investigadores (4). Algunas propiedades del agente fueron investigadas por Greenway y colaboradores, quienes informaron que ciertos animales de laboratorio, como cobayos y ratas y ratones recién nacidos, eran susceptibles al virus; sin embargo, no pudieron cultivar el virus *in vitro* (5). Recién en 1961, JUNV pudo ser propagado por Mettler y colaboradores en cultivos de células HeLa (6) y sus antígenos ser detectados en células cultivadas por técnicas de inmunofluorescencia (7). Posteriormente, se confirmó que no había transmisión por artrópodos y que el roedor *Calomys musculinus* era el principal reservorio natural (8), aunque también se observó que el virus podía infectar otros cricétidos, como por ejemplo el *Akodon azarae* y *Calomys callidus*, que habitan zonas limítrofes a las del reservorio natural (9, 10).

Hacia 1970, utilizando técnicas de fijación de complemento e inmunofluorescencia indirecta, se había ubicado a JUNV dentro del grupo de arenavirus serológicamente relacionados al virus Tacaribe (TACV)(11), habiéndose logrado confirmar también similitudes en las estructuras observadas por microscopía electrónica de transmisión (12). Para 1976 se había logrado determinar que JUNV era un virus ARN (13), y que pertenecía al creciente grupo de los arenavirus, confirmándose similitudes en la composición de ARN viral purificado, estructuras observadas por microscopía electrónica de transmisión y características de infección en modelos animales (5, 14, 15). Fue también en 1976 que se identificaron seis polipéptidos estructurales distintivos por medio de electroforesis en gel de poliacrilamida, reconociéndose una proteína de nucleocápside y una glicoproteína de envoltura (16).

Cuando la enfermedad fue reportada por primera vez, la extensión del área afectada era aproximadamente de 16.000 km² con un número estimado de 270.000 personas en riesgo (17). Desde entonces, la zona endémica se ha extendido en forma ininterrumpida y actualmente incluye parte de las provincias de Buenos Aires, Entre Ríos, La Pampa, Santa Fe y Córdoba, estimándose que la población en riesgo de infección es de más de 5 millones de personas (18, 19)(Figura 1).

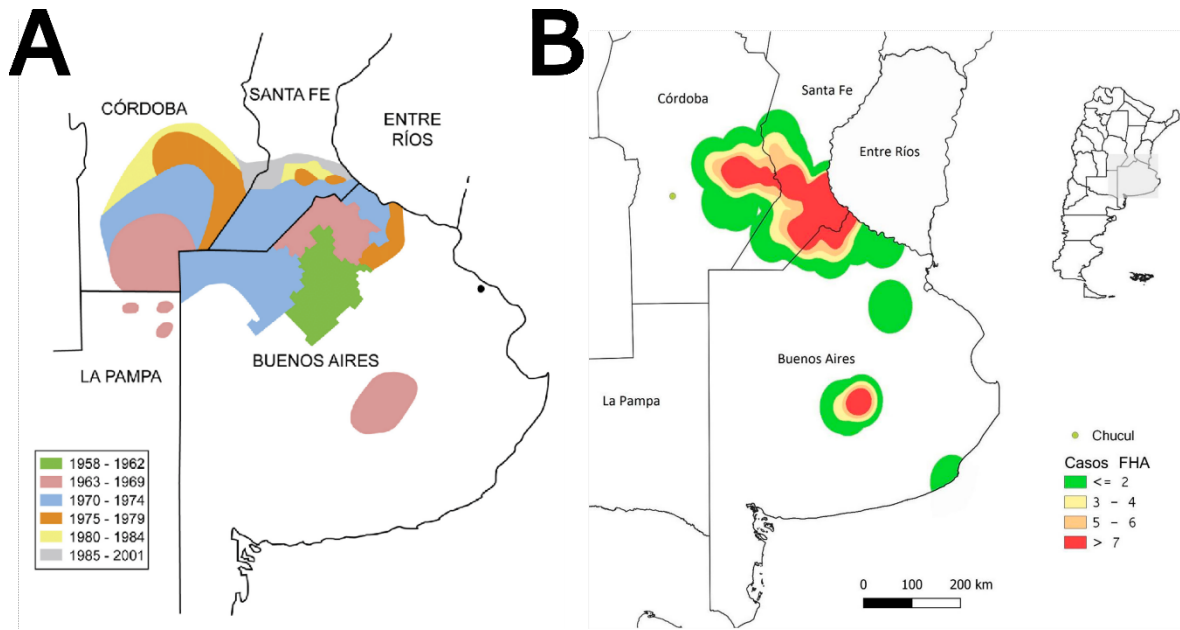


Figura 1. A) Evolución de la zona endémica desde su descubrimiento, adaptado de bibliografía (344). B) Zona endémica y número de casos desde 2010 al 2020 de JUNV en la República de Argentina, modificado de (18).

A.2 Epidemiología y transmisión

El vector natural y reservorio del virus es el roedor *Calomys musculinus*, aunque también se ha encontrado en otros roedores *Calomys* (20, 21). La distribución y comportamiento del roedor explica la zona endémica de esta enfermedad (22). Es probable que JUNV se mantenga en las poblaciones reservorio de *C. musculinus* principalmente a través de la transmisión horizontal. Cuando se transmite verticalmente, el virus puede tener importantes efectos nocivos sobre la formación física de la cría. Los *C. musculinus* recién nacidos infectados con JUNV en el laboratorio tienen alta mortalidad (>70%), y los sobrevivientes muestran retraso en el crecimiento y una fecundidad muy reducida. En contraste, los animales adultos infectados en el laboratorio muestran una respuesta dividida. La mitad de los animales desarrollan anticuerpos y eliminan el virus; mientras que el resto desarrolla viremia crónica y eliminación persistente del virus en la orina y la saliva, sin ningún efecto nocivo evidente. Estudios de campo de poblaciones naturales demostraron que la infección con JUNV entre *C. musculinus* silvestres fue más frecuente entre los machos que entre las hembras, correlacionándose positivamente con la edad y la presencia de heridas o cicatrices. Estos datos sugieren que los encuentros agresivos entre los machos adultos pueden ser un mecanismo importante de transmisión dentro poblaciones del reservorio [revisado en (23), <http://www.expertconsult.com>].

La transmisión a seres humanos se da principalmente a partir de los fluidos del animal infectado, tales como la orina, saliva y sangre. La infección ocurre por contacto con material contaminado, a través de escoriaciones de la piel o por las mucosas oral o conjuntival, o bien la inhalación de aerosoles (Weissenbacher et al., 1987)(Figura 2). Infrecuentemente, puede también existir infección a partir de humanos infectados [revisado en (23)].

A.3 Historia natural

El periodo de incubación de la FHA varía entre 6 y 12 días, al final del cual se manifiestan los síntomas iniciales. Durante la primera semana, los síntomas se asemejan a una gripe pudiendo incluir fiebre, mialgia, artralgia, dolor de cabeza, náuseas, vómito y diarrea. De producirse hemorragias, estas suelen ocurrir durante la segunda semana y se limitan a ser nasales, conjuntivas oculares y/o en las encías, aunque en casos severos y menos comunes se producen hematomas, metrorragia, hematuria, melena y otros síntomas hemorrágicos. En casos extremos, los pacientes pueden experimentar shock y convulsiones [revisado en (24)]. Adicionalmente, pueden aparecer síntomas neurológicos. El paciente puede presentarse irritable, letárgico y con un ligero temblor en la mano y la lengua. Pueden encontrarse presentes otros síntomas como ataxia moderada, hiperestesia cutánea y una disminución de los reflejos tendinosos profundos y la tonicidad muscular (25). Después de la segunda semana, más del 80% de los pacientes se recuperan (Figura 2). Sin tratamiento, la tasa de mortalidad es aproximadamente del 30% [revisado en (1)].

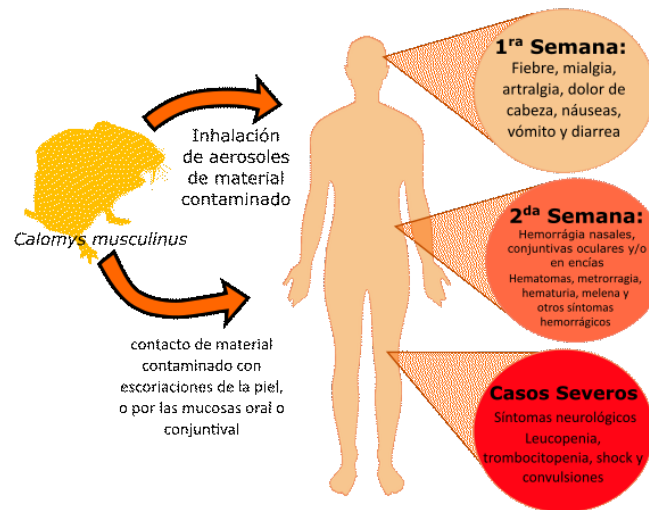


Figura 2. Resumen esquemático de infección y sintomatología de JUNV en seres humanos.

A.4 Diagnóstico

El diagnóstico de la FHA se basa en información clínica y de laboratorio. Un número de plaquetas menor a $100.000/\text{mm}^3$ en combinación con un número de leucocitos en sangre inferior a $2.500/\text{mm}^3$ para pacientes en áreas endémicas se consideran indicadores útiles para identificar individuos infectados (26). La confirmación puede establecerse mediante un ensayo RT-PCR (del inglés, *reverse transcription-polymerase chain reaction*)(27).

Estudios clínicos han revelado que pacientes de FHA a menudo presentan niveles elevados de interferón (IFN) tipo I (IFN- $\text{I}\alpha$) en suero (2.000 a 64.000 UI/ml) (28), que correlacionan positivamente con el agravamiento de la evolución clínica (29).

A.5 Tratamiento

El tratamiento se basa en medidas que mantengan la homeostasis del paciente. La transfusión de plasma de pacientes convalecientes en dosis de 3.000 unidades/kg peso de anticuerpos neutralizantes durante la fase inicial de la FHA reduce la mortalidad a menos del 1% (30). Alrededor

del 10% de los casos tratados con plasma inmune desarrollan un síndrome neurológico tardío (SNT). Este SNT comienza después de un período libre de síntomas y se caracteriza por síntomas febriles, signos cerebelosos y parálisis de los nervios craneales (31). El SNT se ha registrado exclusivamente en pacientes tratados con plasma inmune, la patogénesis no está clara y no está relacionada con el donante, la gravedad de la enfermedad, el título de anticuerpos en el plasma transfundido u otras variables examinadas (31). El compuesto antiviral no específico ribavirin, ha mostrado eficacia cuando se administra en estadios tempranos de la infección (30).

A.6 Prevención

En la década de 1990 se introdujo una vacuna contra la FHA, que emplea la cepa atenuada Candid 1 (C#1). Esta vacuna resultó muy efectiva para reducir la tasa de infección y prevenir el desarrollo de la FHA (32). Sin embargo, la vacuna no puede emplearse en embarazadas ni inmunodeprimidos por ser a virus atenuado, y la FHA persiste como una infección potencialmente letal y por lo tanto reviste interés nacional. La vacuna Candid 1 fue incorporada al Calendario Nacional de Vacunación en el año 2007 (Resolución 48/2007) después de registrarse en la ANMAT (Registro Nº 53205) tras estudios en humanos que demostraron una eficiencia superior al 95,5% de estudios previos (33). Previamente, la vacuna era producida en el Instituto Salk, Pensilvania, Estados Unidos. Actualmente es producida en el Instituto Nacional de Enfermedades Virales Humanas “Dr. Julio I. Maiztegui” (INEVH) y es recomendada para mayores de 15 años que residan o desarrollen actividades en el área endémica de la enfermedad que no hayan recibido la vacuna anteriormente. La vacunación debe llevarse a cabo por lo menos un mes antes de la realización de actividades de riesgo para asegurar la protección. Información adicional se puede encontrar en los Lineamientos Técnicos de la Vacunación contra la Fiebre Hemorrágica Argentina Candid 1 publicada por el Ministerio de Salud Argentina y elaborada por la Dirección de Control de Enfermedades Inmunoprevenibles (DiCEI) y el Instituto Nacional de Enfermedades Virales Humanas (INEVH).

A.7 Modelos experimentales

A.7.1. Modelos animales

Muchos de los estudios de infección por mammarenavirus han utilizado cobayos (*Cavia porcellus*) como modelo experimental. Este animal es particularmente susceptible a la infección por JUNV ya que se ha establecido 1-2 UFP como dosis letal media (DL_{50}) para algunas cepas de JUNV. Entre las cepas de cobayos utilizadas se encuentran la cepa endogámica 13 y la cepa exogámica Hartley (34). La infección de JUNV no presenta diferencias significativas entre esas cepas de cobayos y muestra similitudes respecto a la infección en humanos tales como la presencia de trombocitopenia y leucopenia (35). También, en forma similar a los humanos, la infección de cobayos produce IFN- λ que aumenta a medida que la infección progresa (36).

Los ratones adultos son refractarios a la infección por JUNV (37). En contraste, son susceptibles los ratones neonatos. La cepa C3H/HeJ es susceptible por inoculación intracerebral hasta los 5 meses de edad (38). No obstante, la enfermedad es diferente a la observada en humanos. El modelo de ratón neonatal ha sido empleado frecuentemente para determinar factores contribuyentes a la atenuación de arenavirus del Nuevo Mundo (ver B1 Clasificación) (39). También se han utilizado ratones manipulados genéticamente (*knockout*) adultos, como los dobles *knockout* para el receptor de IFN-I

e IFN-II (40). Varios primates no humanos (PNH) adultos han sido reportados como susceptibles, incluyendo macacos Rhesus (*Macaca mulatta*) (41, 42) y el mono tití (*Callithrix jacchus*) (43).

A.7.2 Modelos celulares

Como se mencionó, JUNV fue aislado por primera vez en cultivos de células HeLa (6). Otras líneas celulares también fueron tempranamente usadas en estudios con JUNV, como las células Vero utilizadas para el método de titulación por formación de placas (44). Por otro lado, un extenso número y variedad de cultivos primarios han sido utilizados para estudiar aspectos específicos de la infección de JUNV, incluyendo fibroblastos murinos (45), células derivadas de cerebro murino (15), células de medula ósea (46), células endoteliales humanas derivadas de cordón umbilical (47, 48), astrocitos de rata (49) y de ratón (50), células cardíacas (19), progenitores hematopoyéticos humanos (51), células dendríticas plasmacitoides humanas (52), y monocitos y macrófagos humanos (53). La infección de macrófagos primarios de cobayo produce una mayor replicación de JUNV en comparación con la infección de macrófagos de hámster o ratón (54). La explicación se obtuvo posteriormente empleando virus de la leucemia murina que expresan glicoproteínas de JUNV u otros mamarenavirus patógenos, donde se demostró que el receptor de transferrina 1 de cobayo (gpTfR1) es el principal receptor de JUNV, mientras que los ortólogos de hámster y ratón no son compatibles con la entrada / infección del virus (54, 55).

La línea celular epitelial A549 derivada de adenocarcinoma de pulmón humano exhibe varias características estructurales y ultraestructurales de los neumocitos tipo II. En la última década, se la ha utilizado con frecuencia como modelo en estudios de JUNV, incluyendo apoptosis (56), exclusión por superinfección (57), respuestas de interferón (58), autofagia (59, 60) y la formación de cuerpos nucleares (61).

B. Virus Junín

B.1 Clasificación

Actualmente, la familia *Arenaviridae* está constituida por cuatro géneros denominados *Mammarenavirus*, *Reptarenavirus*, *Antennavirus* y *Hartmanivirus* cuyas características y hospedantes han sido determinados (Tabla 1), con un quinto género denominado *Innmovirus* constituido actualmente por una única especie cuya biología permanece desconocida (63). Los virus de los cinco géneros comparten una o más de las siguientes características: (a) morfología esférica o pleomórfica; (b) genoma de ARN monocatenario *ambisense* sin extremos 3' poliadenilados; (c) complementariedad de secuencia genómica en los extremos 5' y 3'; (d) secuencias nucleotídicas que pueden formar una o más estructuras de horquilla estables en la región intergénica; (e) ARN mensajeros con CAP pero sin poliadenilación; (f) e inducción de infección persistente (frecuentemente asintomática) en los reservorios naturales (dónde fueron identificados), en la cual ocurre viremia crónica y/o viruria (63).

Tabla 1. Resumen de los cuatro géneros caracterizados de la familia *Arenaviridae*.

Género	Proteínas virales	Particularidad	Genoma (aprox.)	Reservorio	Virus modelo [Especie ICTV]
<i>Antennavirus</i>	N L SSP, G1 y G2	No codifica proteína Z	3 segmentos: S 2090 nts M 3260 nts L 6500 nts	<i>Antennarius striatus</i> (pez sapo estriado)	Wënling frogfish arenavirus 1 (WIFV-1) [<i>Antennavirus striale</i>]
<i>Hartmanivirus</i>	N L SSP, G1 y G2	No codifica proteína Z Produce enfermedad de cuerpos de inclusión en reservorio	2 segmentos: S 3370 nts L 5950 nts	<i>Boa constrictor</i>	Virus de serpiente del Instituto Haartman 1 (HISV-1) [<i>Hartmanivirus haartmani</i>]
<i>Reptarenavirus</i>	N L Z G1 y G2	GP2 más similar a ebolavirus que miembros de bunyavirus El complejo de glicoproteína no cuenta con SSP Produce enfermedad de cuerpos de inclusión en reservorio	2 segmentos: S 3500 nts L 6900 nts	<i>Boa constrictor</i> Boas anuladas	Virus Golden Gate (GOGV) [<i>Reptarenavirus aurei</i>]
<i>Mammarenavirus</i>	N L Z SSP, G1 y G2	Primer grupo identificado de la familia <i>Arenaviridae</i> Varias especies infecta humano y producen fiebres hemorrágicas	2 segmentos: S 3500 nts L 7200pb	Roedores, típicamente de la superfamilia <i>Muroidea</i> *	Virus de la coriomeningitis linfocítica (LCMV) [<i>Mammarenavirus choriomeningitidis</i>]

*Única excepción es TACV encontrado en murciélagos filostómidos y garrapatas estrella solitaria.

Los géneros *Reptarenavirus* y *Hartmanivirus* han surgido del análisis metagenómico, secuenciación y subsecuente aislamiento de nuevos arenavirus, tales como el agente etiológico de la Enfermedad de Cuerpos de Inclusión que afecta a serpientes *alethinophidian* en cautiverio (64-66). El género *Hartmanivirus* cuenta con especies cuyos genomas no codifican para la proteína Z o algún homólogo a esa proteína encontrada en *mammarenavirus* y *reptarenavirus* (67).

El género *Reptarenavirus* tiene la característica de contar con glicoproteínas GP2 relacionadas más próximamente a las glicoproteínas de los ebolavirus (orden Mononegvirales, familia *Filoviridae*) que aquellas de otros virus del orden *Bunyavirales*, incluyendo los otros géneros de los arenavirus.

Adicionalmente, el complejo de glicoproteínas que conforman las espículas de superficie de los virus de este género no cuenta con un péptido señal pequeño (SSP) asociado (68). Curiosamente, a diferencia de otros arenavirus, presentan una propensión a tener coinfecciones con múltiples segmentos genómicos S y L sin necesariamente ser en relación 1:1, aunque esto puede atribuirse a condiciones de cautiverio de las serpientes infectadas (69-71).

Se han identificado dos arenavirus trisegmentados que infectan pez sapos (*Antennarius striatus*). Estos han sido propuestos como miembros del género más reciente *Antennavirus* (72). Como en el caso del género *Hartmanivirus*, el genoma no codifica la proteína Z.

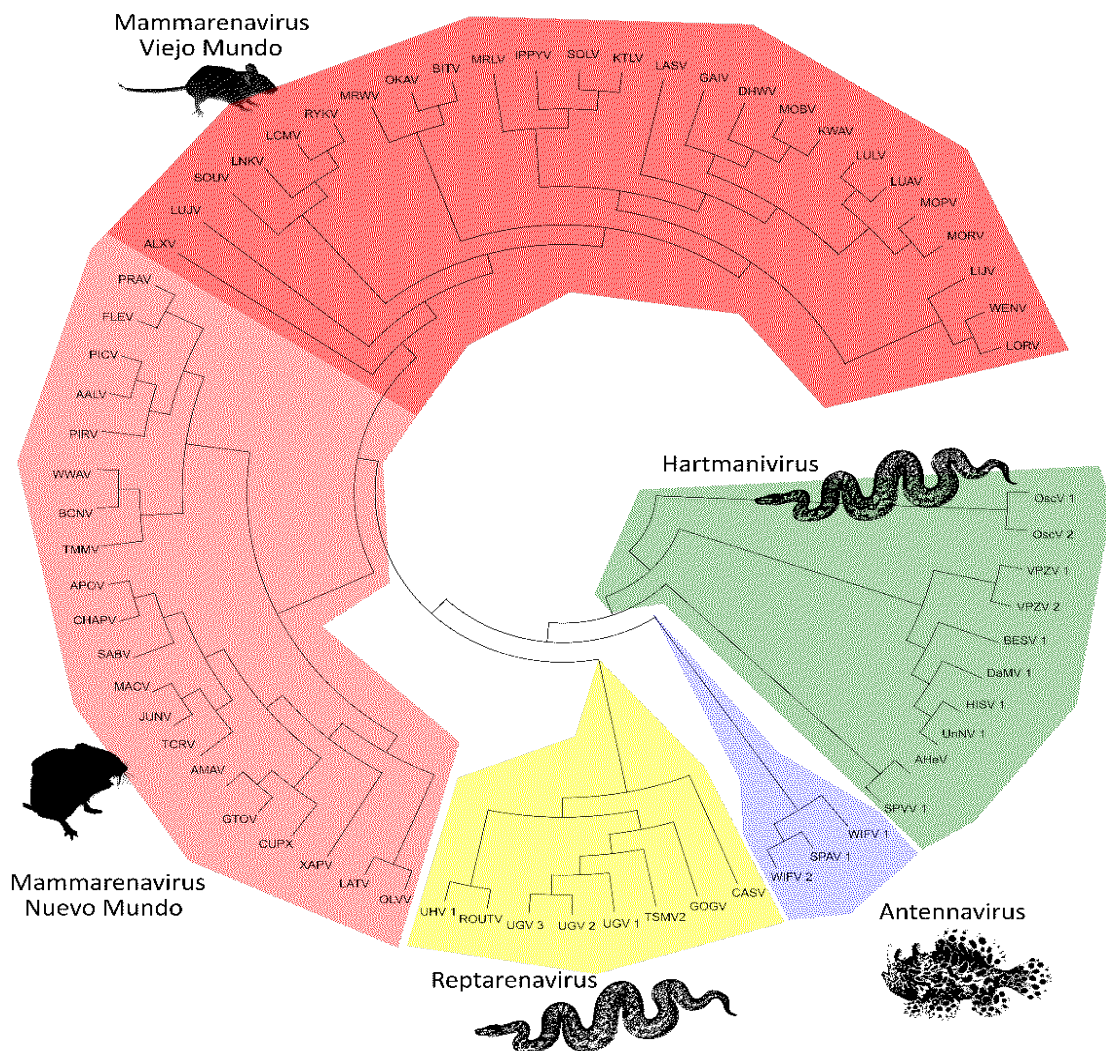


Figura 3. Árbol filogenético de la familia Arenaviridae basados en el segmento L. Se construyó en base a los arenavirus confirmados por el ICTV 2019 (63, 77) utilizando NGPhylogeny.

El género *Mammarenavirus* debe su nombre a su infectividad en mamíferos, y está dividido en dos grupos monofiléticos: Nuevo Mundo (*New World* o NW) y Viejo Mundo (*Old World* o OW). Los reservorios naturales de estos virus son en su mayoría roedores. En casi todos los casos, los roedores reservorio pertenecen a la superfamilia *Muroidea*, y la distribución geográfica de cada virus está generalmente restringida a aquella de sus reservorios naturales (73, 74). Notablemente, los nuevos métodos de secuenciación han permitido encontrar nuevos mammarenavirus en áreas completamente inesperadas, como los virus Wenzhou (WENZ) (75) y Alxa (76) en China y el virus de Mecsek (MEMV) en Europa (62).

El grupo de los virus OW incluye al prototipo de la familia, el virus de la coriomeningitis linfocítica (LCMV), que posee distribución global, y los virus Lassa (LASV) y Lujo (LUJV) que pueden producir enfermedades hemorrágicas en humanos y son endémicos en África occidental y del sur, respectivamente (78-80). También incluye especies que no se han asociado a enfermedad en el ser humano como los virus de Ippy (IPPV), Mobala (MOBV), Mopeia (MOPV), Morogoro y Kodoko (81-83). Mientras que el LCMV se encuentra en el ratón doméstico (*Mus musculus*), lo que explica que sea el único con distribución mundial, las especies africanas se encuentran principalmente en roedores de los géneros *Mastomys* y *Praomys* (todos de la familia *Muridae*, subfamilia *Murinae*) (84, 85).

Los mammarenavirus NW se han categorizado en base a la secuencia del gen de la proteína N, en los clados A, B y C (86). Se ha propuesto un cuarto clado D de especies que han surgido de un evento de recombinación ancestral, cuyas secuencias para la proteína N presentan una relación más cercana a las del clado A, mientras que las secuencias del gen GPC (glicoproteínas) están más relacionadas al clado B. De esta manera, es conveniente expandir la filogenética de los mammarenavirus al segmento S del genoma viral. Este cuarto clado incluye los virus de Whitewater Arroyo (WWAV), Tamiami (TAMV) y Bear Canyon (BCNV) (87). El clado A incluye a los virus Flexal (FLEV), Paraná (PARV), Pirital (PIRV), Alpahuayo (ALLV) y Pichindé (PICV). El clado B contiene a los virus Chaparé (CHPV) y Machupo (MACV) en Bolivia, Sabiá (SABV) en Brasil, Guanarito (GTOV) en Venezuela, y Junín (JUNV) en Argentina, todos asociados a fiebre hemorrágica en humanos. Además, incluye a los virus Tacaribe (TCRV), Amaparí (AMAV) y Cupixi (CPXV) (Figura 4), no asociados a enfermedad humana (86, 88, 89).

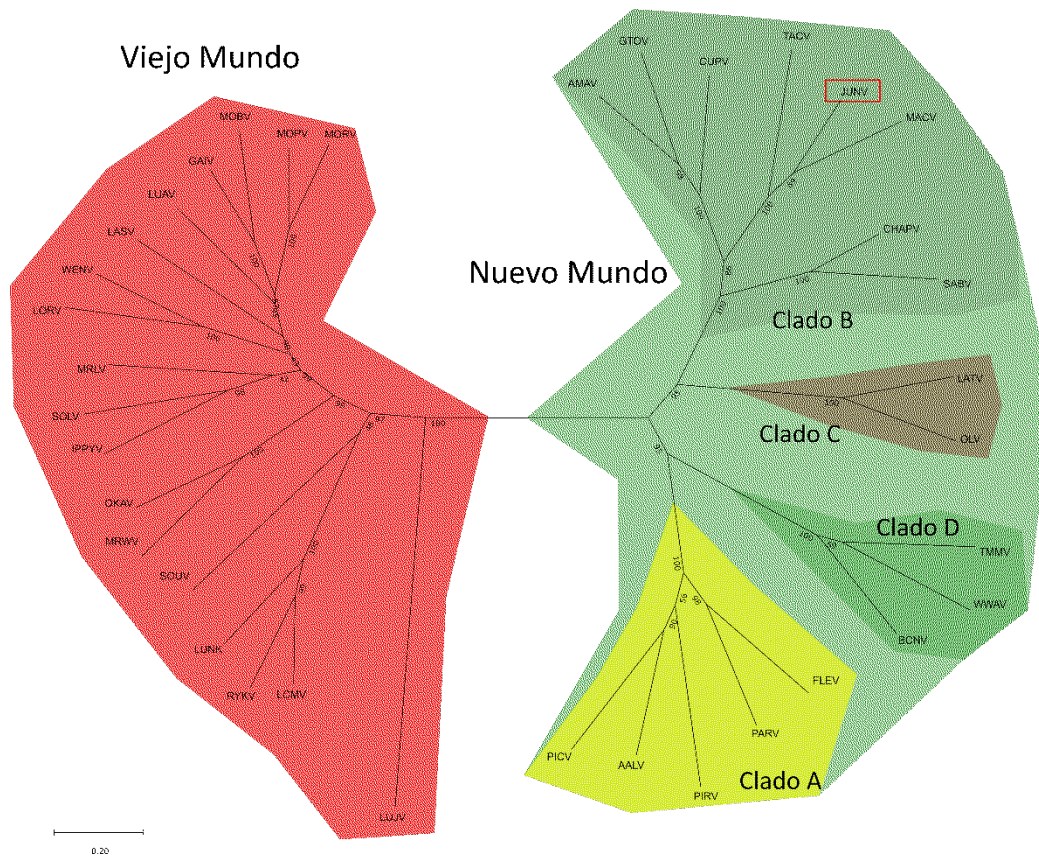


Figura 4. Árbol filogenético del género *Mammarenavirus* basados en el segmento S. Las secuencias fueron alineadas con el algoritmo MAFFT en Ugene y recortada utilizando gblocks. La historia evolutiva se infirió utilizando el método de Máxima Verosimilitud (ML) y el modelo de 3 parámetros de Tamura. El árbol de consenso *bootstrap* fue inferido a partir de 1000 repeticiones. Se utilizó una distribución Gamma discreta para modelar las diferencias de tasas evolutivas entre sitios [5 categorías (+G, parámetro = 0,7352)]. Para esto último se utilizó el programa Mega-X.

B.2 Estructura y composición

Los viriones de JUNV son envueltos, pleomórficos, con aproximadamente 100-120 nm de diámetro y su cápside es de simetría helicoidal. A nivel ultraestructural, pueden observarse un número variable de ribosomas en su interior que originaron el nombre de la familia al otorgar una apariencia arenosa [revisado en (1), entre otros]. La envoltura es adquirida durante la liberación del virus a través de la membrana plasmática de la célula huésped y presenta proyecciones en forma de espículas que contienen al SSP y las glicoproteínas G1 y G2. Dentro de la matriz y sobre el lado interno de la membrana se encuentra la proteína viral de matriz Z (90)(Figura 5).

El genoma de JUNV consiste de dos segmentos de ARN monohebra de polaridad negativa, uno de los cuales es denominado S (*small*) y el otro L (*large*) de aproximadamente 3,5 y 7,3 x 10³ nucleótidos, respectivamente. Estos segmentos se encuentran unidos a proteínas virales de nucleocápside (N) y a la ARN-polimerasa ARN-dependiente (L) dentro de las partículas infecciosas.

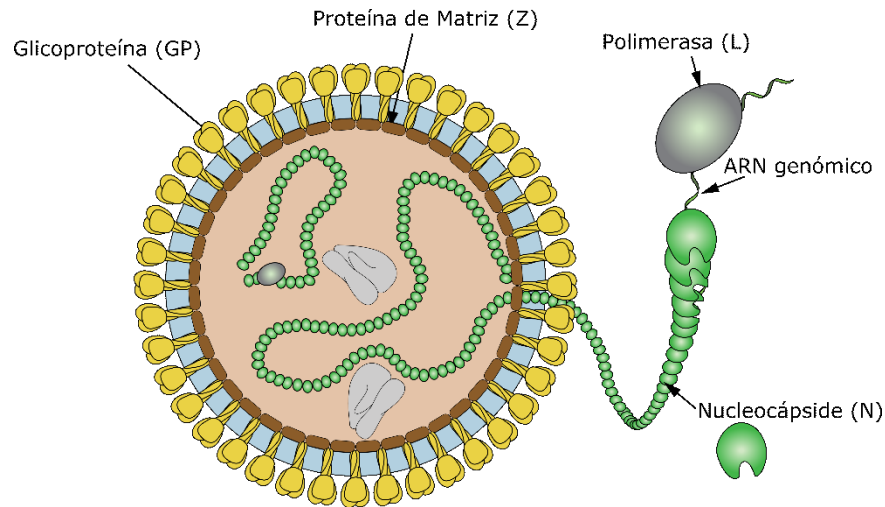


Figura 5. Estructura del virión de JUNV. Una envoltura lipídica (azul) está recubierta de glicoproteínas GP (amarillo) con proteína de matriz Z en la cara interna (marrón). Dentro de la partícula los dos segmentos de ARN genómico están recubiertos de nucleocápside N (verde) y unidos a una ARN polimerasa ARN dependiente L (gris oscuro) para formar las ribonucleocápsides. Los ribosomas se encuentran libres en el interior del virión. Adaptado de viralzone.expasy.org.

B.3 Organización genómica

El genoma de los mammarenavirus presenta una organización ambisentido. El segmento S codifica en sentido positivo para el precursor GPC de las glicoproteínas de superficie G1 y G2, que junto con el péptido señal SSP forman los trímeros que conforman la espícula, y en sentido negativo, la proteína de nucleocápside N. El segmento L codifica para la polimerasa L, y en sentido opuesto, para la proteína ligadora de zinc denominada Z. En ambos segmentos se encuentra una región intergénica que formaría una estructura secundaria estable (91).

Las regiones no traducidas (92) en los extremos 5' y 3' de cada segmento de ARN se encuentran conservadas entre los mammarenavirus. Al igual que otros virus de ARN de cadena negativa, estas regiones no traducidas (RNT) son complementarias entre sí y, por lo tanto, pueden formar estructuras circulares (93). Las RNT también juegan un papel importante en la formación de nucleocápsides virales y favorecen el reconocimiento de la polimerasa viral L de los extremos terminales virales para iniciar la replicación y la transcripción (93)(Figura 6).

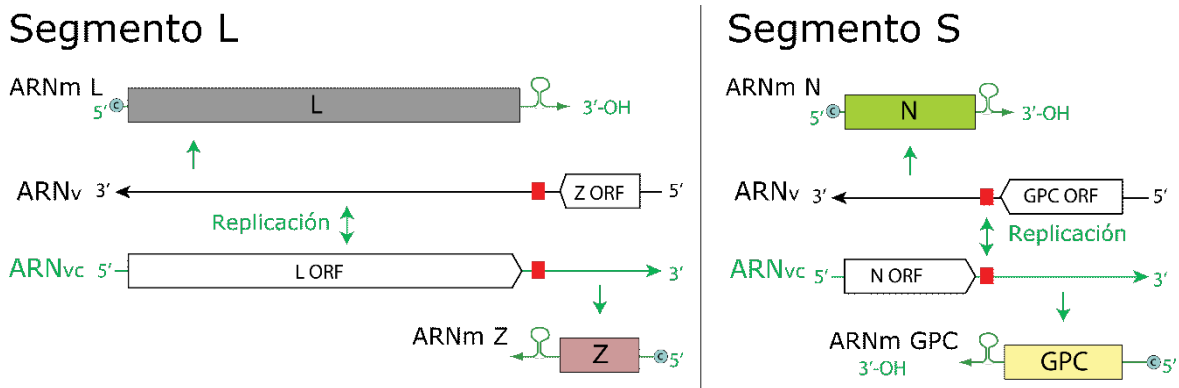


Figura 6. Diagrama representativo de la organización genómica, transcripción y replicación de JUNV. A la izquierda, el segmento L monocatenario de orientación negativa (ARNv) se puede transcribir en el ARN mensajero para la polimerasa L (ARNm L) y debe replicar en la cadena complementaria (ARNvc) para transcribir el ARN mensajero de la proteína Z (ARNm Z). En el recuadro derecho, el segmento S genómico (ARNv) puede transcribir el ARN mensajero para la proteína N (ARNm N) y debe replicarse la cadena complementaria (ARNvc) para poder transcribir el mensajero complementario de la glicoproteína GPC (ARNm GPC). En ambos segmentos se puede observar los marcos de lectura en orientación opuesta y separados por una región intergénica (rojo). Adaptado de viralzone.expasy.org.

B.4 Ciclo de replicación

B.4.1 Ingreso celular

Los mammarenavirus OW y los NW clado C, utilizan alfa-distroglucano como receptor para el ingreso a las células (94, 95). Esta proteína funciona como una molécula transmembrana de anclaje que conecta el citoesqueleto de actina con la matriz extracelular. Adicionalmente, se han reportado 4 receptores adicionales o correceptores para LCMV y LASV: DC-SIGN, LSECtin y Tyro3 y Axl de los receptores TAM de tirosina quinasa (96, 97). El mecanismo de entrada de los arenavirus OW es independiente de la endocitosis mediada por clatrina, contrario a los arenavirus NW (98). Algunos estudios han demostrado que LASV y LCMV entran en la célula a través de los cuerpos multivesiculares/endosomales tardíos siguiendo la ruta de degradación del distroglicano (99).

Mientras aún no se ha identificado el receptor celular de los arenavirus NW del clado A, los virus del clado B utilizan el receptor de transferrina 1 (55). Esto ha sido demostrado tanto en primates no humanos como en cobayos (100, 101). Mientras JUNV logra usar el receptor de transferrina 1 de cobayos (gpTfR1), el virus falla en utilizar la proteína ortóloga en ratones y hámsteres (54). Se ha observado que una mutación dentro de la proteína hTfR1 permitiría convertirla en un receptor funcional para TCRV (102).

Se han propuesto otros receptores de superficie celular como vías alternativas de entrada para JUNV, incluyendo los receptores de fosfatidilserina (PS), lectinas tipo C y canales de calcio dependientes de voltaje (103-105). Adicionalmente, se han demostrado que la infección de JUNV aumenta con la presencia de la molécula de adhesión intercelular específica de células dendríticas humanas-3 unidoras a no-integrina (hDC-SIGN) o su homólogo DC-SIGN o L-SIGN, independientemente de la presencia de hTfR1 (105).

Cómo se mencionó previamente, la internalización de los arenavirus NW, incluyendo el JUNV, utiliza la endocitosis mediada por clatrina (98). Estudios posteriores demostraron que la entrada es

dependiente de la proteína dinamina II y del sustrato del receptor del factor de crecimiento epidérmico 15 (EPS15), dos factores altamente asociados con la vía clásica de la endocitosis mediada por clatrina (106, 107). La infección temprana también requiere que se encuentren intactos el citoesqueleto de actina y la flexibilidad de la red dinámica de microtúbulos (108).

La fusión de las membranas es mediada por pH, dándose durante la acidificación de endosomas maduros (109). Se ha estimado que el pH óptimo de fusión para la envoltura de JUNV es aproximadamente de 5 (109). La acidificación desprende G1 del complejo receptor e induce cambios conformacionales en G2, permitiendo a SSP y G2 la fusión y liberación del contenido genómico viral al citoplasma (110). El JUNV utiliza los endosomas tempranos mediados por Rab5 y los endosomas tardíos mediados por Rab7 (106).

B.4.2 Replicación

Durante la replicación genómica, es sintetizada una copia de largo completo del ARN genómico S y L. Los ARN genómicos no son capaces de dirigir la síntesis de polipéptidos virales y por tanto no son infecciosos. Debido a la codificación ambisentido, tanto el ARN genómico como el antígenómico sirven de moldes para la transcripción de ARNm viral. Estos ARNm tienen una caperuza (CAP) 5' pero no son poliadenilados (111). Las proteínas virales se traducen utilizando un mecanismo dependiente del límite, cuya eficiencia se basa en la interacción entre las señales estimuladoras en el 5' RNT y un elemento modulador negativo en el 3' RNT (112).

Las primeras proteínas virales sintetizadas son N y L, las cuales representan los factores virales mínimos requeridos para la síntesis eficiente del ARN viral (113, 114). GPC y Z son transcritos desde el ARN antígenómico de los segmentos S y L. El ARN antígenómico también sirve de molde para la generación de segmentos de ARN genómicos, los cuales, junto con N y L, se asocian en complejos de ribonucleoproteína (RNPs). Finalmente, las RNPs son empaquetadas y los viriones progenie son liberados de la célula. La envoltura lipídica deriva de la membrana plasmática de la célula huésped que envuelve a las partículas de virus brotante y presenta las espículas conformadas por G1, G2 y SSP (115, 116).

Se ha demostrado que la región intergénica interviene en la replicación viral debido a las estructuras de horquilla altamente estables que comprenden tallos abultados en estas regiones (tanto en el segmento L como S) para soportar la síntesis de ARN subgenómico (117). Ni las secuencias corriente abajo ni corriente arriba de la horquilla desempeñan un papel en la acumulación de ARN (117). En un estudio con LCMV, se observó que la región intergénica cumple funciones individuales en la terminación de la transcripción para mejorar la expresión génica y en el ensamblaje y/o brotación del virus, que son necesarios para la propagación eficiente de la infectividad del LCMV (118).

Se ha reportado que una deleción de 35 nucleótidos y una mutación puntual no sinónima en la secuencia intergénica del segmento L de MACV atenúa una cepa patogénica en cobayos y ofrece protección contra GTOV (119). En LCMV, se determinó que las regiones intergénicas del segmento L y del segmento S cumplen funciones distintas. De hecho, el reemplazo de la región intergénica no codificante del segmento del genoma L por la región intergénica del segmento del genoma S para

generar un LCMV recombinante, rLCMV (RIG/SS), resultó altamente atenuado *in vivo* pero indujo protección contra un desafío letal con LCMV de tipo salvaje (120).

B.4.3 Ensamblado y brotación

En los mammarenavirus, a excepción de TACV, la brotación es dependiente de dominios sinónimos ASAP y YLCL en dominios tardíos de la proteína Z. Además, la brotación de TACV es independiente de Tsg101 (componente del complejo ESCRT-1) pero requiere de la actividad de Vps4A/B (121). Para su función en la brotación viral, tanto TACV, LCMV y LASV requieren de miristoilación en G2. Adicionalmente, para aquella función y la interacción con el complejo glicoproteico, esta miristoilación ocurre en el residuo 2 de Z (122).

Respecto al empaquetamiento de los viriones, hay evidencia que el residuo L79 de la proteína Z en TACV y JUNV es esencial para las interacciones Z-N y la incorporación de tanto N como GP en las partículas brotantes. La integridad del dominio *Really Interesting New Gene* o RING (ver más adelante), no es estrictamente requerida para la liberación eficaz de partículas contenedoras de Z. No obstante, tanto la ruptura de la estructura RING como el reemplazo de L79 con alanina reduce la infectividad de VLPs (del inglés, *virus-like particles*) sin modificar las propiedades de autobrotación de Z, como también perjudica las interacciones intracelulares Z-N y la incorporación de N dentro de VLPs (123)(Figura 7).

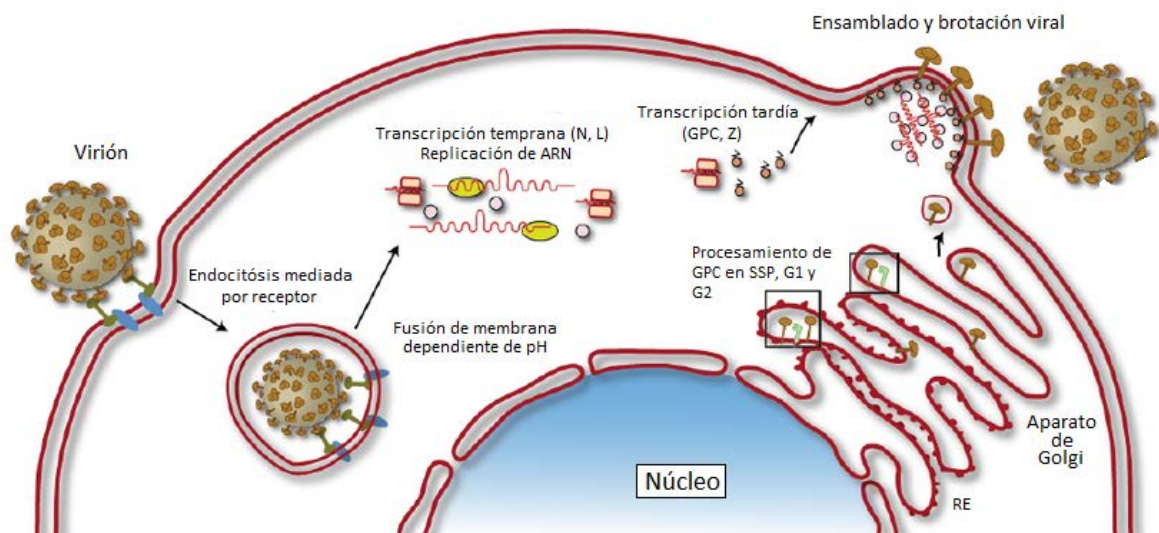


Figura 7. Ciclo de replicación de los mammarenavirus. Adaptado de (124).

B.5 Aislamientos y cepas

Desde su descubrimiento se han obtenido numerosos aislamientos de JUNV, algunos de las cuales fueron posteriormente caracterizados. Dicha caracterización se basó en el efecto en modelos animales, ya sea cobayos, ratones o primates, definiendo así tipo de infección asociado, virulencia y patogenicidad contribuyendo a la transformación de estos aislamientos en cepas. En algunos casos el aislamiento provino de pacientes de FHA. Por ejemplo, la cepa Romero corresponde a un aislamiento de JUNV de un paciente de FHA severa pero no letal y ha sido frecuentemente utilizada

como modelo de virus salvaje para comparar contra una cepa atenuada (125). La virulencia fue probada en cobayos de tres semanas de edad y la DL_{50} fue determinada a los 11-14 días posinfección (dpi), siendo menor a 10 UFP. La cepa Ledesma, en cambio, proviene de un caso fatal de FHA que produce enfermedad neurológica en primates (41). Otro caso es la cepa Espíndola proveniente de un caso fatal de FHA, que a diferencia de la cepa Ledesma, produce una enfermedad hemorrágica en primates (41).

Con el objetivo de mantener una perspectiva de la evolución, propagación y virulencia de las cepas circulantes de JUNV, se realizan aislamientos en forma constante de diferentes orígenes. Varios de éstos provienen de pacientes de FHA y terminan siendo secuenciados y cargados en base de datos, generalmente siguiendo la nomenclatura "P" y un número. Por ejemplo, nuestro grupo de investigación ha trabajado con la cepa P3441, proveniente de un aislamiento de paciente humano en un caso letal de FHA en 1977 y originalmente provista por la Dra. A. Ambrosio del Instituto Nacional de Enfermedades Virales Humanas "Dr. Julio I. Maiztegui" (51-53). En otros casos, se han caracterizado virus aislados de su reservorio natural *Calomys musculus*, tal como la cepa MC2, obtenida del roedor en 1967 y utilizada como modelo patogénico (14), o de virulencia intermedia (126).

Varias cepas vivas atenuadas de JUNV se derivaron de la cepa prototipo XJ de JUNV aislada de humanos por Parodi y colaboradores (3) mediante pasajes en células cultivadas, cerebros de ratón y/o cobayos. La cepa XJCl3 se derivó de la cepa XJ por pasajes en células MA-111 (riñón de conejo). La cepa XJCl3 se atenuó significativamente en cobayos y ratones en comparación con la cepa XJ parental y provocó títulos altos de anticuerpos neutralizantes en animales inoculados (127). Se administró una formulación de vacuna basada en la cepa XJCl3 a 636 voluntarios humanos, la mayoría de los cuales desarrollaron una infección subclínica o síntomas leves. La inmunización indujo títulos elevados de anticuerpos neutralizantes que fueron detectables hasta nueve años después de la inmunización en el 90% (153 de 165) de los receptores de la vacuna analizados (128). Sin embargo, dado que la reserva de vacuna de virus de siembra de la cepa XJCl3 se estableció en cerebro de ratón lactante y se clonó en una línea celular heteroploide, se interrumpieron los ensayos clínicos.

Otra cepa atenuada de JUNV, la XJ0, se derivó de un virus parental común con XJCl3. La inmunización de cobayos con la cepa XJ0 indujo inmunidad protectora frente a la exposición a JUNV virulento a partir de los 3 dpi, y a los 30 días protegió al 100% de los animales (129). Aunque estos resultados demostraron el potencial de la cepa XJ0 para el desarrollo de una vacuna viva atenuada, los estudios se interrumpieron cuando se detectó una infección persistente por JUNV en los órganos linfohemopoyéticos de cobayos inoculados [revisado en (130)].

La cepa de JUNV C#1 fue desarrollada como un esfuerzo de colaboración entre el Instituto de Investigación Médica de Enfermedades Infecciosas del Ejército de los EE. UU. (USAMRIID) y el Ministerio de Salud y Acción Social de Argentina. C#1 tiene un historial de pasajes definido (Figura 8), que se originó a partir de la cepa prototipo XJ de JUNV. Se buscó obtener cepas que adquirieran atenuación con pasajes consiguientes, a los cuales para su seguimiento fueron llamados XJ junto a un número (por ejemplo, XJ13, por tener 13 pasajes). Después de dos pasajes en cobayos y cuarenta y tres pasajes en ratones por medio de inyección intracerebral, se obtuvo la cepa XJ44, la cual fue

caracterizada como atenuada en cobayos y virulenta en ratón con un índice de letalidad de 1,37 +/- 0,45. El homogeneizado de cerebro del último pase de cerebro de ratón se usó para infectar células FRhL-2 (línea celular diploide de pulmón de *rhesus* fetal certificada), en las que se pasó diecinueve veces el virus seguido de clonado a través de dos pasos de dilución limitante. Las semillas maestra y secundaria se obtuvieron después de una única ronda de amplificación en células FRhL-2, finalmente obteniendo lo que se nombraría C#1. La alta eficacia protectora y la falta de neurovirulencia de C#1 se demostraron en cobayos y NHP (131, 132). Las pruebas de campo dirigidas a la evaluación de la inmunogenicidad y eficacia de C#1 en humanos involucraron a 6.500 trabajadores agrícolas masculinos en una región de Argentina endémica de FHA y demostraron una alta eficacia protectora ($\geq 84\%$). Además, no se detectaron efectos adversos graves asociados con la vacunación (133). Actualmente, C#1 es la única vacuna aprobada en Argentina para la prevención de la FHA y al presente se ha utilizado para vacunar a más de 200.000 personas en riesgo.

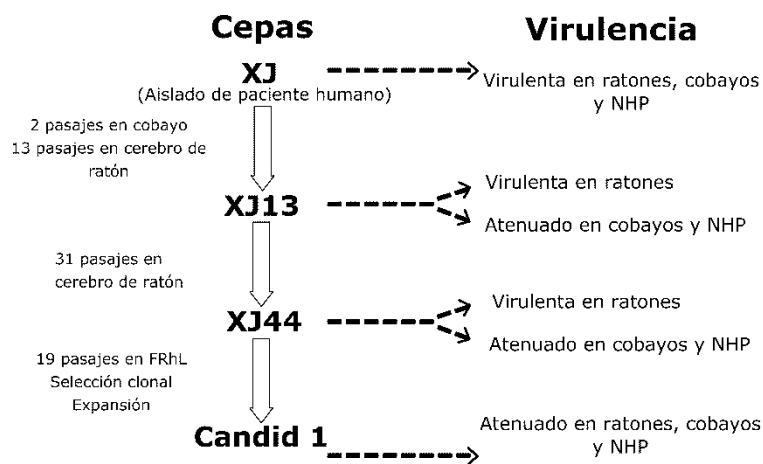


Figura 8. Desarrollo de la cepa vacunal de JUNV Candid 1. NHP = primate no humano, FRhL = células fetales de pulmón del mono rhesus. Adaptado de (39).

Todas las cepas incluidas en este estudio han sido secuenciadas en su totalidad y toda información pertinente se encuentra en bases de datos de acceso libre, tal como la *National Center of Biotechnology Information* (NCBI). Su grado de relación filogenético se muestra en la Figura 9. No obstante, no todas han sido utilizadas con la misma frecuencia para el estudio de JUNV. Generalmente, debido a cuestiones de bioseguridad, la cepa vacunal C#1 es la más utilizada como referencia de JUNV, lo cual de por sí ofrece una perspectiva parcial del virus en situación salvaje (cabe recordar que C#1 es el resultado de una extensa manipulación dentro de condiciones de laboratorio). Cuando se busca comparar contra una cepa patogénica semejante al JUNV silvestre, se han utilizado las cepas Romero y XJ13. Desafortunadamente, el transporte y manipulación de cepas patogénicas restringen su distribución entre laboratorios, y no se ha llegado a definir una única cepa patogénica modelo. De hecho, la falta de coordinación entre laboratorios ha resultado en repetidas cargas de secuencias en base de datos para una misma cepa.

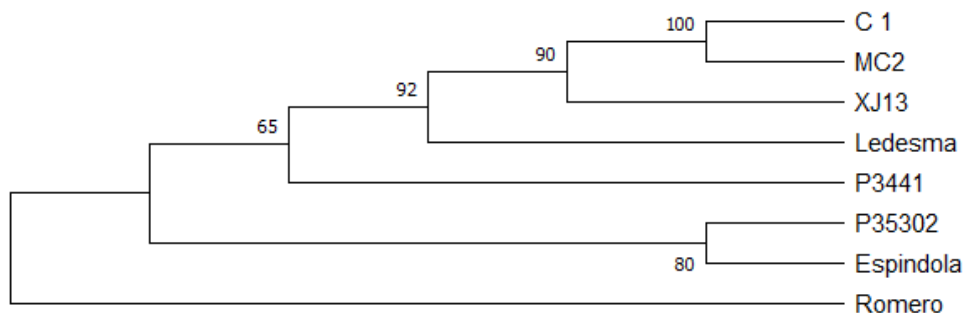


Figura 9. Árbol filogenético para diferentes cepas de JUNV realizado a partir del alineamiento con ClustalW de secuencias del segmento S de cada cepa por medio del método de Máxima Verosimilitud y el modelo evolutivo de 3 parámetros de Tamura con distribución gamma (ML + T92 + G; parámetro G = +0,1109) y 5 categorías. Los números en los nodos representan los valores del test estadístico de *bootstrap*. Un mayor número de *bootstrap* indica mayor ajuste de los datos y consistencia para la bipartición del taxón.

Tabla 2. Cepas de JUNV que se encuentran secuenciadas en forma total disponibles en NCBI incluidas en este trabajo.

Cepa	Origen	Patogenicidad	Número de Acceso		Año (cargado a base de datos)
			Segmento S	Segmento L	
C#1	Laboratorio	Atenuada	AY746353 FJ969442	AY746354 AY819707	2004 2009
XJ13	Laboratorio	Patogénica (en cobayos)	AY358023 NC_005081 FJ805378	AY358022 NC_005080 FJ805377	2003 2003 2009
P35032	Paciente de FHA no letal	Patogénica	KU978800	KU978799	2016
MC2	<i>Calomys musculus</i>	Intermedia	D10072	AY216507	1989/2003
Romero	Paciente humano de FHA no letal "común"	Patogénica	AY619641 JN801476	AY619640 JN801477	2004 2012
Ledesma (P3766)	Paciente humano de FHA letal neurológica	Patogénica letal	KR260734	KR260733	2015
Espindola (P3790)	Paciente humano de FHA letal hemorrágica	Patogénica letal	OL774853	OL774852	2021

*Común: incluye fiebres y otros síntomas que se desarrollan en los pacientes de FHA y que en la mayoría de los pacientes se resuelven sin desarrollo de síntomas hemorrágicos o neurológicos, definido previamente (134).

En nuestro laboratorio hemos utilizado las cepas P3441 y C#1 como modelos patogénico y atenuado respectivamente en varios estudios. No obstante, hasta la fecha, nunca se había buscado profundizar en las diferencias moleculares entre ambas cepas. De hecho, la cepa P3441 nunca fue secuenciada en su totalidad. Actualmente se cuenta con la información en base de datos y herramientas disponibles para poder realizar comparaciones entre P3441 respecto a otras cepas patogénicas de JUNV a fin de dilucidar sus determinantes moleculares de patogenicidad.

Hipótesis y objetivos

Hipótesis de trabajo

Las diferencias de patogenicidad entre las cepas de JUNV radican en diferencias puntuales en la secuencia genómica y aminoacídica que resultan en características biológicas distintivas de las proteínas.

Objetivo General

Conocer, identificar y caracterizar los determinantes moleculares de patogenicidad de JUNV P3441.

Objetivos Particulares

- Obtener y analizar la secuencia de la cepa patogénica JUNV P3441.
- Comparar las diferencias existentes entre la cepa P3441 y otras cepas de JUNV.
- Clonar las proteínas virales de JUNV P3441 en vectores de clonado.
- Subclonar las proteínas virales de JUNV P3441 en vectores de expresión eucariota.
- Comprobar la expresión y la antigenicidad de las proteínas recombinantes en células eucariotas.
- Estudiar aspectos moleculares desconocidos para JUNV que permitan caracterizar la cepa P3441.

Capítulo 1. Secuenciación y clonados

1.1 Preparación y diseños

Para obtener la secuencia de la cepa P3441, se planificó generar moléculas de ADN por PCR de todo el genoma viral. Con este fin, se diseñaron *primers* con las siguientes consideraciones:

- Los productos de PCR tendrían un largo menor a 950pb, el límite superior del método seleccionado de secuenciación, como también de esta manera reducir la probabilidad de errores y favorecer su manipulación.
- El largo de cada segmento sería mayor a 600pb, el límite inferior del método para asegurar una secuenciación completa con cada uno de los *primers* utilizados para generar cada segmento.
- La temperatura de fusión sería 55°C aproximadamente para todos los pares de *primers* a fin de facilitar combinaciones posteriores.
- Cada fragmento tendría una región de solapamiento con el fragmento anterior y posterior, a excepción de los extremos de cada segmento genómico del virus, para asegurar la verificación de potenciales errores asociados a los extremos del método de secuenciación y obtener una continuidad de la secuencia.

Los *primers* se diseñaron en base a la secuencia de C#1 en la base de datos nucleotídica de NCBI con números de acceso FJ969442 y AY819707 para el segmento S y L, respectivamente. Cuando hubo opciones disponibles, se eligieron aquellos que:

- Presentaban un porcentaje de citosina-guanina entre el 40 y 60% en la secuencia.
- Tenían un tamaño entre 18 y 25 nucleótidos.
- No tenían 4 o más de un mismo nucleótido o dinucleótido seguidos.
- No formaban estructuras secundarias complejas.

Finalmente, se encargaron 17 pares de *primers*, 11 para el segmento L y 6 para el segmento S. Para facilitar la nomenclatura, se designaron por número, siendo 1 el *primer* directo del extremo 5' del segmento L del genoma viral y 34 el *primer* reverso del extremo 3' terminal del segmento S.

1.2 Resultados

1.2.1. Amplificación y secuenciación

Una vez obtenido el ARN de P3441, se dividió en alícuotas y, después de medir la concentración y calidad del producto, se realizó una reacción de transcripción reversa. Inicialmente se realizaron dos reacciones en paralelo utilizando *primers* aleatorios y una mezcla de los *primers* reversos de los 17 segmentos.

Con los ADNc resultantes, se realizaron ensayos de PCR con todos los pares de *primers* utilizando la polimerasa Taq. Una vez verificados los productos amplificados en electroforesis en gel de agarosa, se repitió la PCR con polimerasa Pfu. Las reacciones que presentaron la banda esperada fueron purificadas. En algunos casos, fue necesario realizar varios ensayos de PCR para obtener suficiente masa.

Los productos obtenidos de las RT-PCR se enviaron a la empresa MacroGen Inc. (República de Corea), excepto la de los segmentos 7-8, 21-22 y 27-28 que no se pudieron amplificar. Estas secuencias se

obtuvieron posteriormente durante el clonado de las proteínas y su secuenciación. Aunque se logró secuenciar la mayoría de la región intergénica del segmento L, cuando se comparó con otras cepas, se observó la ausencia de 24 nucleótidos (Figura 1.1A). Curiosamente algo similar se observó en la secuencia de la cepa Ledesma. Como se puede ver en la Figura 1.1B y la Tabla 3, la región intergénica del segmento S de las otras cepas estudiadas, mantienen un alto porcentaje de identidad.

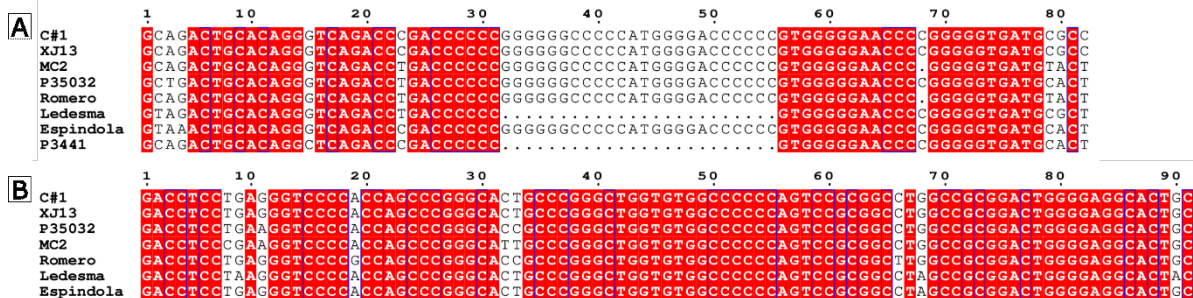


Figura 1.1 Alineamiento de secuencias intergénicas de los segmentos L (A) y S (B) para diferentes cepas de JUNV. Se muestra lo obtenido para la secuencia de P3441 para el segmento L.

Tabla 3. Comparación de regiones intergénicas de los segmentos S y L para distintas cepas de JUNV respecto a C#1.

	Segmento S		Segmento L	
	Similitud (%)	Diferencias (nº nts)	Similitud (%)	Diferencias (nº nts)
XJ13	100	0	0	0
P35032	98	2	95	4
MC2	97	3	95	4
Romero	97	3	95	4
Ledesma	97	3	96*	3*
Espindola	99	1	95	4
P3441	-	-	93*	4*

*excluyendo la región ausente

Las secuencias de la cepa P3441 obtenidas se comunicaron al NCBI y quedaron registradas con las siguientes nomenclaturas: polimerasa L (OR806657), proteína de matriz Z (OR806658), la proteína de nucleocápside N (OR806659) y el precursor de glicoproteínas GPC (OR806660). Los alineamientos comparativos se muestran a nivel aminoacídico en los capítulos siguientes.

1.2.2 Clonado de las proteínas virales

1.2.2.1 Proteína Z

El marco de lectura para la proteína Z de P3441 fue amplificado con *primers* que incluían las secuencias para las enzimas de restricción BamHI y HindIII. Los fragmentos amplificados se clonaron en el vector de entrada *pGEM-T Easy Vector* (Promega) que permite la ligación aprovechando la poliadenilación de las moléculas de ADN producidas con la polimerasa Taq. Posteriormente a su verificación con las enzimas de restricción y secuenciación, se procedió a la realización de *minipreps*. Z de C#1 ya había sido clonada previamente en el laboratorio por el Dr. Agustín Ure. Posteriormente, se procedió con la subclonado (ver capítulo 2).

1.2.2.2 Proteína N

El marco de lectura de la proteína N de P3441 fue amplificado como se describió previamente, utilizando *primers* que incluían el sitio de restricción para las enzimas NheI y KpnI, excluyendo el codón *stop* 3' terminal para que pueda traducirse la proteína recombinante con el fluoróforo del vector diana. Debido a que pEGFP-N3 y pDsRed-N1 difieren en dos pares de bases entre el sitio de clonado múltiple (el cual es idéntico para ambos plásmidos) y el fluoróforo del vector, se amplificó el marco de lectura de N de P3441 dos veces utilizando un *primer* reverso distinto, obteniéndose un fragmento para su posterior subclonado en pEGFP-N3 y otro para pDsRed-N1. N de C#1 se amplificó directamente del plásmido pGEM-N de C#1 mencionado previamente. Se clonaron los fragmentos obtenidos en el vector de entrada *pGEM-T Easy Vector* y se verificaron por digestión con enzima de restricción y secuenciación, obteniéndose los plásmidos designados pGEM-N de P3441 N3, pGEM-N de P3441 N1 y pGEM-N de C#1 N1. Posteriormente, se procedió con la subclonado (ver capítulo 3).

1.2.2.3 Precursor GPC

Para GPC los *primers* no incluían sitios de restricción propios y se incluyó el codón *stop*. Se clonaron los fragmentos obtenidos en el vector de entrada *pGEM-T Easy Vector* y se verificaron por digestión con enzima de restricción y secuenciación, obteniéndose el plásmido designado pGEM-GPC de P3441. Posteriormente, se procedió con la subclonado (ver capítulo 4).

1.2.2.4 Proteína L

Debido al tamaño de la proteína L, se diseñó amplificar el marco de lectura de L en 4 fragmentos que incluyesen en sus extremos sitios de restricción internos de la proteína (NheI, NdeI y NcoI), utilizando *primers* con sitios de restricción para los extremos de los fragmentos terminales para poder subclonar en el vector de destino (Tabla 4). El diseño de los *primers* fue en base a la secuencia de P3441 que se había obtenido previamente.

Se intentó amplificar por PCR cada fragmento con polimerasa Taq y los *primers* diseñados para después clonar directamente en el vector *pGEM-T East Vector*. Aunque se pudieron amplificar y clonar en el vector de destino los fragmentos L1, L2 y L3, no se logró para el fragmento L4. Por lo tanto, se reformuló la estrategia para clonar L en 3 fragmentos utilizando nuevos *primers* que incluían sitios de restricción preexistentes en la secuencia de P3441. Se utilizó el mismo *primer* directo 5' de L1 y el primer *inverso* 3' de L4 del plan anterior (Tabla 5).

Tabla 4. Fragmentos diseñados para el primer intento de clonado de la proteína L de P3441.

Nombre	Tamaño (pb)	Enzima de restricción 5'	Enzima de restricción 3'
L1	598	BamHI	NheI
L2	2475	NheI	NdeI
L3	987	NdeI	NcoI
L4	2634	NcoI	XhoI

Tabla 5. Fragmentos diseñados para el segundo intento de clonado de la proteína L de P3441.

Nombre	Tamaño (pb)	Enzima de restricción 5'	Enzima de restricción 3'
L1 NEW	2166	BamHI	KpnI
L2 NEW	2397	KpnI	BstEII
L3 NEW	2152	BstEII	XhoI

Se lograron amplificar los tres fragmentos por PCR y su clonado en *pGEM-T Easy Vector*. Posterior a su verificación con enzima de restricción y secuenciación, se procedió con la subclonado (ver capítulo 5).

1.2.2.5 Vector diana

Una complicación que surgió durante el clonado de los genes codificantes de las proteínas virales fue la doble digestión de los vectores pEGFP-N3 y pDsRed-N1 donde se subclonarían las proteínas. Debido a diversos factores, resultó difícil asegurar la doble digestión de ambos vectores ya sea porque la calidad de las enzimas no fuera óptima o porque se encontraban muy próximos como para poder discernir si la digestión había sido exitosa. Por lo tanto, se buscó clonar un fragmento postizo en estos vectores de expresión eucariota en el centro del sitio de clonado múltiple para luego purificar los plásmidos que hubieran sido exitosamente digeridos por las dos enzimas.

Se digirió el plásmido pGEM-N de C#1, provisto por la Dra. López (Cevan) con la enzima de restricción EcoRI, liberando el marco de lectura completa de la proteína N. El fragmento liberado fue posteriormente purificado. Al mismo tiempo, se digirieron pEGFP-N3 y pDsRed-N1 con la misma enzima. Los plásmidos linealizados fueron purificados y posteriormente tratados con fosfatasa alcalina intestinal de ternero para impedir la recircularización durante la ligación. Luego se realizó una reacción de ligación empleando la enzima T4 entre los plásmidos (por separado) con el fragmento del plásmido pGEM-N de C#1 (ver parte 1.2.2. de este capítulo).

Los productos de la reacción de ligación fueron transformados. Las colonias se verificaron por *colony-PCR* y aquellas positivas fueron repicadas para realizar *minipreps*. Finalmente, se hizo una verificación con un ensayo de digestión con EcoRI para confirmar la liberación del inserto clonado y su utilización en las clonados posteriores, designados como pDsRed-X y pEGFP-X.

1.3 Discusión

En este trabajo se siguió un método de secuenciación que aseguraba que sólo fragmentos de tamaño reducido del genoma viral salieran del laboratorio. Con este fin, se diseñaron varios pares de *primers* y se hicieron diversas RT-PCR. Los productos amplificados fueron purificados y secuenciados en su mayoría. Las excepciones incluyeron el par 27-28, que debía amplificar la región intergénica del segmento S. Esta región forma una estructura secundaria estable y por ende no sorprende que hayan surgido complicaciones en su amplificación. Afortunadamente, no hubo complicaciones para amplificar la región intergénica del segmento L. De los otros dos pares de *primers* que no lograron amplificar los fragmentos esperados, es posible que el par 21-22 tuviese un problema similar ante la estructura secundaria estable del extremo 3' del segmento L. Considerando que posteriormente hubo dificultad en amplificar el extremo C terminal del marco de lectura para la proteína L, es posible que esa región sea más difícil de amplificar. Afortunadamente, se pudo completar la secuencia de P3441 durante los varios intentos de clonado, cuando se enviaron fragmentos clonados para verificar su secuencia.

La mayoría de los *primers* utilizados para la secuenciación demostraron cumplir su función y quedan disponibles para estudios de regiones particulares del genoma de JUNV. Intentos fallidos y complicaciones consecutivas incentivaron la contratación de servicios de clonado comerciales para la compleción de este trabajo. Aun así, dichas empresas también compartieron sus propios problemas y limitaciones.

La realización de trabajos similares en el futuro deberá tener en cuenta las complicaciones observadas a lo largo de este trabajo. La utilización de métodos alternativos de clonado, tales como el método de Gibson o Golden Gate podrían ser una alternativa válida. En el ensamblado de Gibson se realiza en una única reacción isotérmica que combina una exonucleasa que degrada extremos con secuencia solapante, la actividad de una polimerasa de ADN que contrarresta la exonucleasa y una ligasa que detiene el proceso al unir segmentos que consiguieron solaparse. Con este método se reduciría el número de pasos y reactivos a manipularse. Por otro lado, el *Golden Gate* hace uso de vectores de entrada y de destino con sitios de restricción para enzimas tipo IIS, que digieren corriente abajo del sitio de reconocimiento y generan extremos no palindrómicos evitando el riesgo de recircularización en el clonado, aunque requiere mayor nivel de planificación que los otros métodos de clonado.

Capítulo 2. Proteína Z

2.1 Introducción

La proteína de matriz Z de los mammarenavirus tiene un largo que oscila entre 90 a 99 aminoácidos, siendo en JUNV de 94. Su peso molecular aproximado es 11 kDa. Esta proteína está involucrada en varias funciones del ciclo de replicación del virus, incluyendo la regulación de la síntesis del ARN viral, el proceso de brotación del virión y el ensamblado viral. También se la ha asociado como modulador de los niveles de interferón de tipo I (IFN-I) celular (135).

2.1.1 Estructura de la proteína Z

La proteína Z presenta tres dominios característicos: un dominio central RING y los dominios N y C terminales que flanquean el dominio central. Estos dos últimos presentan una estructura altamente flexible que permite que adopten diferentes estados conformacionales (Figura 2.1) (136).



Figura 2.1. Gráfico esquemático de los dominios de la proteína Z. El brazo N terminal (rojo) presenta un sitio de miristoilación en el segundo residuo de la cadena. El dominio central RING (verde) es una estructura secundaria estable y coordina cationes zinc. El brazo C terminal (azul) tiene dominios tardíos L en cantidades, secuencias y distribución variable entre las especies de los mammarenavirus.

Los dominios RING presentan motivos distintivos que pueden encontrarse en diferentes proteínas celulares y virales. Generalmente, cuentan con un largo de 40 a 60 aminoácidos y consisten típicamente de una única hélice alfa y múltiples láminas beta. Su característica principal es el motivo Cisteína₃HistidinaCisteína₄ con el cual coordina para dos cationes zinc. Dentro de este motivo se reconocen dos sitios, denominados I y II, involucrados directamente en la coordinación catiónica. Estos sitios se encuentran en lados opuestos del dominio. Aunque el sitio I es muy similar a la estructura de otros dominios RING conocidos, la conformación del sitio II es una característica propia de la proteína Z de los arenavirus (136). Se ha demostrado que el dominio RING, particularmente el sitio I, es importante en la función de autoensamblado de Z (137-139).

En los dominios tardíos de la proteína Z se encuentra un motivo de cuatro péptidos identificados en las proteínas de matriz de otros virus ARN y en la proteína Gag de los retrovirus. Este tetrapéptido puede consistir en las secuencias aminoacídicas P[T/S]AP y/o PPxY, donde 'x' representa cualquier aminoácido (140). Adicionalmente, la proteína Z también cuenta con el motivo YxxL dentro del dominio RING que pertenece a los dominios tardíos en otros virus. Estos tres motivos se encuentran involucrados en las interacciones proteína-proteína entre la proteína viral Z y componentes del complejo de clasificación endosomal para el transporte (ESCRT), incluyendo TSG101. Por su parte, el ESCRT es importante en la vía de clasificación de proteínas vacuolar (VPS), en el que las proteínas asociadas a la membrana se clasifican a través de una serie de compartimentos endosómicos para su eventual degradación en el lisosoma (140, 141).

Todos los mammarenavirus, ya sean OW o NW, tienen el motivo altamente conservado YxxL dentro del dominio central RING, y los motivos P[T/S]AP y PPPY en el dominio C terminal. No obstante, estos últimos motivos varían entre los mammarenavirus NW y OW, tanto en el número como en la posición relativa. Mientras que LCMV presenta un motivo PPPY y STAP, especies africanas cuentan con motivos PPPY y PTAP cercanos, con la excepción del LUJV. En la mayoría de los mammarenavirus NW, la proteína Z cuenta con un motivo P[T/S]AP en el extremo C terminal.

En el extremo N terminal de la proteína Z de todos los mammarenavirus se encuentra un sitio de miristoilación conservado, el cual es importante para la función de la proteína en el ensamblado de los viriones (142, 143). En este sitio se une un ácido mirístico de manera irreversible, covalente y cotraduccional. La miristoilación N terminal requiere de la enzima de la célula hospedante N miristoil transferasa y miristoil-CoA como sustrato. La reacción remueve la metionina terminal de la cadena peptídica por medio de la metionina aminopeptidasa del hospedante (144). El miristato es un ácido graso de catorce carbonos y confiere a la proteína la capacidad de anclarse en la membrana, y así facilitar la interacción de Z con otras proteínas virales, así como su autoensamblado (143, 145, 146). Se han identificado aminoácidos básicos en el dominio N terminal seguidamente del sitio de miristoilación (136), que podrían llegar a asociarse a fosfolípidos de la membrana y mejorar el anclaje de la proteína en la membrana plasmática.

2.1.2 Replicación viral

La proteína Z es un elemento regulatorio de la replicación viral. Principalmente, es un inhibidor de la replicación por medio de su interacción directa con la ARN polimerasa ARN dependiente L (147). Esta inhibición es dependiente de la concentración de la proteína Z respecto de la proteína L (148). Tanto para LCMV como para LASV, se observó que Z requiere de un dominio RING intacto, mientras que los dominios N y C terminales no son esenciales para esta función (146, 149). En particular, los residuos involucrados en la coordinación de los cationes zinc y un triptófano altamente conservado dentro del dominio RING son esenciales para la actividad inhibitoria de Z en la síntesis de ARN (150).

2.1.3 Brotación y ensamblado viral

La proteína Z interacciona con las glicoproteínas virales (145, 151), llegando a formar en conjunto VLPs (151), lo cual coincide con la organización estructural de partículas de arenavirus observadas por medio de criomicroscopía electrónica (152). La miristoilación N terminal de Z es fundamental para la interacción con las glicoproteínas, no así los dominios RING y C terminal (145). Esta interacción se da con SSP aún en ausencia de GP1 y GP2 (145).

Tanto el empaquetamiento como la brotación de partículas virales son debidos a la proteína Z. De hecho, la proteína sola genera VLPs. Esta actividad de brotación ha sido atribuida a los dominios tardíos (153, 154). El dominio P[T/S]AP accede a la vía ESCRT por medio de la interacción con el componente Tsg101 (gen de la susceptibilidad tumoral 101) de ESCRT-I, una de las cuatro poliproteínas de la maquinaria de la vía (155-157), mientras que el motivo YxxL media la unión con Alix/AIP1, el cual une a ESCRT-I y III (158-160). Al menos en LCMV y LASV, se ha demostrado que Tsg101 está involucrado en la brotación (154, 161). No obstante, la proteína Z de TCRV induce brotación inclusive cuando carece de los dominios P[T/S]AP y PPxY clásicos, y no requiere Tsg101 (162).

2.1.4 Regulación de funciones celulares mediada por Z

La proteína Z puede interactuar con diferentes factores celulares para favorecer la replicación viral, incluyendo la proteína de la leucemia promielocítica (PML) (163), la fracción nuclear de la proteína ribosomal P0 (164), el factor de iniciación de traducción eucariota 4E (eIF4E) (165) y la proteína del homeodominio rica en prolina (PRH) (166).

La proteína PML es un componente importante de los cuerpos nucleares (NB), un tipo de organela sin membrana que pueden ser alterados durante condiciones de enfermedad o estrés, tales como una infección viral (167, 168). Se observó que durante la infección con LCMV, PML se relocaliza del núcleo al citoplasma donde forma grandes cuerpos. Las proteínas Z de LCMV y LASV son suficientes para redistribuir PML al citoplasma y formar complejos específicos *in vitro* (163). PML promueve la apoptosis (169-171), de manera que su interacción con Z permitiría evadir la apoptosis de la célula hospedante durante la infección. Adicionalmente, se ha asociado PML con un rol en la defensa antiviral (172, 173), por lo que la interacción con Z podría ser otro factor que favorece la evasión de la respuesta inmune innata de los mamarenavirus.

Las proteínas ribosomales P (P0, P1 y P2) están involucradas en varias funciones importantes de los ribosomas, formando parte de la subunidad ribosomal mayor. Entre sus funciones se incluye la unión al ARN 28S y asociación con factores de elongación eucariota y aminoacil ARNt. Durante la infección con LCMV, se observó colocalización entre la proteína Z y las proteínas ribosomales P (164). Además, los niveles de expresión de P1 y P2 resultaron significativamente reducidos, a diferencia de P0 que mantuvo su nivel de expresión. Curiosamente, P0 fue encontrado dentro de viriones de LCMV, sugiriendo una unión fuerte con Z (164).

El factor de iniciación de la traducción eucariota eIF4E facilita la unión de los ribosomas a las estructuras CAP 5' de ARN mensajeros, favoreciendo así la traducción dependiente de eIF4E (174). Se ha observado que la proteína Z de LASV y LCMV selectivamente reprime la expresión proteica por medio de su interacción con eIF4E tanto en células infectadas como transfectadas (165). La proteína Z no parece afectar los niveles de expresión de eIF4E o su estabilidad (139). Por medio de estudios estructurales, se determinó que la proteína Z se une a la superficie dorsal del factor eIF4E, induciendo cambios conformacionales en la superficie distal (la cual cuenta con un sitio ligador de CAP). Esto se logra con los residuos dentro del primer sitio coordinador de zinc en el dominio RING (136).

La leucemia promielocítica aguda (LPA) se caracteriza por un bloqueo en la diferenciación de las células mieloides. Como resultado de una translocación cromosómica en estos pacientes, la proteína PML se altera al igual que los NB. La alteración de PML y de los NB en la LPA está relacionada con una pérdida del control del crecimiento y la posterior leucemogénesis. Utilizando técnicas de 2 híbridos de levadura, se observó que PML y la proteína Z arenoviral, se unen a la proteína PRH a través de sus dominios RING. PRH funciona en la hematopoyesis y puede actuar como un represor transcripcional. PRH tiene un patrón puntiforme tanto en el núcleo como en el citoplasma de las células K562 de leucemia mielógena crónica y en la línea celular LPA, NB4. Los estudios de inmunoprecipitación y colocalización indican que PML y PRH interactúan en ambas líneas celulares. El efecto sobre el crecimiento celular de PML y las acciones hematopoyéticas de PRH plantea la posibilidad de que la interacción entre PML y PRH represente un vínculo entre el control del crecimiento y la hematopoyesis (166).

2.1.5 Evasión de la respuesta inmune innata

Los interferones tipo I son mediadores secretados en vertebrados que inducen un estado antiviral en las células diana, limitando la replicación viral y la propagación del virus. No sorprende entonces que los virus hayan desarrollado mecanismos para contrarrestar este sistema antiviral. De hecho, la mayoría de los mammarenavirus cuentan con mecanismos para inhibir la producción de IFN-I (175-178). Se ha demostrado que la proteína Z puede interactuar directamente con RIG-I y MDA5, bloqueando la interacción de estas proteínas con MAVS y, de esta manera, inhibir los eventos de señalización necesarios para la inducción de IFN-I (179, 180). Algunos autores han propuesto que únicamente las proteínas Z de las especies patógenas de mammarenavirus utilizan este mecanismo (180).

2.2 Resultados

2.2.1 Secuencia

En la Tabla 6 se resumen las observaciones resultantes de los alineamientos comparativos de las secuencias nucleotídica y aminoacídica de la proteína Z de la cepa P3441 con las conocidas para otras cepas de JUNV cepas mencionadas en la Introducción general.

Tabla 6. Porcentaje de identidad y número de nucleótidos y aminoácidos respecto a P3441

	Nucleótidos		Aminoácidos	
	Similitud (%)	Diferencias (nº nts)	Similitud (%)	Diferencias (nº aa)
Candid 1	96	9	97	2
XJ13	96	9	97	2
P35032	94	15	97	2
MC2	95	13	100	0
Romero	95	13	100	0
Ledesma	94	15	98	1
Espindola	96	11	98	1

Entre las cepas estudiadas no hubo más de 2 aminoácidos de diferencia, mientras en otras no hubo ninguna. La mayoría de las diferencias se encontraron dentro de los primeros 20 aminoácidos, en el dominio N terminal de la proteína (Figura 2.2).

Con el fin de predecir el impacto de estas diferencias presentes en la estructura de la proteína Z de P3441, se generó una estructura tridimensional de JUNV Z, utilizando el programa de inteligencia artificial AlphaFold2 por medio del servidor ColabFold (181). De esta manera, se obtuvieron 5 modelos de JUNV Z para la cepa P3441.

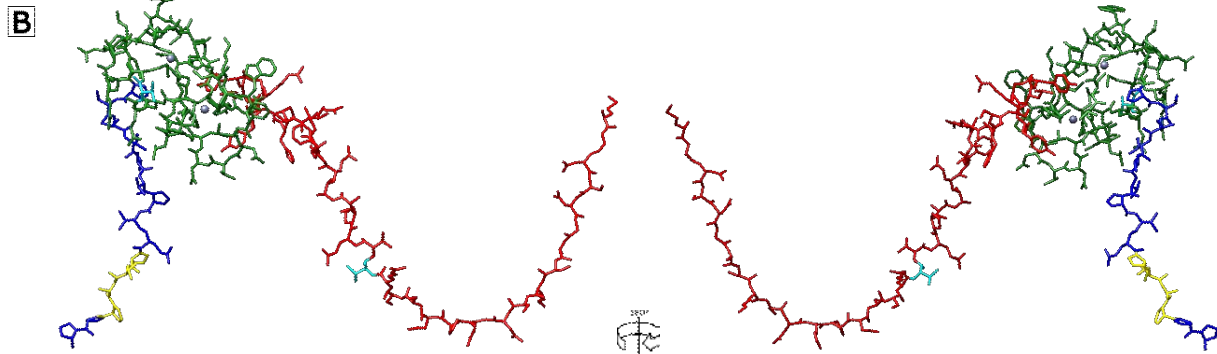
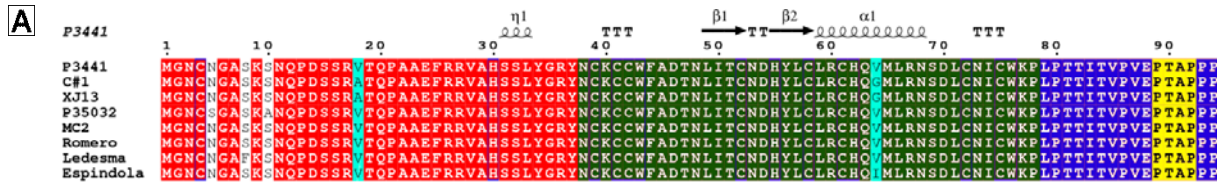


Figura 2.2. (A) Alineamiento de secuencias de amino ácido para la proteína Z de Junín de la cepa P3441, C#1 y otras cepas patogénicas. Las equivalencias se encuentran marcadas en color, y la estructura secundaria de la proteína puede observarse por encima del alineamiento con α indicando alfa-hélices, β lámina beta, η hélices 3_{10} , y T para giro beta. (B) Estructura tridimensional de la proteína JUNV Z obtenido por ColabFold (ver materiales y métodos). Cada imagen corresponde a lados opuestos de la proteína. Se muestran los dominios N terminal (rojo), C terminal (azul), RING (verde) y tardío L (amarillo). En cian se muestran los sitios de interés 18 y 64 y como esferas grises los cationes Zn coordinados en el dominio central.

Los modelos obtenidos fueron validados utilizando los programas PROCHECK y VERIFY3D (ver materiales y métodos) dentro del sitio web SAVES v6.0. Las estructuras fueron posteriormente refinadas por medio del servidor ReFOLD3 (182). La presencia de átomos coordinados dentro de la estructura de la proteína (en este caso, cationes zinc en el dominio RING) fueron verificados y, llegado el caso, agregados por medio de los programas PyMol y Chimera. El modelo con la mejor puntuación (*score*) de refinamiento para cada algoritmo fue utilizado para predecir el impacto de mutaciones dentro de la estructura. Por medio del servidor MAESTROweb (183) se pudo visualizar como diferencias de otras cepas respecto a P3441 podrían ser mutaciones estabilizadoras o afectar en la estabilidad estructural del modelo presentado. Los resultados se muestran en la Tabla 7.

Tabla 7. Cambio de estabilidad ($\Delta\Delta G_{\text{pred.}}$) calculada como mutación independiente de un residuo en la estructura para la proteína Z de P3441 a un residuo en otra cepa (X).

Mutación (P3441 > Cepa X)	$\Delta\Delta G_{\text{pred.}}$	$C_{\text{pred.}}$
N5S (X = P35032)	-0,131	0,882
S8A (X = Ledesma)	-0,108	0,893
S10A (X = P35032)	-0,063	0,920
V18A (X = C#1)	-0,125	0,889
V64G (X = C#1)	0,367	0,835
V64I (X = Espindola)	-0,280	0,886

Valores de $\Delta\Delta G_{\text{pred.}} > 0$ se consideran desestabilizadores. $C_{\text{pred.}}$ Es un valor de confianza de la predicción, siendo $C_{\text{pred.}} = 0$ de mínima confianza e $C_{\text{pred.}} = 1$ máxima confianza de predicción.

Las predicciones obtenidas poseen una confianza aceptable ($C_{pred} > 0,8$). En la mayoría de los casos parecería que las diferencias en las otras cepas son estabilizadoras respecto a P3441, aunque no serían una influencia importante. En contraste, el análisis para V64G predice una diferencia desestabilizadora ($\Delta\Delta G_{pred} > 0$), la cual podría impactar en la flexibilidad y estabilidad de la estructura secundaria del dominio RING.

Para profundizar este hallazgo, se utilizó el servidor DynaMut (184) que, a diferencia de MAESTROweb, intenta incluir otros factores en su análisis, tal como la dinámica molecular, comparando el efecto de mutaciones para diferentes conformaciones (Figura 2.3).

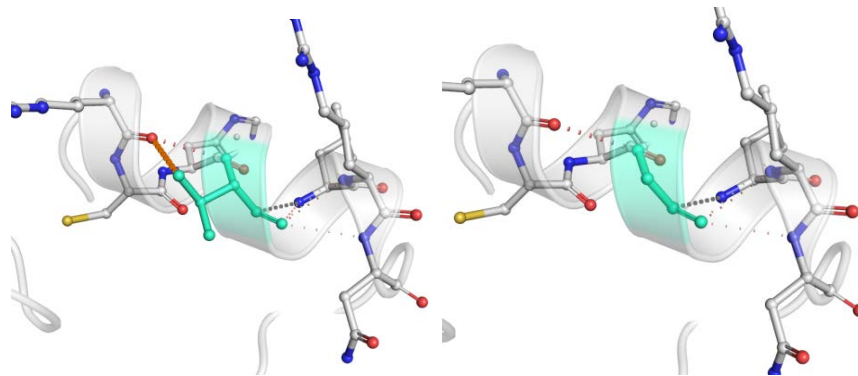


Figura 2.3. Esquema del residuo 64 en el dominio RING de JUNV Z. A la izquierda se encuentra la estructura para P3441 y otras cepas patogénicas. A la derecha se encuentra la estructura para C#1. En verde se ve la valina (izquierda) y glicina (derecha) correspondiente a cada cepa. En gris se ven los carbonos (C), en rojo los oxígenos (O) y en azul los nitrógenos (N). En amarillo se muestran átomos de azufre (S).

La predicción de DynaMut fue de un $\Delta\Delta G = -0,510$ kcal/mol, lo que significa una mutación desestabilizadora. Otros métodos de predicción de $\Delta\Delta G$ que DynaMut calcula en paralelo resultaron en valores negativos y, en algunos casos, mayores. Adicionalmente, el servidor predijo un aumento en la diferencia de entropía vibracional ($\Delta\Delta S > 0$), que corresponde a un aumento en la flexibilidad molecular.

2.2.2 Clonado en vector de expresión eucariota

Con el objetivo de subclonar Z del vector *pGEM-T Easy Vector* a un vector de expresión eucariota que permitiese expresar proteína recombinante con un fluoróforo C terminal, se digirió el plásmido pGEM-Z obtenido previamente (capítulo 1) con las enzimas de restricción HindIII y BamHI. Finalmente, se ligó el fragmento liberado del vector de entrada en el plásmido linealizado pEGFP (Figura 2.4). El diseño experimental replica lo alcanzado por la Dra. Paula Arrías para su trabajo final de grado, en el cual logró clonar la proteína en pEGFP.

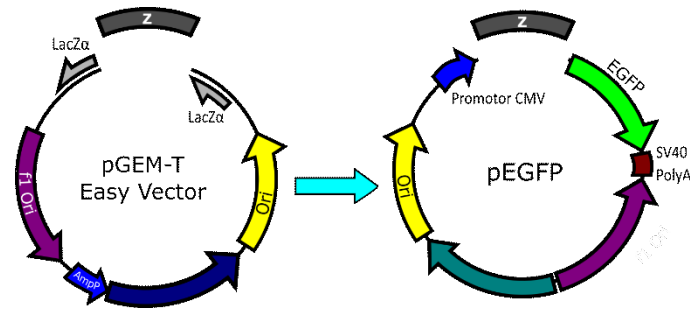


Figura 2.4. Esquema del clonado del marco de lectura para la proteína Z de JUNV en plásmidos pEGFP. Se partió de un vector que permitía el clonado directo de fragmentos amplificados de PCR para poder realizar la posterior digestión con las enzimas de restricción HindIII y BamHI y seguido de la ligación en el vector de expresión eucariota empleando ligasa T4.

Los plásmidos resultantes fueron confirmados por digestión y secuenciación, se realizaron *minipreps* y luego se transfectaron en células A549 para realizar un análisis por microscopía confocal. Mientras que el producto del plásmido vacío se distribuyó en las células, concentrándose principalmente en el núcleo, se pudo observar que las proteínas recombinantes Z-EGFP presentaban una distribución puntual y citosólica, como era esperable para esta proteína viral (Figura 2.5).

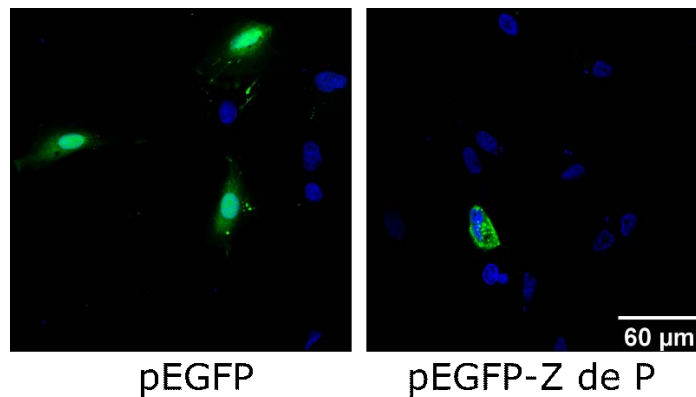


Figura 2.5. Imágenes de fluorescencia confocal de células A549 transfectadas con el plásmido pEGFP vacío o con el ORF de Z de P. En azul se observan los núcleos celulares por medio de tinción de ADN doble-cadena con verde metilo. El verde corresponde a la fluorescencia del fluoróforo EGFP.

2.2.3 Localización, interacciones y antigenicidad

Como fue descrito previamente, Z es una proteína con varias funciones. Entre ellas, interactúa con varias proteínas y actúa sobre varias vías de señalización. Con el fin de poder estudiar la utilidad de los plásmidos generados, se siguió la línea de investigación previa de nuestro grupo de trabajo con la proteína PML y la evasión de la respuesta inmune. En primer lugar, se infectaron células A549 con las cepas C#1 y P3441 y a 1 dpi se realizaron ensayos de inmunofluorescencia con el objetivo de estudiar cómo afectaba la formación de NB utilizando anticuerpos específicos para PML y para JUNV N (para reconocer a las células infectadas) (Figura 2.6). Posteriormente se transfectaron células A549 con los plásmidos de EGFP con la proteína Z y se realizó la misma marcación para PML (Figura 2.7). En contraste a lo observado en LCMV, no hubo translocación de PML-NB al citoplasma durante la infección con JUNV y la expresión de Z no modificó la distribución de los PML-NB, al menos a 1 dpt.

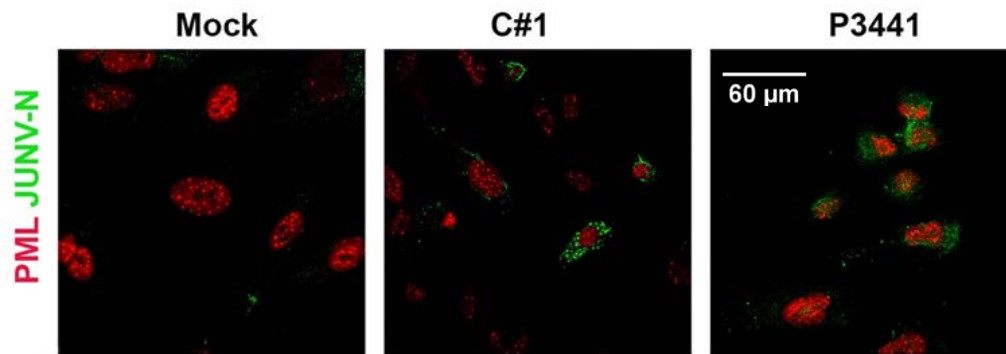


Figura 2.6. Imágenes de fluorescencia por microscopía confocal para células A549 no infectadas (Mock) o infectadas con JUNV (cepa C#1 o P3441) a 1dpi. En verde se encuentra la fluorescencia correspondiente a FITC de la proteína N de JUNV y en rojo corresponde a AF647 para la proteína PML.

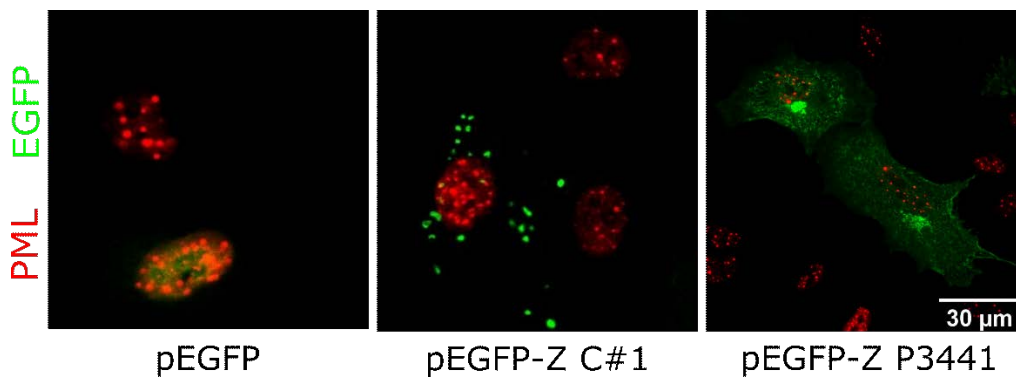


Figura 2.7. Imágenes de fluorescencia por microscopía confocal para células A549 transfectadas con el plásmido pEGFP o pEGFP-Z de C#1 o P3441. En verde se encuentra la fluorescencia correspondiente al fluoróforo EGFP vacío o de la proteína Z recombinante y en rojo la fluorescencia AF647 correspondiente a la localización de PML.

Finalmente, se estudió si las proteínas recombinantes Z presentaban características antigénicas similares a las virales. Se repitieron las transfecciones en células A549 con los plásmidos y se realizaron ensayos de inmunofluorescencia utilizando anticuerpos específicos de la proteína Z. No obstante, el anticuerpo secundario presentó mucha marcación inespecífica y como el otro anticuerpo secundario disponible tenía la misma fluorescencia (FITC), no se pudo realizar ese estudio.

2.3 Discusión

Por primera vez se pudo determinar que Z de P3441 difiere con otras cepas en base de datos, tanto atenuadas como patogénicas. No obstante, no hubo diferencias con las cepas Romero y MC2. Adicionalmente, no se encontraron más de 2 aminoácidos de diferencia entre cepas, lo cual se explica por el tamaño reducido de esta proteína que limita el número de mutaciones potenciales existentes. También se debe considerar que ante las múltiples funciones de la proteína Z, probablemente exista una presión selectiva en gran parte de la proteína.

La mayoría de las diferencias se encontraron dentro de los primeros 20 aminoácidos, en el dominio N terminal de la proteína, donde no hay estructura secundaria estable. Es posible que estos residuos tengan un impacto limitado en la estabilidad de la proteína, no así es el caso para las diferencias en

el sitio 64, que se encuentra dentro del dominio RING, más específicamente en una estructura alfa hélice.

Utilizando herramientas bioinformáticas, se realizaron predicciones del impacto de esta diferencia en el residuo 64. Acorde a lo obtenido, parecería que aunque la glicina de C#1 y la valina de P3441 presentan propiedades químicas similares, el menor tamaño de glicina confiere mayor libertad y por ende desestabiliza la estructura secundaria. Esto es, en P3441 el dominio RING tendría mayor estabilidad estructural. Las similitudes entre ambos tipos de aminoácidos (no polares, carencia de carga, etc.) podría explicar por qué la proteína de Z para C#1 puede desenvolver sus funciones esenciales para mantener la replicación viral mientras al mismo tiempo tener una evasión de la respuesta antiviral de la célula infectada atenuada o más ineficiente, lo que explicaría parcialmente o completamente los fenotipos de infección observados.

Los plásmidos generados para P3441, que expresan el gen de la proteína Z unido al fluoróforo EGFP bajo control de un promotor eucariota, resultaron útiles para estudiar la proteína Z al poder observar que la sobreexpresión de Z no afectaba la distribución de los cuerpos nucleares, al menos a 1 dpt. Finalmente, se estudió si las proteínas recombinantes Z presentaban características antigénicas similares a las virales. Se repitieron las transfecciones en células A549 de los plásmidos y se realizaron ensayos de inmunofluorescencia utilizando anticuerpos específicos para la proteína Z. No obstante, limitaciones técnicas debidas a la gran inespecificidad del anticuerpo secundario no permitieron la realización de dicho estudio.

Para continuar los análisis de Z, se planifican realizar estudios a tiempos más tardíos junto con la proteína PML, a fin de confirmar si la diferencia entre las secuencias de C#1 y P afectan en la unión/interacción con PML.

Capítulo 3. Proteína N

3.1 Introducción

La proteína de nucleocápside N de los mammarenavirus oscila entre los 558 y 593 aminoácidos de largo. En el caso de JUNV, esta proteína posee 564 aminoácidos y un peso molecular de 65 kDa. La proteína N de los mammarenavirus está asociada a la replicación viral (113, 185) y a la evasión de la respuesta inmune (175, 186, 187).

3.1.1 Estructura de la proteína N

La proteína N presenta un dominio N terminal globular que se une al ARN y un dominio C terminal con actividad exonucleasa (Figura 3.1). Ambos dominios se encuentran divididos por una pequeña región flexible (*linker*) de aproximadamente 30 aminoácidos de largo (188, 189). En el dominio N terminal se encuentra la región que confiere la capacidad de unión a otras proteínas N (190, 191), mientras que en el dominio C terminal se encuentra el sitio de unión a la proteína Z (192), la actividad antagonista a interferón, asociada principalmente a un motivo DIEG(R) (en JUNV DIEG) (175, 187, 193-195), y una estructura exonucleasa en un motivo DEDDh (189, 196).

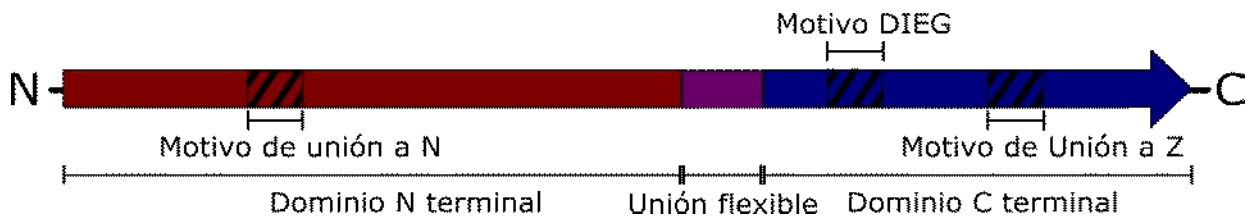


Figura 3.1. Diagrama representativo de los dominios dentro de la estructura de N. El dominio N terminal (rojo) incluye una secuencia (motivo) involucrado en la interacción entre proteínas N. El dominio C terminal (azul) cuenta con una secuencia DIEG asociada a la interacción con IKKe y un motivo asociado a la unión con la proteína Z.

3.1.2 Replicación viral

La proteína N recubre las moléculas de ARN viral, protegiéndolas de la degradación e impidiendo la formación de ARN doble cadena con ARN de polaridad opuesta. No obstante, también interactúa con proteínas del hospedante para favorecer su ciclo viral.

La proteína N puede unirse a sí misma y formar estructuras multiméricas. Esta habilidad se ha asociado a una región dentro de los primeros cincuenta aminoácidos del extremo N terminal de la proteína (190). Es por medio de estas uniones que se coordina la estructura de las ribonucleoproteínas, las cuales se asocian a la proteína L para generar estructuras multiméricas flexibles (197) formando las unidades biológicamente activas para la transcripción de ARN mensajero viral y la replicación del genoma (113, 114, 198). De hecho, existe evidencia que la replicación de JUNV ocurre en estructuras puntuales discretas dentro del citoplasma denominadas complejos de replicación-transcripción o RTC (del inglés, *replication-transcription complexes*), donde se localizan ARNs y proteínas N recién sintetizadas (199).

La proteína N es un factor importante y necesario para la formación y estructura de los RTC (200-202). En LCMV se ha sugerido que la fosforilación de un único residuo conservado (T206) de N guía la nucleación de los RTC (203). En células infectadas con TCRV, los RTC también presentan varias proteínas ribosomales como también los factores de iniciación eucariotas eIF4G y eIF4A (199). Finalmente, en LASV se ha encontrado evidencia que N podría tener un rol en la optimización de la

síntesis de ARN y en la manipulación y control de ARN viral erróneo, como también de ARN mutante (204).

3.1.3 Regulación de funciones celulares mediada por N

Por medio de experimentos de coimmunoprecipitación, se ha demostrado interacción entre N con la proteína G3BP1, la cual se ubica en los gránulos de estrés citoplasmáticos (205), sugiriendo que N recluta esa proteína. Considerando que los gránulos de estrés son focos de degradación de ARN mensajero en situaciones de estrés celular, podría especularse que esto podría servir como fuente de CAPs necesarios para iniciar los procesos de traducción. Sin embargo, se ha visto que JUNV inhibe la formación de gránulos de estrés inhibiendo la fosforilación de eIF2 α cuando la célula se estresa con tratamiento de arsenito (206), por lo que resta aclarar este supuesto rol de la proteína N.

La proteína N también interactúa con las ribonucleoproteínas nucleares heterogéneas del huésped (HnRNPs), algo que se ha observado tanto para JUNV como para otros mammarenavirus (205, 207-209). Estas proteínas participan en el procesamiento del preARNm y en la traducción, tráfico y estabilidad del ARNm (210, 211) y tendrían un importante rol en el ciclo de replicación viral. Acorde a estudios de silenciamiento de las hnRNP A1 y A2 previos a la infección, se observó una reducción de la síntesis proteica y de la producción de viriones de JUNV (208). Adicionalmente, la hnRNP K, que también puede unirse a la proteína N, parecería ser otro factor necesario para la multiplicación de JUNV (205, 209).

3.1.4 Evasión de la respuesta inmune innata

Los mammarenavirus inducen expresión de IFN-I por medio de diferentes mecanismos. En JUNV, se ha visto que la C#1 induce la respuesta inmune innata en macrófagos, detectándose una inducción en la producción de TNF e IFN-I (212). Similarmente, se conoce que el ARN arenaviral es detectado por los receptores citosólicos RIG-I y PKR (213, 214).

En el capítulo 2 se describió que Z interactúa directamente con RIG-I y MDA5 resultando en el bloqueo a la proteína MAVS y subsecuentemente, de la inducción de IFN-I (179, 180). Esto sólo ocurriría en los mammarenavirus patogénicos (180). En contraste, la proteína N jugaría un rol antagónico a la expresión de IFN-I en la infección de mammarenavirus tanto patogénicos como no patogénicos, como se describe a continuación.

El motivo DIEG(R) en el dominio C terminal de la proteína N interactúa con el dominio quinasa de IKK ϵ , bloqueando la fosforilación de IRF3 mediada por TBK1 (175, 187, 193, 195). También se ha identificado en TACV y LASV un dominio exonucleasa de la familia DEDDh con actividad 3'-5' exonucleasa con preferencia por ARN doble cadena (ARNdc) (196, 215). Este dominio sirve para conjugar dos cationes divalentes en su sitio activo por medio de cuatro residuos ácidos conservados (aspartato y glutamato) y un residuo base general (histidina) (216). Se ha sugerido que la proteína N digiere al ARNdc que puede llegar a servir como un patrón molecular asociado a patógeno (PAMP) para sensores de ARN citosólicos como RIG-I y MDA5 (188, 195, 215). No obstante, estudios en los cuales se utilizó un anticuerpo monoclonal contra ARNdc de alta afinidad como el 9D5, se demostró que los mammarenavirus NW realizan una discreta acumulación citosólica de ARNdc durante la infección en contraste con los mammarenavirus OW, donde esta acumulación no se produce (217), lo cual favorecería una respuesta antiviral de la célula huésped en el primer caso a detrimento de la replicación y propagación viral como sería el segundo caso. TACV sería una excepción, debido a que una mutación dentro del dominio exonucleasa que no afecta el motivo DIEG(R), no redujo la

capacidad para inhibir la respuesta de IFN-I, sugiriendo que otros factores presentes en la proteína N tienen mayor rol en el antagonismo a IFN-I (218).

3.2 Resultados

3.2.1 Secuencia

Se compararon las secuencias nucleotídicas y aminoacídicas de la proteína N de P3441 con las conocidas para otras variantes. En la Tabla 8 se resumen las observaciones.

Tabla 8. Porcentaje de identidad y número de nucleótidos y aminoácidos respecto a P3441.

	Nucleótidos		Aminoácidos	
	Similitud (%)	Diferencias (nº nts)	Similitud (%)	Diferencias (nº aa)
Candid 1	95	69	98	8
XJ13	97	42	99	5
P35032	96	58	98	6
MC2	94	97	96	19
Romero	97	36	99	4
Ledesma	96	55	99	3
Espindola	97	39	98	6

La cepa que presentó el mayor número de diferencias fue la MC2, con un porcentaje de identidad de aminoácidos del 96% que reflejan 19 aminoácidos de diferencia (Figura 3.2). La cepa más similar a P3441 fue la cepa Ledesma seguida de la cepa Romero, las cuales corresponden a cepas obtenidas de un caso fatal de FHA con síntomas neurológicos y a un paciente no letal de FHA, respectivamente. Con el interés de determinar factores que intervengan en la patogenicidad, se analizó con mayor énfasis los sitios donde hubo diferencias en las cepas C#1 y/o XJ13 respecto a las cepas patogénicas. Estos sitios fueron marcados en cian (Figura 3.2). Se aceptaron sitios donde MC2 compartía diferencias con C#1 respecto a P3441, ya que podrían ser sitios que podrían participar en la patogenicidad intermedia de esta cepa. Los 7 sitios de interés (cian), 4 se encuentran en el dominio N terminal y 3 en el dominio C terminal. C#1 comparte con MC2 las diferencias en los residuos 158, 268 y 322 del dominio N terminal y comparte con XJ13 las tres diferencias de los sitios de interés en el dominio C terminal.

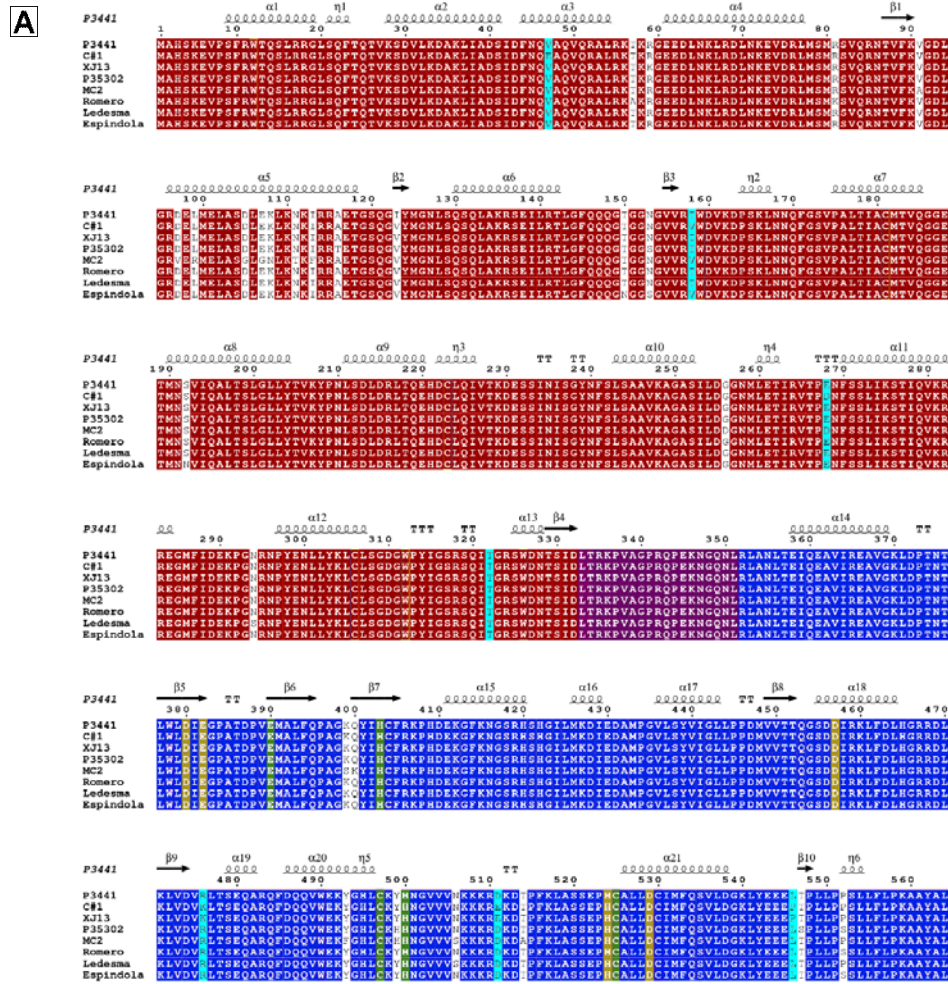


Figura 3.2. (A) Alineamiento de secuencias de aminoácido para la proteína N de JUNV de la cepa P3441, C#1 y otras cepas. Las equivalencias se encuentran marcadas en color, con las diferencias presentes en blanco o cian, y la estructura secundaria de la proteína puede observarse por encima del alineamiento, con α indicando alfa hélices, β lámina beta, η hélices 3_{10} , y T para giro beta mientras los dominios N terminal en rojo, el C terminal en azul, y la región unidora flexible en púrpura. (B) Estructura tridimensional de la proteína JUNV N obtenido por ColabFold (ver materiales y métodos). Cada imagen corresponde a lados opuestos de la proteína. Se muestran los dominios N terminal (rojo), C terminal (azul), región de unión flexible (púrpura), el motivo DEDDh (amarillo) y los residuos coordinadores de un segundo catión (verde). En cian se muestran los sitios con diferencias de interés 47, 158, 268, 322, 476, 511 y 547 y los cationes Zn coordinados por la proteína están representados como esferas grises.

De los sitios de interés, 6 estarían en giros o zonas flexibles, siendo el sitio 47 el único que se encuentra en una estructura secundaria definida (alfa hélice). Aparte de eso, los residuos parecen estar en regiones expuestas de la estructura.

Seguidamente, se analizaron cuales mutaciones de P3441 podrían impactar en la estabilidad estructural de la proteína. Para ello se generaron estructuras tridimensionales para luego analizarlas siguiendo los pasos mencionados en el capítulo 2.

El modelo que obtuvo la mejor puntuación de calidad (*score*) en los diferentes servidores web utilizados, fue utilizado posteriormente para un análisis de los sitios de interés. Utilizando el servidor MAESTROweb, se determinó la energía de estabilización final para cada sitio de interés, calculándose como mutación individual del residuo en P3441 a aquella correspondiente a C#1. Los resultados se muestran en la Tabla 9.

Tabla 9. Cambio de estabilidad ($\Delta\Delta G_{pred.}$) calculada como mutación independiente de un residuo en la estructura para la proteína N de P3441 a un residuo en otra cepa (X).

Mutación (P3441> Cepa X)	$\Delta\Delta G_{pred.}$	$C_{pred.}$
V47E (X = C#1)	-0,166	0,871
I158V (X = Espindola, MC2, C#1)	-0,105	0,928
E268D (X = MC2, C#1)	0,632	0,727
T322I (X = MC2, C#1)	-0,130	0,908
R476K (X = XJ13, C#1)	0,916	0,716
D511E (X = XJ13, C#1)	-0,124	0,936
L546P (X = XJ13, C#1)	0,388	0,936

Valores de $\Delta\Delta G_{pred.} > 0$ se consideran desestabilizadores. $C_{pred.}$ Es un valor de confianza de la predicción, siendo $C_{pred.} = 0$ de mínima confianza e $C_{pred.} = 1$ máxima confianza de predicción.

Las predicciones contaron con una confianza aceptable ($C_{pred.} > 0,8$) para la mayoría de los casos. Cuatro sitios de interés fueron predichos como mutaciones estabilizadoras ($\Delta\Delta G_{pred.} < 0$) para C#1, aunque sus valores no sugieren una influencia considerable. Por otro lado, tres de las diferencias entre C#1 y P (residuos 268, 476 y 546) hacen la estructura de N para C#1 más inestable ($\Delta\Delta G_{pred.} > 0$), y en mayor magnitud a las mutaciones estabilizadoras calculadas.

Con el servidor MAESTROweb se predijo que la estructura de N para C#1 es más inestable en su totalidad respecto a la estructura de P3441 ($\Delta\Delta G_{pred} = 0,169$) y la confianza predicha ($C_{pred.}$) fue de 0,768. Cuando se estudió lo mismo por dominio, se predijo que el dominio N terminal (residuos 47, 158, 268 y 322) de la estructura de C#1 era más estable respecto a P3441 ($\Delta\Delta G_{pred} = -0,233$) con una confianza predicha de 0,967. Por el contrario, se predijo una estructura más inestable para el dominio C terminal (residuos 476, 511 y 546) de C#1 y XJ13 ($\Delta\Delta G_{pred} = 0,722$) con una confianza de 0,792. A continuación, se utilizó el servidor DynaMut para continuar el análisis de las diferencias entre C#1 y P. Los resultados se muestran en la Tabla 10. En los cálculos de este servidor, $\Delta\Delta G < 0$ son para mutaciones desestabilizadoras, opuesto a la manera que se calcula con MAESTROweb.

Tabla 10. Cambio de estabilidad ($\Delta\Delta G_{\text{pred}}$ DynaMut) y entropía vibracional ($\Delta\Delta S_{\text{vib}}$) calculada como mutación independiente de un residuo de la proteína N de P3441 a un residuo de otra cepa (X).

Mutación (P>C#1)	$\Delta\Delta G_{\text{pred}}$. DynaMut	$\Delta\Delta S_{\text{vib}}$
V47E	-0,411	-0,077
I158V	-0,844	0,498
E268D	-0,354	0,091
T322I	-0,069	0,257
R476K	-0,289	-0,002
D511E	-0,063	0,206
L546P	-1,3898	0,438

$\Delta\Delta G_{\text{pred}}$. DynaMut corresponde a la diferencia de energía libre entre el modelo salvaje y el mutante calculado por el servidor DynaMut, el cual define valores negativos (< 0) como mutación desestabilizadora. $\Delta\Delta S_{\text{vib}}$ es la diferencia de energía de entropía vibracional entre el modelo salvaje y el mutante, con valores positivos (> 0) para mutaciones que aumentan la flexibilidad molecular.

A diferencia de lo obtenido con MAESTROweb, los resultados sugieren que la estructura de N de C#1 es estructuralmente más inestable que para P3441 en su totalidad. Para todos los sitios de interés, la mutación de la secuencia de P3441 a C#1 resulta en inestabilidad ($\Delta\Delta G < 0$). No obstante, el impacto de cada mutación difiere acorde al sitio. Según los resultados, la mayoría de las mutaciones predichas aumentan la flexibilidad molecular, o en su defecto la reducen levemente. Acorde al valor de energía de desestabilización calculada, las mutaciones se pueden dividir en tres grupos: reducido ($\Delta\Delta G < 0,1$), intermedio ($0,1 < \Delta\Delta G < 0,5$) y elevado ($\Delta\Delta G > 0,5$). Las mutaciones T332I y D511E, respecto al resto, serían las menos desestabilizadoras, con valores comparativamente bajos para la diferencia de energía libre calculada. Al mismo tiempo, generarían una flexibilización de la estructura ($\Delta\Delta S_{\text{vib}} > 0$) intermedia. De manera inversa, las mutaciones V47E, E268D y R476, con valor de energía de desestabilización intermedia, reducirían levemente la flexibilidad estructural ($\Delta\Delta S_{\text{vib}} < 0$) o aumentarían la flexibilización levemente.

Las mutaciones con niveles predichos de energía de desestabilización elevados corresponden a I158V y L549P, que generarían la mayor flexibilización estructural. Considerando que DynaMut difiere del servidor MAESTROweb por considerar la dinámica conformacional dentro de sus cálculos de energía, no sorprende que la mutación L546P produce la mayor inestabilidad de las mutaciones predichas. Los residuos prolina afectan la capacidad de rotación de la cadena peptídica, lo cual impactaría en las conformaciones de la proteína. Al mismo tiempo, disminuiría la capacidad de interactuar con otros residuos (Figura 3.3), lo cual también desestabilizaría la estructura. Aunque el residuo prolina restringiría las conformaciones en el sitio de mutación, la disrupción de la estructura original explicaría el aumento de entropía vibracional correspondiente a una flexibilización estructural. Por otra parte, se observa una disminución en el número de interacciones con cadenas laterales distantes en la mutación I158V (Figura 3.3), siendo la cadena lateral de la isoleucina más larga que la de la valina.

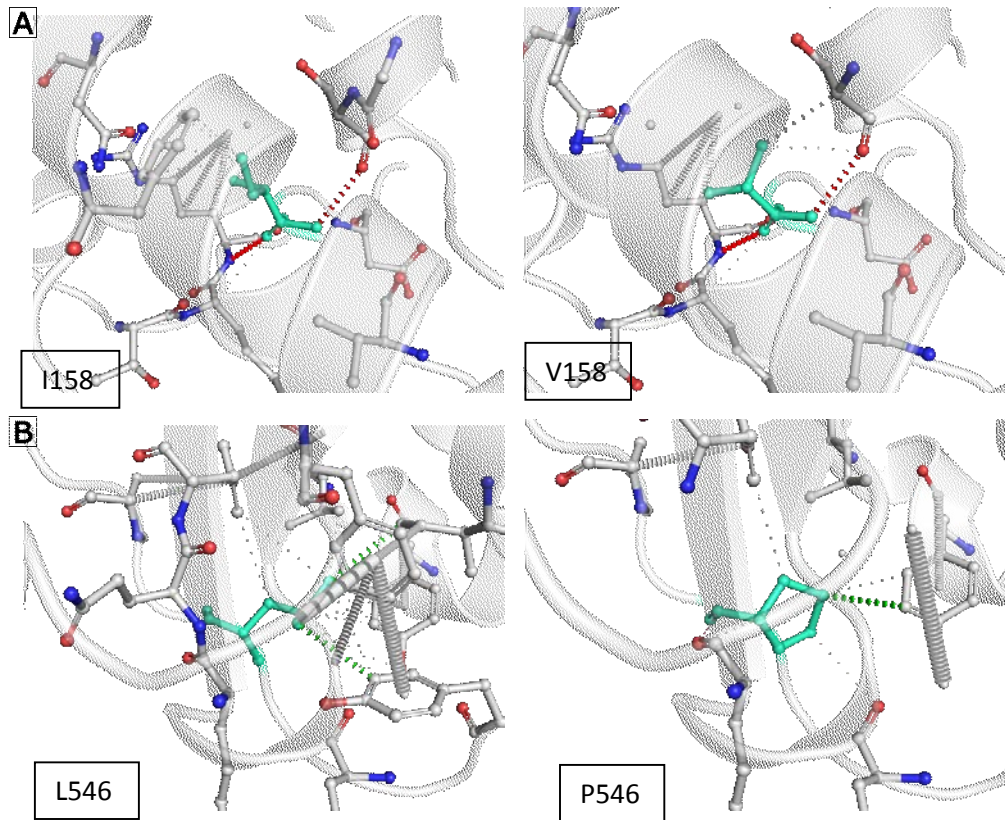


Figura 3.3. Esquema de estructura e interacciones entre cadenas laterales para los residuos 158 (A) y 546 (B) en JUNV N. A la izquierda se encuentran las estructuras para P3441 y otras cepas patogénicas. A la derecha se encuentran las estructuras para C#1. (A) Los residuos isoleucina (izquierda), valina (derecha), (B) leucina (izquierda) y prolina (derecha) se encuentran en verde. En gris se ven los carbonos (C), en rojo los oxígenos (O) y en azul los nitrógenos (N). Las interacciones con otros residuos se ven como líneas punteadas.

3.2.2 Generación de vectores eucariotas de N

Una vez clonada la proteína, se intentó subclonarla dentro de un vector de expresión bajo promotor eucariota que permitiese expresar la proteína recombinante N unida a un fluoróforo C terminal. Se utilizaron los plásmidos pGEM-N y pDsRed-X descritos en el capítulo 1, digiriendo con las enzimas de restricción NheI y KpnI. Se purificaron los fragmentos de tamaño esperado y se ligaron con la enzima T4 para así obtener los plásmidos pDsRed-N de P3441 y C#1 (Figura 3.4).

Después de varios intentos fallidos de clonar el marco de lectura de N en pEGFP-N3, se contrató el servicio de clonado de *WatsonBio Sciences* (EE.UU) para que subclonaran N utilizando las enzimas NheI y KpnI (Figura 3.4).

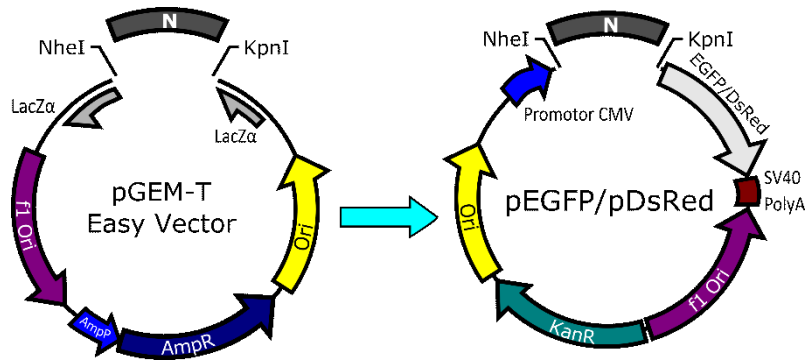


Figura 3.4. Esquema del clonado del marco de lectura para la proteína N de JUNV en plásmidos pEGFP y pDsRed. Se partió de un vector que permitía el clonado directo de fragmentos amplificados de PCR para poder realizar la posterior digestión y ligación en el vector de expresión eucariota.

Los plásmidos resultantes fueron transfectados en células A549. Las proteínas recombinantes N-DsRed y N-EGFP presentaron una distribución puntual y citosólica, como era esperable para la proteína N y a diferencia de la distribución celular (concentrándose principalmente en el núcleo) del fluoróforo DsRed y EGFP observada en las transfecciones con los plásmidos vacíos (Figura 3.5 y Figura 3.6).

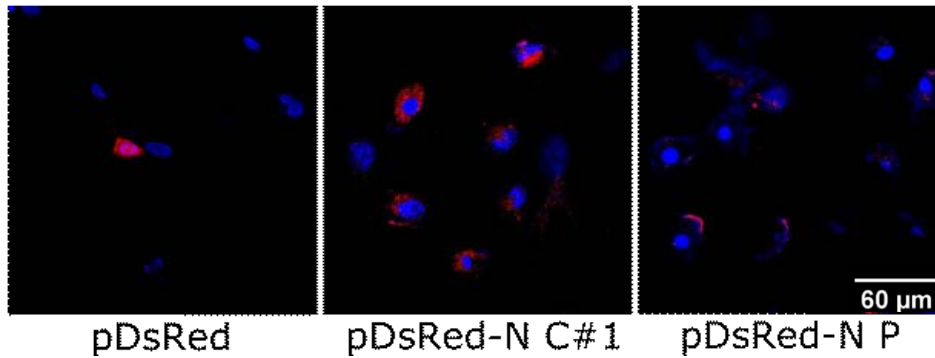


Figura 3.5. Imágenes de microscopía confocal de células A549 transfectadas con el plásmido pDsRed vacío o conteniendo N de C#1 o P (P3441). En azul se observan los núcleos celulares por medio de tinción de ADN doble-cadena con verde metilo. En rojo se encuentra la fluorescencia del fluoróforo DsRed.

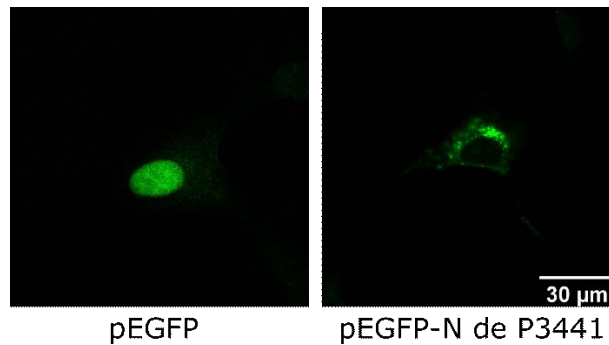


Figura 3.6. Imágenes de microscopía confocal de células A549 transfectadas con el plásmido pEGFP vacío o conteniendo N de P. En verde se encuentra la fluorescencia del fluoróforo EGFP.

3.2.3. Localización, interacciones y antigenicidad

Como fue descrito previamente, N es una proteína involucrada en varias funciones, incluyendo la evasión de la respuesta inmune innata. De manera preliminar, se estudió la distribución de la proteína N de C#1 y P3441 respecto a la proteína MAVS en células infectadas, comparándose con células no infectadas. Como se puede ver en la Figura 3.7, existe una colocalización entre la proteína N tanto de P3441 como C#1 y MAVS.

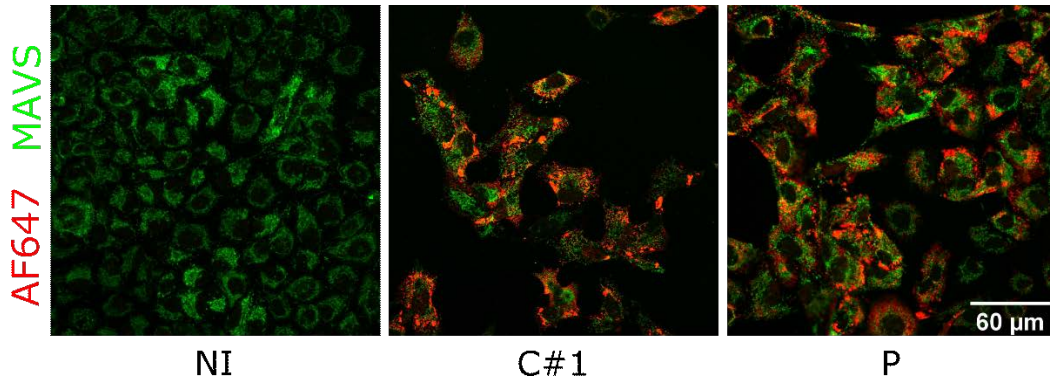


Figura 3.7. Imágenes de microscopía confocal de células A549 no infectadas (NI) o infectadas con JUNV C#1 o P3441 (P) a 3dpi. La fluorescencia verde corresponde a FITC para la proteína MAVS mientras que la fluorescencia roja corresponde a AF647 para la proteína N.

Posteriormente, se repitió el estudio de MAVS utilizando células A549 transfectadas con los plásmidos de N, donde se observó que se repetía el patrón de colocalización observado para la infección del virus (Figura 3.8).

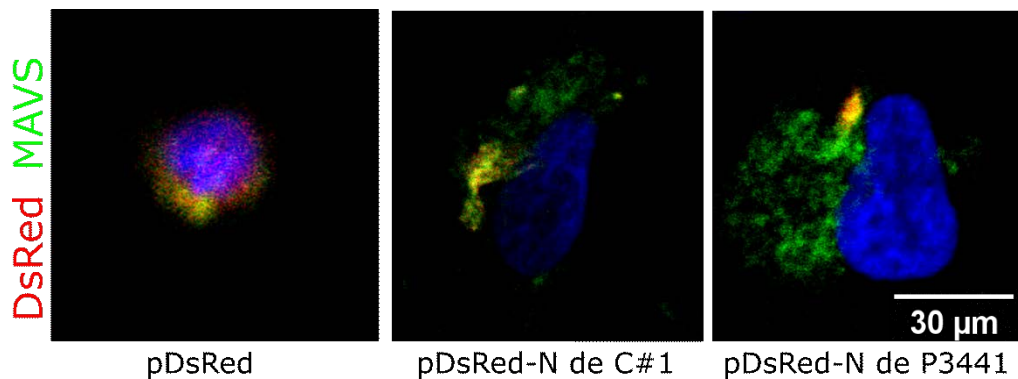


Figura 3.8. Imágenes de microscopía confocal de células A549 transfectadas con el plásmido pDsRed o pDsRed con la proteína N de C#1 o P3441 (P). La fluorescencia verde corresponde a FITC para la proteína MAVS mientras que la fluorescencia roja corresponde a DsRed para el fluoróforo vacío o de la proteína recombinante de N. La fluorescencia azul corresponde a DAPI para marcación nuclear.

Finalmente se estudió la antigenicidad de las proteínas recombinantes en ensayos de inmunofluorescencia con anticuerpos específicos para la proteína N (Figura 3.9). Se pudo observar una colocalización entre ambos fluoróforos que sugiere que las proteínas N recombinantes mantienen la capacidad de ser reconocidas por anticuerpos dirigidos contra las proteínas nativas.

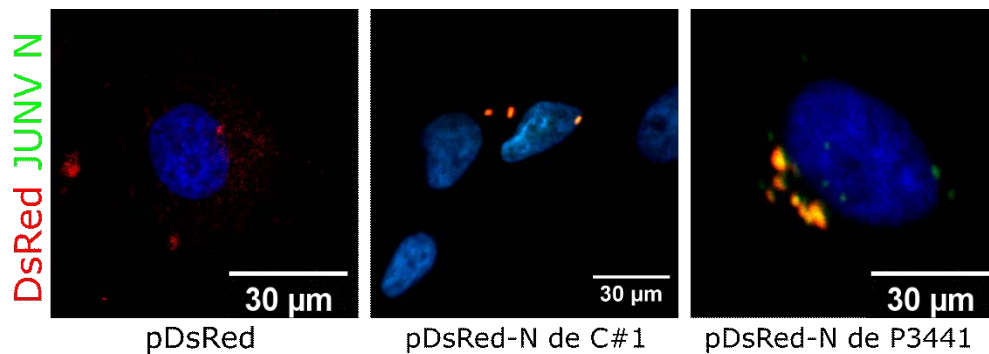


Figura 3.9. Imágenes de microscopía confocal de células A549 transfectadas con el plásmido pDsRed o pDsRed con la proteína N de C#1 o P3441 (P). La fluorescencia verde corresponde a FITC para la proteína N mientras que la fluorescencia roja corresponde a DsRed para el fluoróforo vacío o de la proteína recombinante de N. La fluorescencia azul corresponde a DAPI para marcación nuclear.

3.3 Discusión

La secuencia de la proteína N de la cepa P3441 presentó diferencias respecto a otras cepas. Estas diferencias se analizaron dentro de modelos tridimensionales por diversas herramientas bioinformáticas. Acorde a lo observado, ambos dominios de N presentan diferencias entre las cepas y no parece haber un dominio que tolere o favorezca mutaciones por encima del otro. Notablemente ninguna diferencia se encontró dentro de la región entre dominios, la cual teóricamente toleraría el mayor número de mutaciones, aunque podría deberse a su acotada extensión. El análisis futuro de más cepas permitiría confirmar las regiones de mayor conservación y las que toleren el mayor número de mutaciones.

Dentro de las diferencias observadas entre C#1 y P3441, se determinó que las posiciones 268 y 546 proveen la mayor diferencia teórica en la estabilidad entre la estructura de N. Los servidores utilizados mostraron discrepancias para los otros sitios de interés, sugiriendo que el impacto de dichos sitios depende del análisis utilizado. Acorde a nuestro análisis, las diferencias 268, 476 y 546 hacen la estructura de N para C#1 más inestable ($\Delta\Delta G_{pred.} > 0$), y en mayor magnitud a las mutaciones estabilizadoras calculadas. Esto no sorprende para el residuo 546, el cual en C#1 y XJ13 es un residuo prolina que restringe la conformación espacial de la cadena peptídica (219) y de hecho es utilizado frecuentemente para desorganizar estructuras en ensayos de mutagénesis (220). Por el contrario, las mutaciones en los residuos 268 y 476 son semejantes químicamente entre la secuencia de P3441 respecto a C#1. El aspartato (D) y glutamato (E) son aminoácidos con carga negativa o ácidos, a diferencia de la lisina (K) y la arginina (R) que son de carga positiva o básicos a pH 7. No obstante, aunque presenten similitudes químicas, tienen diferencias en el número de átomos y tamaño. De la misma manera que se ha visto diferencias entre valina (V) e isoleucina (I) entre cepas, es posible que las similitudes químicas permitan que estas diferencias entre C#1 y P mantengan la funcionalidad de la proteína N mientras se produzcan diferencias de actividad debido a las diferencias entre los aminoácidos mencionados.

Se lograron generar plásmidos para la expresión en células eucariotas de la proteína N con fluoróforo tanto para C#1 como P3441. Se verificó la antigenicidad de estas proteínas recombinantes por ensayos de inmunofluorescencia con anticuerpos específicos para la proteína N nativa.

A fin de estudiar las diferencias en la evasión de la respuesta inmune innata entre las cepas C#1 y P, se realizaron ensayos de inmunofluorescencia para la proteína MAVS y se exploró su potencial interacción con la proteína N. Se observó una colocalización mayor entre MAVS y la proteína N de P3441, lo cual se repitió cuando la proteína N se expresó individualmente por los plásmidos generados alentando su futuro estudio con mayor detalle.

Capítulo 4. Glicoproteínas de superficie

4.1 Introducción

La superficie de los mammarenavirus se encuentra recubierta de espículas conformadas por trímeros de complejos G1, G2 y SSP (221, 222). Para JUNV, el precursor GPC cuenta con 485 aminoácidos, de los cuales 58 aminoácidos corresponden a SSP, 189 a G1 y 238 a G2.

El complejo de glicoproteínas se sintetiza como un precursor (GPC), que postraduccionalmente es escindido por la peptidasa señal en una glicoproteína (GP-C) y el péptido estable señal (SSP) (223, 224). Tanto en mammarenavirus NW como OW, GP-C es posteriormente procesada por la convertasa S1P (proteasa de sitio 1), también conocida como SKI-1 (isoenzima-1 quexina subtilisina)(225-227), generando G1 y G2 dentro de las cisternas trans del Aparato de Golgi (225-228). G2 cuenta con un motivo de retención en el RE, el cual es anulado por SSP, permitiendo así que el complejo glicoproteína pueda salir del RE (229). Sólo las subunidades escindidas son incorporadas en los viriones brotantes (225), aunque el procesamiento de GPC no es esencial para el transporte a la membrana plasmática (230, 231). Residuos específicos alrededor del sitio de escisión influyen la maduración de GPC (226, 232)(Figura 4.1).

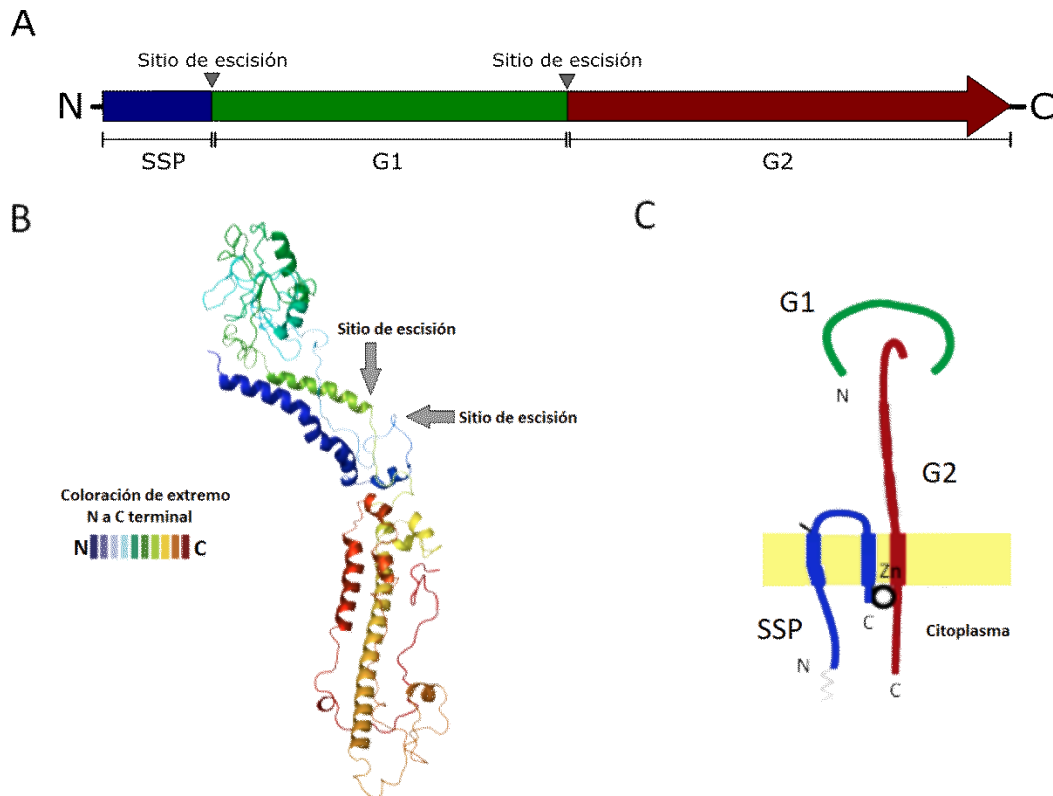


Figura 4.1. Diagrama representativo del precursor de glicoproteína GPC y su procesamiento postraduccional. A) La cadena peptídica se escinde en G1 (verde), G2 (rojo) y el péptido señal pequeño (azul). B) La figura tridimensional del precursor completo está coloreado desde el azul del extremo N terminal al rojo en el extremo C terminal. C) Esquema de la disposición de cada péptido posterior a la escisión en la membrana. Modificado de (233, 234).

El péptido SSP, por su parte, es esencial como factor de maduración de GPC actuando en trans y se mantiene asociado de manera estable con GP por medio de un dominio que interactúa con un catión

zinc en la cola citoplasmática de G2 (235-237). Mientras que G1 y G2 son N-glicosilados en múltiples sitios (238), SSP adquiere un miristol en el residuo glicina en la posición 2 (239), similar a la proteína Z. SSP tiene un largo y vida media inusual (mayor a 6 horas) para un péptido señal (223, 240). La ausencia de miristoilación en SSP no afecta la formación del complejo trimérico de SSP/G1/G2 pero sí el proceso de fusión (239).

El péptido SSP cuenta con dos regiones hidrofóbicas separadas por un bucle hidrofílico con un residuo de cadena lateral cargado positivamente que está conservado en los mammarenavirus (K33), diferente al único dominio transmembrana del péptido señal estándar (240). En JUNV, SSP cuenta con sus extremos N y C terminales en el citosol y la lisina K33 en el lado extracelular (222). Esta lisina K33 es importante para la fusión dependiente de pH de la glicoproteína con la membrana celular del hospedante (234, 241).

La N-glicosilación es un proceso importante para el plegamiento correcto y el transporte intracelular de glicoproteínas (242). De hecho, se ha observado en mammarenavirus que N-glicanos son necesarios para la proteólisis y transporte a la superficie celular del complejo GP-C (238), y se requieren para la infectividad de los viriones (243). No todos los mammarenavirus poseen el mismo número de sitios de glicosilación. Por ejemplo, LCMV cuenta con 8 sitios de N-glicosilación (6 en G1 y 2 en G2) mientras que LASV tiene 11 (7 en G1 y 4 en G2) (238, 242). A su vez, no todos los sitios de glicosilación son igualmente importantes. Por ejemplo, en LASV la falta de glicosilación en los residuos 89 y 365 inhibe la infectividad mientras que la ausencia de glicanos en las posiciones 109 y 119 sólo inhiben parcialmente (244). Adicionalmente, se observó que la N-glicosilación de diferentes sitios permitía la evasión de la respuesta inmune de LASV como también modificaba la proporción de células T CD4+ y CD8+ detectadas ante inmunización de ADN en ratón (244). En LCMV, la glicosilación del sitio 173 inhibe la neutralización enmascarando el epítipo GP-1D (243). Se ha propuesto que las glicoproteínas de ciertos mammarenavirus confieren un escudo protector contra la detección y unión a anticuerpos (245).

Se ha establecido que existe una interacción entre JUNV GP y la proteína 53 kDa del compartimiento intermedio ER-Golgi (ERGIC-53) (246). ERGIC-53 es importante para la infectividad de los viriones de JUNV. Interactúa con el precursor GPC pero no con las subunidades procesadas G1 y G2, y es un receptor de transporte para el tráfico de la glicoproteína en estadios tempranos de la vía exocítica e incluso estaría presente en el virión infectivo (246).

4.1.1 Brotación y ensamblado viral

Como se describió en el capítulo 2, la proteína Z interactúa con otras proteínas virales para poder coordinar el ensamblado y brotación correcto de viriones infectivos. De hecho, se ha observado que Z interactúa con el complejo de glicoproteínas SSP/G1/G2 por medio de SSP (145). Existe evidencia que esta interacción difiere entre los mammarenavirus OW y NW. La miristoilación de Z parecería ser suficiente para el ensamblado de glicoproteínas en partículas brotantes de OW (145), mientras que se requiere un dominio RING intacto en Z para la incorporación de las glicoproteínas en los viriones NW (247).

4.1.2 Determinantes moleculares de atenuación

Existe variación en el procesamiento del precursor GPC entre cepas. Por ejemplo, GPC de la cepa C#1 posee un procesamiento menos eficiente que GPC de la cepa Romero, produciéndose una retención

en el ER. Este fenómeno se ha asociado a una mutación puntual (T168A) que afecta un motivo de N-glicosilación en G1 entre la cepa Romero y C#1 (248).

La atenuación en la cepa C#1, generada por medio de pasajes secuenciales fue explicada por la diferencia de aminoácidos en el sitio 427 de G2 (39). La sustitución ocurrida T168A entre los pasajes posteriores a XJ13 y anteriores a XJ44 elimina un motivo de glicosilación ligada a N, lo cual impactaría en la forma de la enfermedad predominante entre visceral y neurológica generada por JUNV (249).

4.2 Resultados

4.2.1 Secuencia

En la Tabla 11 se resumen las observaciones de los alineamientos comparativos para las secuencias nucleotídicas y aminoácidas de las cepas mencionadas en la introducción general con la cepa P3441.

Tabla 11. Porcentaje de identidad y número de nucleótidos y aminoácidos de diferencia para las secuencias de GPC de cepas de JUNV respecto a P3441.

	Nucleótidos		Aminoácidos	
	Similitud (%)	Diferencias (nº nts)	Similitud (%)	Diferencias (nº aa)
Candid 1	96	54	98	7
XJ13	97	39	99	3
P35032	95	61	98	7
MC2	93	87	98	6
Romero	96	42	99	4
Ledesma	95	53	99	4
Espindola	97	38	99	4

Se puede observar una diferencia apreciable entre las cepas de JUNV respecto a P3441 en la secuencia nucleotídica, especialmente con MC2. Sin embargo, fueron C#1 y P35032 quienes presentaron el mayor número de aminoácidos de diferencia.

Dentro de SSP sólo se observó un único sitio de diferencia entre las cepas. Mientras P3441 comparte una valina con C#1 y MC2, las cepas más virulentas presentan una isoleucina (Figura 4.2). Considerando su cercanía a K33, se incluyó este sitio V35 como residuo de interés para análisis posteriores.

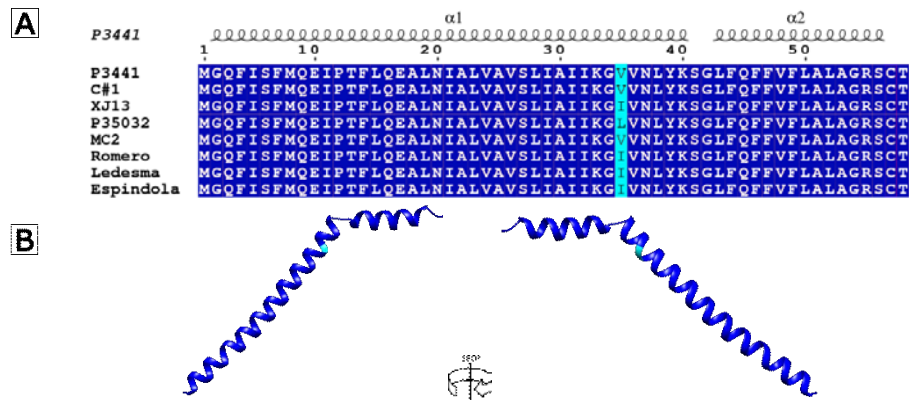


Figura 4.2. (A) Alineamiento de secuencias de aminoácidos para la proteína SSP de la cepa P3441, C#1 y otras cepas patogénicas de JUNV. Las equivalencias se encuentran marcadas en negro, y la estructura secundaria de la proteína puede observarse por encima del alineamiento, con α indicando alfa hélices, β lámina beta, η hélices 3_{10} , y T para giro beta. (B) Estructura tridimensional de la proteína JUNV SSP obtenido por ColabFold (ver materiales y métodos). Cada imagen corresponde a lados opuestos de la proteína. En cian se muestra el residuo de interés en la posición 35.

La mayoría de las diferencias entre las cepas respecto a P3441 radican en la proteína G1. Como se puede ver la Figura 4.3, estos sitios se encuentran distribuidos entre los residuos de la posición 50 y 165, en algunos casos encontrándose próximos con uno a tres aminoácidos conservados de por medio. Esta región ha sido predicha de poseer una estructura terciaria compuesta de hojas plegadas beta y alfa hélices. De las cuatro diferencias entre P3441 y C#1, marcados en cian en la Figura 4.3, se identificaron tres en sitios sin estructura secundaria y el cuarto en una hélice alfa.

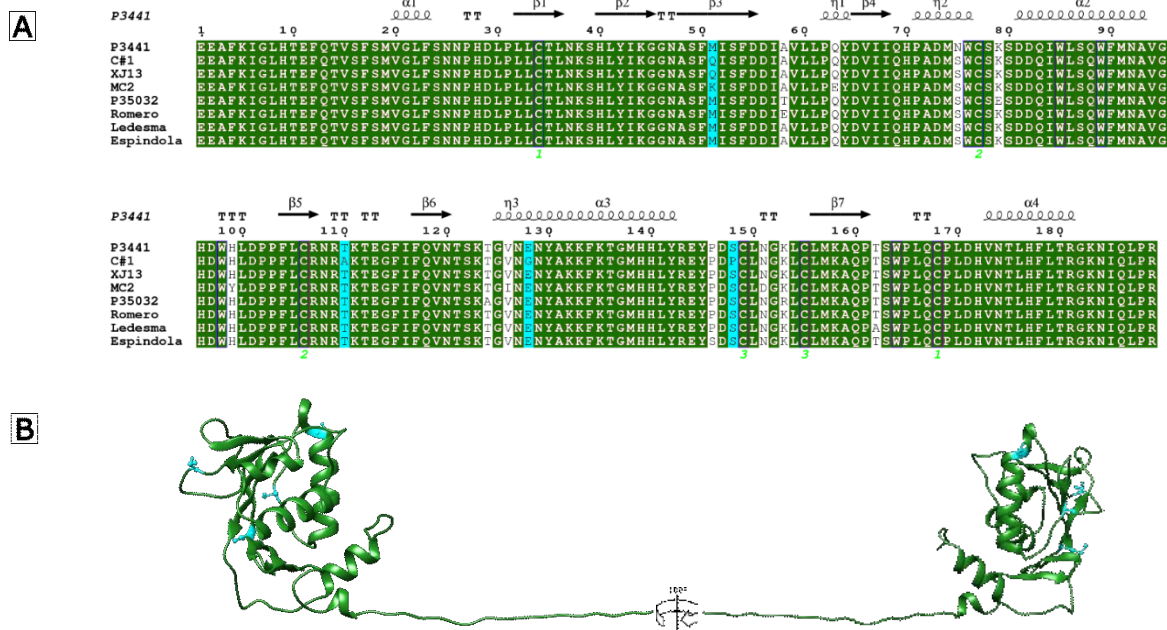


Figura 4.3. (A) Alineamiento de secuencias de aminoácidos para la glicoproteína G1 de la cepa P3441, C#1 y otras cepas patogénicas de JUNV. Las equivalencias se encuentran marcadas en negro, y la estructura secundaria de la proteína puede observarse por encima del alineamiento, con α indicando alfa hélices, β lámina beta, η hélices 3_{10} , y T para giro beta. (B) Estructura tridimensional de la proteína JUNV G1 obtenido por ColabFold (ver materiales y métodos). Cada imagen corresponde a lados opuestos de la proteína. En cian se muestran los residuos de interés en la posición 51, 110, 128 y 148.

A diferencia de G1, en G2 sólo se observaron tres diferencias respecto a P3441, de los cuales dos fueron únicas para C#1 respecto a las otras cepas (Figura 4.4).

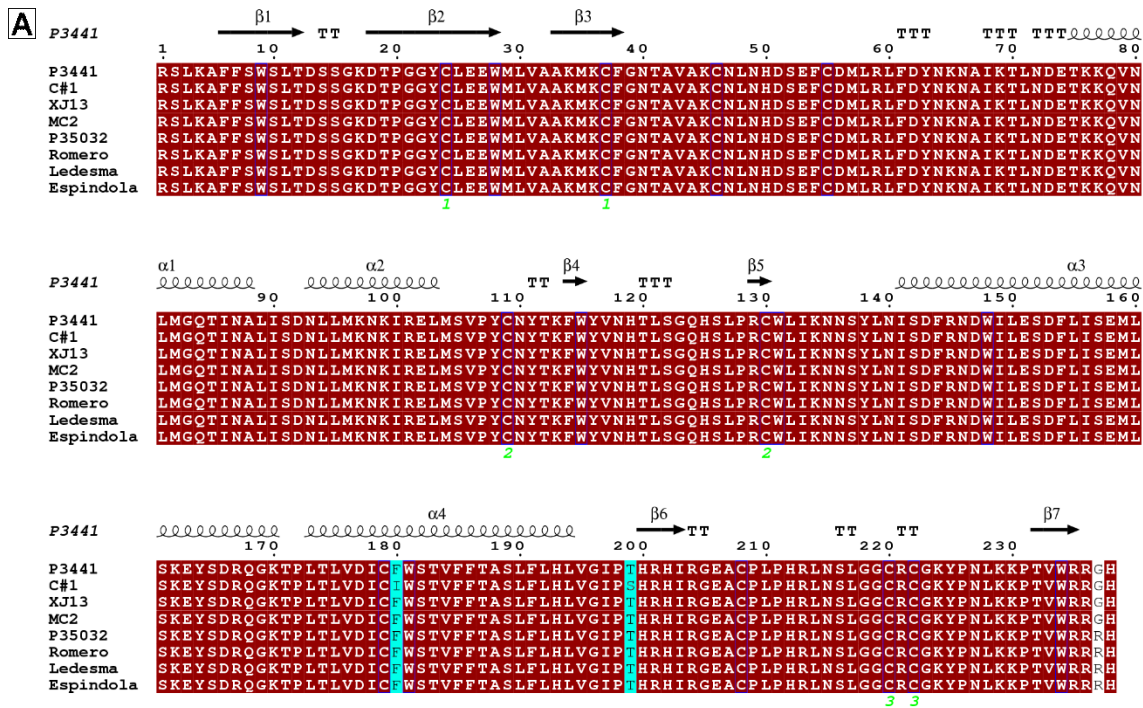


Figura 4.4. (A) Alineamiento de secuencias de aminoácidos para la glicoproteína G2 de la cepa P3441, C#1 y otras cepas patogénicas de JUNV. Las equivalencias se encuentran marcadas en negro, la estructura secundaria de la proteína puede observarse por encima del alineamiento, con α indicando alfa hélices, β lámina beta, η hélices 3_{10} , y T para giro beta. (B) Estructura tridimensional de la proteína JUNV G2 obtenido por ColabFold (ver materiales y métodos). Cada imagen corresponde a lados opuestos de la proteína. En cian se muestran los residuos de interés en la posición 180 y 199.

Para analizar el impacto de estas diferencias en la estructura de los diferentes componentes, se realizó un primer análisis usando el servidor MAESTROweb para una primera serie de predicciones. Los resultados se pueden ver en la Tabla 12 y en todos los casos contaron con una confianza apreciable ($C_{pred.} > 0,8$).

Los $\Delta\Delta G_{pred}$ para las diferencias encontradas en las secuencias de SSP entre P3441 y las otras cepas sugieren que P3441 cuenta con un SSP con mayor inestabilidad alrededor del residuo 35.

En el caso de G1, las predicciones de $\Delta\Delta G$ sugieren que C#1 es estructuralmente más inestable que P3341 y las otras cepas patogénicas. Los $\Delta\Delta G_{pred}$ fueron intermedios o elevados ($> 0,5$), prediciéndose

la mayor inestabilidad en la diferencia del residuo 51. Se incluyó para G1 la diferencia en el residuo 75, la cual es única para P3441 respecto a todas las demás cepas. Las mutaciones combinadas de G1 dieron un $\Delta\Delta G_{pred} = 2,416$, y $C_{pred} = 0,763$.

Finalmente, para G2 las dos diferencias de C#1 con las otras cepas confieren menor estabilidad. F180I tuvo un valor para el $\Delta\Delta G_{pred}$ más elevado que T199S. De hecho, mientras que T199S fue menor inclusive para las predicciones de G1, F180I tuvo un $\Delta\Delta G_{pred}$ similar a la observada para el residuo 51 de G1. La mutación combinada de G2 dio $\Delta\Delta G = 1,584$ y $C_{pred} = 0,738$.

Tabla 12. Cambio de estabilidad ($\Delta\Delta G_{pred}$) calculada como mutación independiente de un residuo en la estructura para las proteínas SSP, G1 y G2 de P3441 a un residuo en otra cepa (X).

Péptido	Mutación	$\Delta\Delta G_{pred}$	C_{pred}
SSP	V35I (X = XJ13, MC2, Romero, Ledesma, Espindola)	-0,648	0,932
	V35L (X = P35032)	-0,743	0,867
G1	M51Q (X = C#1, XJ13)	1,629	0,867
	N75S (X = toda otra cepa)	-0,304	0,939
	T110A (X = C#1)	0,430	0,876
	E128G (X = C#1)	0,532	0,928
	S148P (X = C#1)	0,644	0,889
G2	F180I (X = C#1)	1,568	0,851
	T199S (X = C#1)	0,164	0,927

Valores de $\Delta\Delta G_{pred} > 0$ se consideran desestabilizadores. C_{pred} Es un valor de confianza de la predicción, siendo $C_{pred} = 0$ de mínima confianza e $C_{pred} = 1$ máxima confianza de predicción.

Para profundizar en las diferencias observadas, se utilizó el servidor DynaMut para incorporar un análisis de dinámica molecular en el estudio del impacto estructural. Los resultados se muestran en la Tabla 13.

Tabla 13. Efecto en la estabilidad ($\Delta\Delta G_{pred}$ DynaMut) y entropía vibracional ($\Delta\Delta S_{vib}$) calculada para un cambio de un residuo en la estructura de SSP, G1 o G2 de P3441 a un residuo en otra cepa (X).

Péptido	Mutación (P > X)	$\Delta\Delta G_{pred}$ Dynamut	$\Delta\Delta S_{vib}$
SSP	V35I (X = XJ13, MC2, Romero, Ledesma, Espindola)	-0,061	0,198
	V35L (X = P35032)	-0,053	0,178
G1	M51Q	-0,169	0,212
	N75S	0,331	0,048
	T110A	0,092	0,083
	E128G	-0,333	0,335
	S148P	0,837	-0,219
G2	F180I (X = C#1)	0,974	-0,258
	T199S (X = C#1)	-0,613	0,149

$\Delta\Delta G_{pred}$ corresponde a la diferencia de energía libre entre el modelo salvaje y el mutante calculado por el servidor DynaMut, el cual define valores negativos (< 0) como mutación desestabilizadora. $\Delta\Delta S_{vib}$ es la diferencia de energía de entropía vibracional entre el modelo salvaje y el mutante, con valores positivos (> 0) para mutaciones en las cuales aumenta la flexibilidad molecular.

A diferencia de lo obtenido con el servidor MAESTROweb, se predijo que las diferencias entre las cepas para SSP fueron desestabilizadoras. No obstante, se calculó un valor muy reducido de $\Delta\Delta G_{pred}$ que sugiere un bajo impacto en la estructura de SSP. Para ambas diferencias estudiadas, el $\Delta\Delta S_{vib}$ predicho indica que en estos sitios SSP tiene menos estabilidad en P3441 respecto a otras cepas, ya

sean patogénicas como Romero y Ledesma, o atenuada como C#1. Como se puede ver en la Figura 4.5, la mayoría de los enlaces se mantienen y las cadenas laterales parecen adoptar el mismo espacio. Con isoleucina o leucina se pierde la interacción entre un grupo metilo de la cadena lateral (como en el caso de la valina) con un oxígeno de la cadena lateral de un residuo vecino (Figura 4.5, enlace rojo), lo que explicaría por qué el servidor DynaMut predijo mayor inestabilidad cuando la valina es reemplazada por esos aminoácidos. No obstante, esa interacción es con una cadena lateral sin polaridad (grupo metilo de la valina), lo que implica una menor fuerza de enlace comparada con interacciones del tipo Van der Waals o puente de hidrógeno, explicando el bajo $\Delta\Delta G$ predicho.

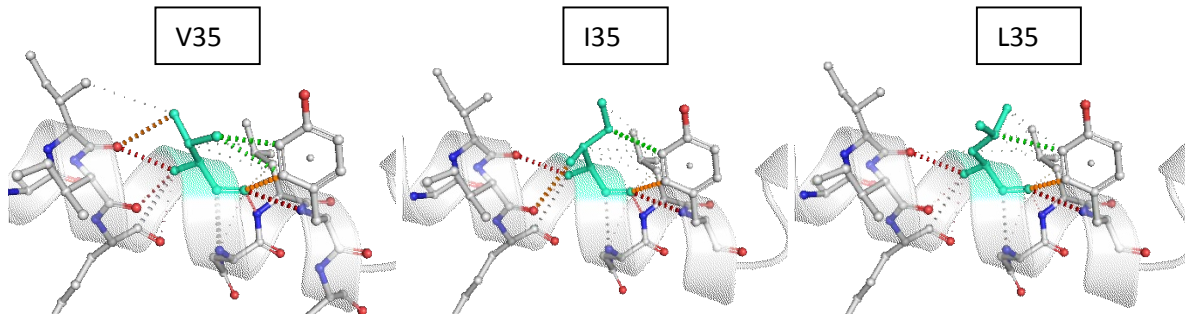


Figura 4.5. Esquema de estructura e interacciones entre cadenas laterales para el residuo 35 en JUNV SSP. A la derecha se encuentran las estructuras para C#1. Los residuos valina (izquierda), isoleucina (centro) y leucina (derecha) se encuentran en verde. En gris se ven los carbonos (C), en rojo los oxígenos (O) y en azul los nitrógenos (N). Las interacciones con otros residuos se ven como líneas punteadas.

El servidor DynaMut mantiene que M51Q y E128 son mutaciones desestabilizadoras de la estructura, pero la magnitud del cambio M51Q es mucho menor a la predicha previamente. En ambos casos hay un aumento en la flexibilidad, que es producto de esta desestabilización. Adicionalmente, con el servidor DynaMut se predijo que S148P y T110A son estabilizadoras, aunque en el primer caso la flexibilidad estructural se reduce cuando se reemplaza S148 con una prolina, probablemente debido a la rigidez que impone la prolina.

Para el caso de M51Q, la metionina y la glutamina cuentan dos cadenas laterales químicamente diferentes, con M teniendo un metil azufre terminal mientras que Q presenta un grupo amida. El tipo de interacciones que el azufre de M puede generar con aminoácidos cercanos no se repiten con la cadena lateral de Q (Figura 4.6, líneas amarillas), lo que explica la desestabilización cuando se da la mutación.

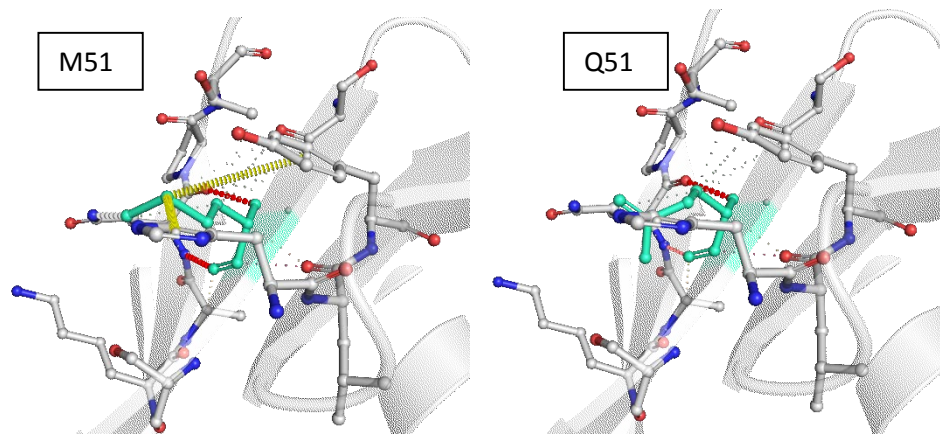


Figura 4.6. Esquema de estructura e interacciones entre cadenas laterales para el residuo 51 en JUNV G1. A la derecha se encuentran las estructuras para C#1. Los residuos metionina (izquierda) y glutamina (derecha) se encuentran en verde. En gris se ven los carbonos (C), en rojo los oxígenos (O) y en azul los nitrógenos (N). Las interacciones con otros residuos se ven como líneas punteadas.

Similarmente, en el caso de E128G (Figura 4.7), el ácido glutámico (E) es uno de los aminoácidos con carga negativa en pH fisiológico, mientras que la glicina (G) no tiene cadena lateral. No obstante, no parece que el ácido glutámico forme enlaces con cadenas laterales cercanas de G1. Probablemente la inestabilidad predicha para E128G se da por la mayor libertad de movimiento adquirida con la glicina, lo cual también condice con la mayor flexibilidad predicha para este cambio.

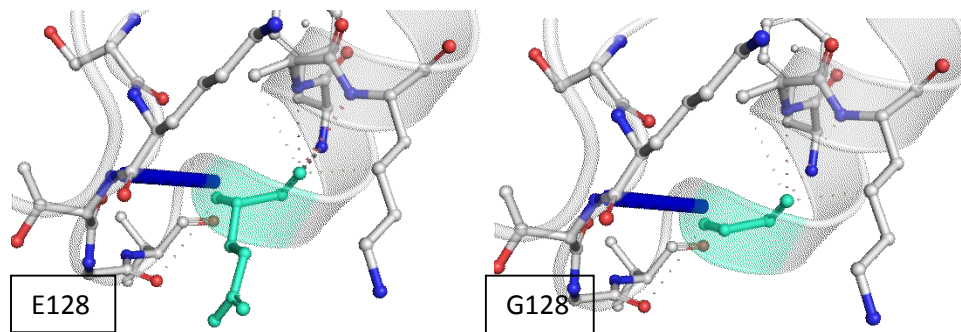


Figura 4.7. Esquema de estructura e interacciones entre cadenas laterales para el residuo 128 en JUNV G1. A la derecha se encuentran las estructuras para C#1. Los residuos ácido glutámico (izquierda) y glicina (derecha) se encuentran en verde. En gris se ven los carbonos (C), en rojo los oxígenos (O) y en azul los nitrógenos (N). Las interacciones con otros residuos se ven como líneas punteadas.

La diferencia N128S es única para P3441 (Figura 4.8) y ocurre entre una asparagina (N) que cuenta con una cadena más larga y un grupo amida respecto a la serina (S) en las otras cepas. Aunque la asparagina parece formar más interacciones con los aminoácidos cercanos, la cadena lateral de S de menor tamaño parece favorecer el movimiento de las cadenas laterales circundantes.

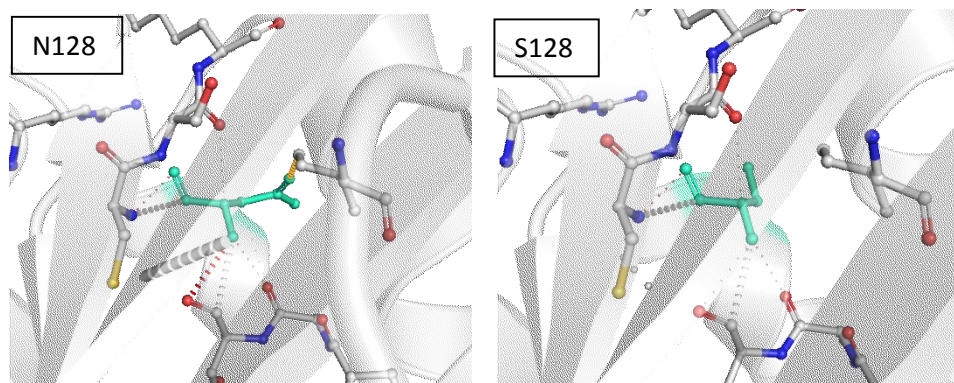


Figura 4.8. Esquema de estructura e interacciones entre cadenas laterales para el residuo 128 en JUNV G1. A la derecha se encuentran las estructuras para C#1. Los residuos asparagina (izquierda) y serina (derecha) se encuentran en verde. En gris se ven los carbonos (C), en rojo los oxígenos (O), en amarillo los azufre y en azul los nitrógenos (N). Las interacciones con otros residuos se ven como líneas punteadas.

Con respecto a las diferencias observadas dentro de G2, el servidor predijo que la posición 180 en C#1 se encuentra estabilizada y menos flexible respecto a P3441. Esto no sorprende debido a que el anillo bencílico de la cadena lateral de la fenilalanina de P3441 es más voluminoso e hidrofóbico respecto a la isoleucina. En la Figura 4.9 se puede ver que este anillo queda orientado restringido hacia dentro de la hélice alfa, mientras que la isoleucina en la misma posición puede extenderse e interactuar con el residuo del giro siguiente, lo cual explica la predicción de estabilización ante un aumento en la energía de enlaces en la posición. No obstante, es posible que la fenilalanina de P3441 y cepas patogénicas tenga un rol en la interacción con otras proteínas o que se requiere de flexibilidad en el sitio tal que se explique la atenuación producida en C#1 ante esta mutación.

Para el sitio 199, por el contrario, se predijo una desestabilización en la mutación S199T. Curiosamente, el residuo serina es químicamente similar a la treonina, ambos contando grupos hidroxilos en sus cadenas laterales que le confieren hidrofobicidad. La diferencia entre los dos radica en un grupo metilo carente en la cadena lateral de la treonina respecto a la serina. Esto desfavorece la capacidad de acercarse a residuos cercanos para formar enlaces, como se visualiza en la Figura 4.9, y por ende explica por qué las predicciones del servidor sugieren una desestabilización en el sitio.

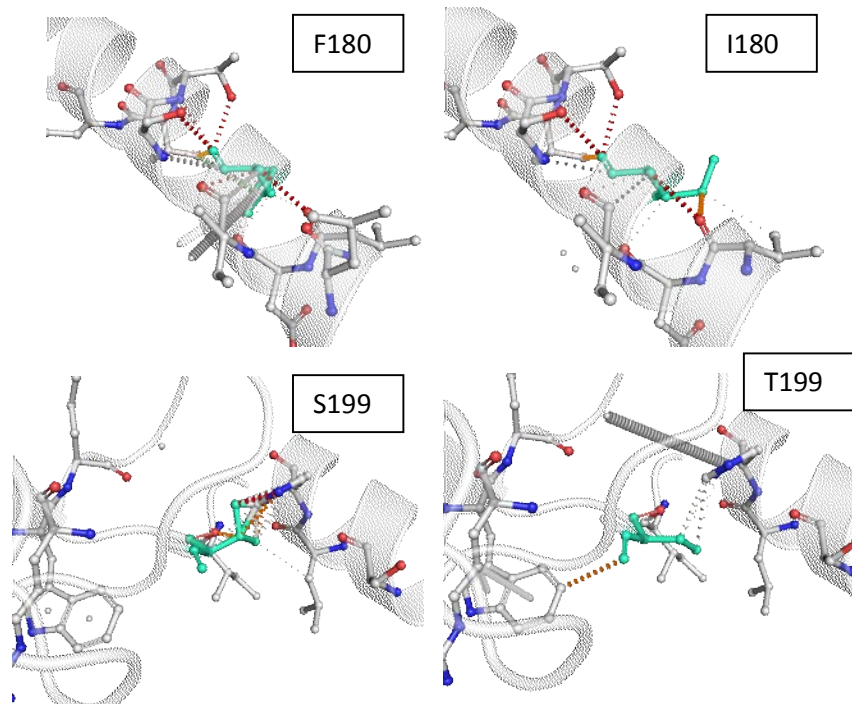


Figura 4.9. Esquema de estructura e interacciones entre cadenas laterales para el residuo 180 y 199 en JUNV G2. A la derecha se encuentran las estructuras para C#1. Los residuos fenilalanina (izquierda, arriba), isoleucina (derecha, arriba), serina (izquierda, abajo) y treonina (derecha, abajo) se encuentran en verde. En gris se ven los carbonos (C), en rojo los oxígenos (O), en amarillo los azufre y en azul los nitrógenos (N). Las interacciones con otros residuos se ven como líneas punteadas.

4.2.2 Generación de vectores de expresión eucariotas para GPC

Una vez clonada la proteína en un vector de entrada (*pGEM-T Easy Vector*), se buscó subclonarla dentro de un vector de expresión bajo promotor eucariota. En este caso se incluyó el codón de terminación para favorecer la formación de las espículas lo más semejante a la proteína natural y sin interferir en el procesamiento postraduccional de GPC. Es decir, el marco de lectura clonado se transcribe y traduce sin un fluoróforo C terminal.

Después de varios intentos fallidos de clonar el marco de lectura de GPC en *pEGFP-N3*, se contrató el servicio de clonado de *WatsonBio Sciences* (EEUU). Los plásmidos *pEGFP-N3* y *pGEM-GPC* se enviaron a la empresa y se utilizó la enzima *EcoRI* para el clonado. La empresa aseguró la orientación correcta del marco de lectura por secuenciación, y fue verificado posteriormente por digestión con enzimas de restricción y ensayos de PCR en el laboratorio.

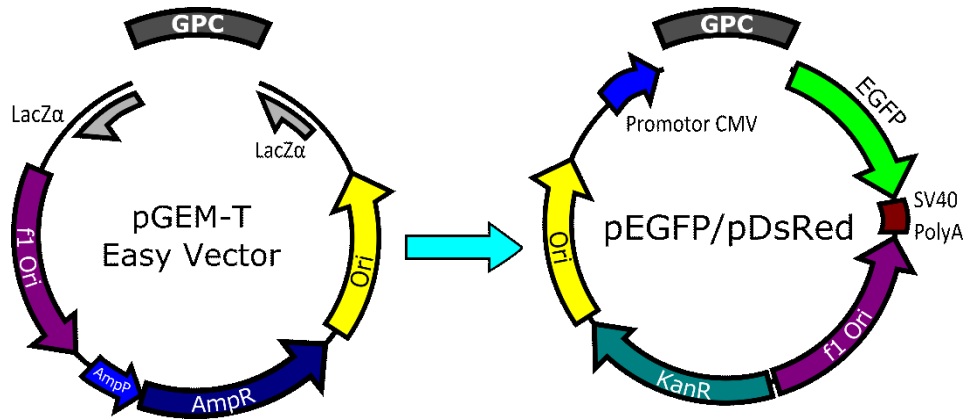


Figura 4.10. Generación del plásmido pEGFP-GPC. Esquema del clonado del marco de lectura para la proteína GPC de JUNV. Se partió de un vector que permitía el clonado directo de fragmentos amplificados de PCR para poder realizar la posterior digestión y ligación en el vector de expresión eucariota.

El subclonado fue exitoso y se realizaron transfecciones en células A549. Se observó una distribución uniforme dentro de las células transfectadas con el vector pEGFP vacío (Figura 4.11) y una distribución puntiforme para las células transfectadas con el vector pEGFP-GPC, revelada por la marcación con un anticuerpo monoclonal específico para GPC (Figura 4.11). En este ensayo se demuestra que el producto del plásmido pEGFP-GPC de P mantiene su antigenicidad por ser reconocido por anticuerpos específicos de GPC nativa.

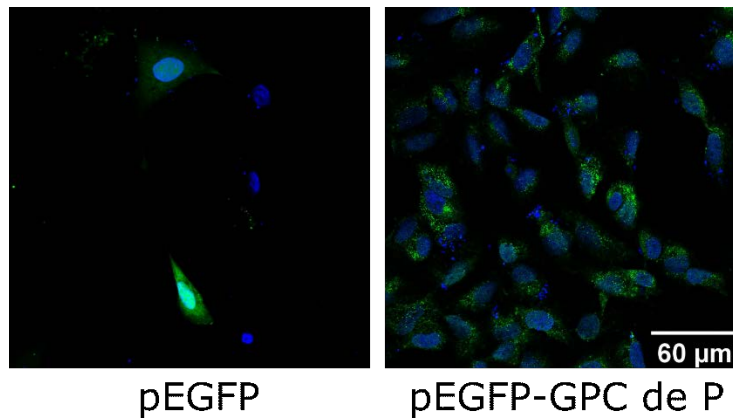


Figura 4.11. Imágenes de fluorescencia por microscopía confocal de células A549 transfectadas con el plásmido pEGFP o con pEGFP-GPC de P3441 (P). En azul se observan los núcleos celulares por medio de tinción de ADN doble cadena con verde metilo. En verde se encuentra la fluorescencia correspondiente a FITC proveniente de inmunofluorescencia usando anticuerpos monoclonales contra GPC.

No se cuenta con un plásmido equivalente para GPC de C#1. No obstante, se han iniciado los trámites para su obtención de otros laboratorios.

4.2.3. Localización e interacciones

El precursor de glicoproteína GPC se procesa postraduccionalmente y se conoce su transporte dentro del complejo de Golgi. Se realizaron ensayos de inmunofluorescencia en células A549 infectadas para determinar la localización de las GPC respectivas, detectadas con anticuerpos monoclonales

específicos contra GPC, dentro del complejo de Golgi, teñido con un anticuerpo contra la proteína GM130 (Figura 4.12).

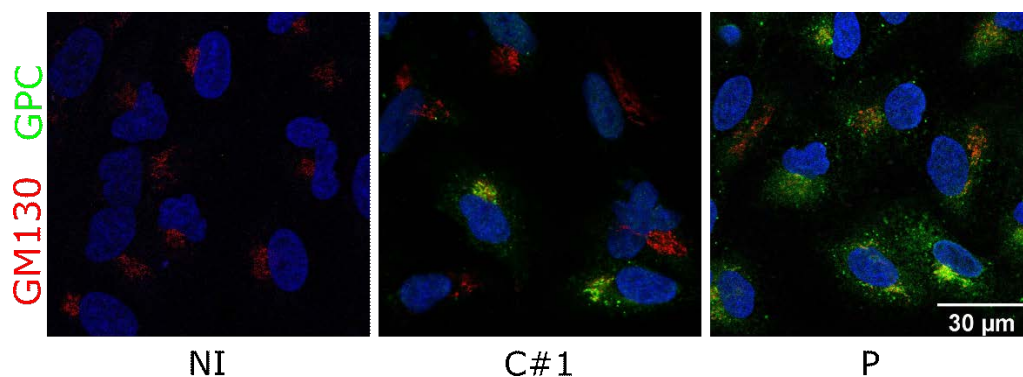


Figura 4.12. Imágenes de microscopía confocal de células A549 no infectadas o infectadas con C#1 y P3441 (P) a 2 dpi con marcación de GM130 y GPC. En azul se observan los núcleos celulares por medio de tinción de ADN doble cadena con verde metilo. En verde se observa usando anticuerpos monoclonales contra GPC y anticuerpo secundario anti especie unido con a FITC. El color rojo corresponde a la marcación con anticuerpos anti GM130 seguida de secundario anti especie marcado con ficoeritrina.

4.3 Discusión

Por primera vez se pudo obtener la secuencia del gen GPC de la cepa P3441 de JUNV. Esta secuencia fue comparada con otras cepas y se identificaron importantes diferencias. La mayoría de ellas radican en G1, y sólo tres diferencias en G2. Esto podría significar que G1 presenta menor presión de selección respecto a G2 o, en otras palabras, mayor tolerancia a mutaciones que no afecten la viabilidad del virus. Considerando que G1 está involucrada en el reconocimiento del receptor celular, es posible que esta mayor tolerancia a mutaciones sea evolutivamente favorable para JUNV al permitir la evasión de anticuerpos y/o a infectar nuevas especies animales. Este fenómeno se ha visto en otros virus, dónde la proteína involucrada en el reconocimiento del hospedante presenta mayor número de mutaciones acumuladas respecto a otras proteínas virales, como el SARS-CoV-1 (250). Entonces, es probable que diferencias en los fenotipos de infección entre las cepas patogénicas de JUNV radiquen en su mayoría en esta glicoproteína, que podría definir diferencias en la eficiencia de unión y entrada a células hospedantes como diferencias en el tropismo a diferentes células diana. Por su parte, G2 se encuentra más restringida debido a su rol en la fusión de membranas para la liberación del genoma viral en el citoplasma, y por ende es de esperarse menor tolerancia a mutaciones. De hecho, se encuentra reportado que el residuo I180 de C#1 produce la mayor atenuación en la cepa vacunal (251). Respecto a la mutación encontrada en SSP, es difícil determinar si esto impacta en la función de SSP debido a que la rigidez estructural no es necesariamente contraproducente en la función de una proteína, en particular siendo SSP utilizado como péptido señal, mientras que la diferencia N75S haría más estable a G1 de P3441 comparada con todas las variantes de JUNV estudiadas.

Acorde a nuestro análisis, los residuos 51 y 128 de G1 podrían ser sitios que hacen estructuralmente más inestable esta glicoproteína en C#1 con respecto a las cepas patogénicas. Por otro lado, resultó llamativo la estabilidad predicha para I180 en C#1 respecto a F180 de P3441 para G2, lo que sugiere que la atenuación producida por este residuo puede ser debido al impacto en la interacción con

componentes celulares. De hecho, la baja tasa de diferencias entre las cepas en esta glicoproteína sugiere que es susceptible a perder funcionalidad ante cambios, comparada con G1 donde radican la mayoría de las diferencias entre los GPC de las cepas de JUNV estudiadas.

El marco de lectura de GPC de P3441 fue clonado dentro de un vector de expresión eucariota y expresado en células A549 exitosamente. Además, se demostró que la secuencia traducida mantiene la vía de procesamiento esperado. Por último, la proteína recombinante fue reconocida por un anticuerpo monoclonal demostrando que es antigénicamente similar a la proteína natural. Este plásmido servirá en proyectos futuros para profundizar estudios ya iniciados sobre la respuesta del aparato de Golgi en las infecciones por C#1 y P3441.

Capítulo 5. Proteína L

5.1 Introducción

La ARN polimerasa ARN dependiente L de los mammarenavirus oscila entre 2080 y 2241 aminoácidos. En el caso de JUNV, esta proteína posee 2210 aminoácidos y un peso molecular de 240 kDa.

5.1.1 Estructura de la proteína L

La proteína L es la más grande de las proteínas de JUNV. Cuenta con cuatro dominios putativos con diferentes funciones identificadas por análisis comparativos de secuencias aminoacídicas entre varias especies de virus con genoma ARN (252-254). Los cuatro dominios, conservados en los mammarenavirus, se encuentran entre las posiciones 1 a 300, 500 a 900, 1000 a 1600 y 1700 a 2000 (255). Los dominios N terminales son conservados únicamente en las proteínas L de arenavirus y bunyavirus (256, 257). El tercer dominio exhibe ocho motivos (A-H) centrados alrededor de aminoácidos invariantes de todas las ARN polimerasas ARN dependientes, seis de los cuales (A-F) comparten homología con otras polimerasas ya caracterizadas (256-258) y dos (G y H) están conservados únicamente en las polimerasas de virus con genoma ARN de polaridad negativa (258, 259). El cuarto dominio C terminal es conservado únicamente en mammarenavirus (260) y aunque no es necesario para la interacción con la proteína Z, es importante para la actividad polimerasa de la proteína (253).

El dominio I tiene actividad endonucleasa (261), mientras que el dominio central de la proteína L (en LASV los residuos 1040 a 1540) presenta la actividad ARN polimerasa ARN dependiente (262, 263). Mutaciones puntuales dentro de esta región central resultaron en una actividad defectuosa para la síntesis de ARN mensajero pero eficiente para ARN antígenómico (263). La parte N terminal de L está involucrada en la síntesis de ARN mensajero (264) (Figura 5.1).

Varios estudios de cristalografía por rayos X y microscopía crioelectrónica han demostrado que la estructura de L se asemeja al complejo ARN polimerasa del virus influenza, permitiendo dividir L entre tres dominios estructurales: uno similar a la proteína polimerasa ácida (PA) de influenza, otro dominio central con actividad ARN polimerasa y finalmente un dominio C terminal similar a la proteína polimerasa básica 2 (PB2) del complejo polimerasa de influenza (265). De hecho, en LASV la proteína L se pudo separar en tres partes manteniendo su funcionalidad (254). Cuando se dividió L de LASV en los residuos G467 o G939, se formaron dos partes (N y C terminal) que se complementaban en *trans* y se reconstituían en un complejo funcional, interactuando entre sí independientemente del sitio en cual se había partido la proteína (254). Adicionalmente, en estos mismos ensayos, se observó que la parte N terminal conformaba una distribución puntual en el citoplasma, a veces perinuclear, semejante a la distribución de la proteína L de LASV completa, mientras que el dominio C terminal se distribuía homogéneamente por el citoplasma (254).

La proteína L de JUNV es catalíticamente activa en presencia de ARN viral de polaridad negativa y es activada en presencia de cationes magnesio o manganeso, pero no de zinc (265).

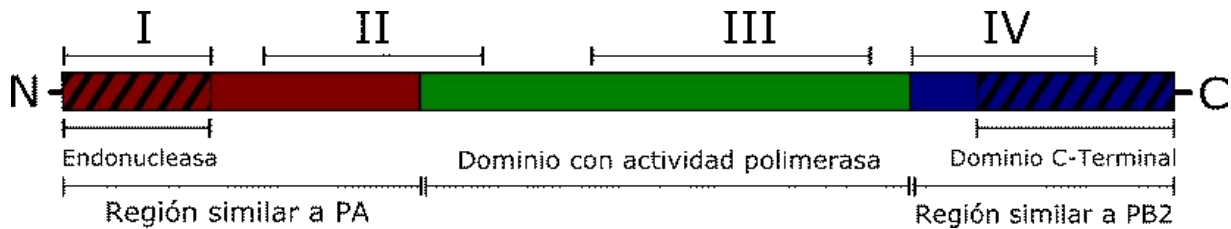


Figura 5.1. Diagrama representativo de los dominios de la polimerasa L. Las regiones conservadas I a IV para mammarenavirus están indicadas en la parte superior del diagrama como fue propuesto (260). La estructura de L posee similitudes con el complejo polimerasa del virus influenza, contando con regiones homólogas. El dominio N terminal (rojo) se asemeja estructuralmente a PA del virus influenza, contando con una secuencia asociada a actividad endonucleasa. El dominio central (verde) cuenta con actividad y estructura de polimerasa semejante al del virus influenza. El dominio que se parece a PB2 de influenza (azul), presenta un dominio C terminal distintivo de la proteína L.

Dos sitios de unión a Z fueron localizados entre los residuos 160 y 290 en LASV y 156 y 292 para TACV, superponiéndose en el sitio similar a dominio endonucleasa y otro en el dominio asociado a actividad polimerasa (253). Ensayos de microscopía crioelectrónica permitieron identificar en L de MACV el sitio de unión a Z en la intersección entre las regiones N y C terminales con la región central (265, 266). Esta unión ocurre tanto cuando la proteína L se encuentra asociada con ARN o libre (147). Vale destacar que las proteínas Z de LASV y MACV pueden inhibir la proteína L del otro virus, aunque con menor eficiencia (267). Semejante a otras polimerasas de virus ARN de polaridad negativa, la proteína L se oligomeriza para cumplir su función en el ciclo de replicación viral (268).

5.1.2 Replicación viral

Cuando se libera al citoplasma, el genoma viral se encuentra encapsulado por la proteína N y unido a proteínas L en sus extremos para formar una ribonucleoproteína pseudo circularizada (200, 269, 270). Para lograr replicar y transcribir, la proteína L necesita remover las subunidades de N además de guiar la encapsidación de moléculas de ARN nacientes. La síntesis de ARNm viral comienza con el proceso de remoción de extremos 5' m⁷ GpppN o CAP (*CAP-snatching*) del ARNm del hospedante, proceso mediado por la misma proteína L (261, 271-274). El CAP adquirido sirve de cebador de la transcripción (261, 271-274), mientras que la replicación se ha propuesto que comienza con un pppG inicial que explicaría la presencia de un nucleótido guanidina ocasional en los extremos 5' de los genomas en ciertas réplicas de secuenciación (275). El ARNm naciente no se une a la proteína N y termina cuando se alcanzan las regiones intergénicas del genoma viral, cuya alta estructura detiene el proceso de transcripción (111, 276, 277). La unión de N a estas regiones intergénicas suprime la terminación de L en esta región durante la síntesis de ARN (118, 277) y de esta manera permite la replicación del genoma (185).

La proteína L tiene la actividad polimerasa necesaria tanto para la transcripción de marcos de lectura virales como para la replicación genómica. Esta última función requiere de la proteína N y es inhibida por la proteína Z (114), tanto para mammarenavirus NW como OW (114, 148, 198, 278). La inhibición de Z ocurre por la unión directa con L, dejando el complejo polimerasa arrestado en un estado inactivo (279).

La interacción entre N y L se da por medio de intermediarios de ARN (280-282), requiriendo la presencia de secuencias de ARN específicas del virus dentro de las regiones no traducibles 5' y 3' del

genoma viral (281). Los complejos L-N pueden adoptar estructuras monoméricas, homotriméricas en forma de anillo e inclusive órdenes mayores (197). Las proteínas N de ciertos arenavirus son capaces de funcionar junto a L de otros (282). Además, en LCMV se observó que un residuo (I53) y la región C terminal de N son esenciales para la función enzimática de L (282).

5.1.3 Virulencia

Estudios filogenéticos han sugerido que la proteína L es el blanco de preferencia en la selección evolutiva tanto en OW (283) como en NW (284), proponiendo a L como el principal determinante de adaptación. Una hipótesis es que L modula la virulencia al contribuir en el establecimiento de nuevos hospedantes. Los mammarenavirus pueden infectar a varias especies animales por fuera de sus reservorios naturales (285), ya sea en otros roedores distantes a sus hospedantes naturales como a primates (286). Con frecuencia, estas infecciones resultan en patología severa o letal (286). Esto sugiere una tolerancia optimizada de parte de los hospedantes naturales que permite el mantenimiento del virus en dichas poblaciones animales. Cambios en la polimerasa podrían ser requeridos para refinar la persistencia viral y/o virulencia después de saltos de hospedante en la evolución de los mammarenavirus. En OW, cambios en la proteína L fueron asociados a diferencias en la eficiencia de replicación y fenotipos de enfermedad en roedores (287-289). Por ejemplo, una replicación viral acelerada en estadios tempranos de la infección demostró ser central en el establecimiento de infecciones persistentes para algunas cepas de LCMV (288, 290). En NW, tres sustituciones en la secuencia aminoacídica de L de PICV transformaron una cepa que producía una enfermedad febril en cobayos (cepa P2), en una cepa letal (P18) (291). La diferencia en el fenotipo de la enfermedad fue asociada a una mayor replicación viral (291). Por otro lado, se ha observado que L de MOPV puede activar la vía de señalización de RLR/MAVS, posiblemente debido a la síntesis de ARN pequeños (292). En este sentido, al menos en OW, la proteína L interactúa con componentes de la respuesta inmune innata. Por ejemplo, en LCMV se determinó la interacción con TRIM21, DDX3X y NKRF, y la replicación de este virus era mayor en ratones *trim21*^{-/-} (293). Aparentemente, la proteína L también interactúa con otros componentes del hospedante que podrían servir de blancos para tratamientos farmacológicos, como el caso de eRF3a/GSPT1 en LASV (294).

5.2 Resultados

5.2.1 Secuencia

En la Tabla 14 se resumen las observaciones de los alineamientos comparativos para las secuencias nucleotídicas y aminoacídicas de la proteína L de las cepas mencionadas en la Introducción general con la cepa P3441.

Tabla 14. Porcentaje de identidad y número de nucleótidos y aminoácidos de diferencia para diferentes cepas de JUNV respecto a P3441.

	Nucleótidos		Aminoácidos	
	Similitud (%)	Diferencias (nº nts)	Similitud (%)	Diferencias (nº aa)
Candid 1	96	208	97	46
XJ13	97	192	98	40
P35032	95	283	97	47
MC2	97	194	97	49
Romero	97	193	97	48
Ledesma	95	279	97	59
Espíndola	97	194	98	40

MC2 y Romero, que contaban con numerosas diferencias en proteínas anteriores, difirieron en una única posición entre sí (residuo 775, Figura 5.2A). C#1 contó con 7 diferencias únicas respecto a las otras cepas, 8 diferencias respecto a otras cepas que compartía con XJ13 y 8 que compartía con XJ13 y Ledesma comparada a las otras cepas. En cambio, P3441 muestra 20 diferencias únicas respecto a las otras cepas. Se determinaron los sitios dónde hubiera diferencias entre las cepas C#1 y/o XJ13 respecto a otras cepas patogénicas y fueron marcados en cian en la Figura 5.2 A y B. De estos sitios, 4 se encuentran en la región N terminal, 5 en la región central y 5 en la región C terminal.

Luego se analizó como las mutaciones en P3441 podrían impactar en la estabilidad estructural de la proteína. Para ello, primero se generaron estructuras tridimensionales, para después verificar, refinar y elegir según los criterios utilizados en el análisis de la estructura de Z en el capítulo 2.

El modelo que obtuvo la mejor puntuación de calidad (*scores* definidos de los diferentes servidores web utilizados), se utilizó para un análisis de los sitios de interés. Con el servidor MAESTROweb se determinó la energía de estabilización final para cada sitio de interés, calculándose como mutación individual del residuo en P3441 a aquella correspondiente a C#1. Los resultados se muestran en la Tabla 15. En todos los casos, las mutaciones fueron predichas como estabilizadoras de la estructura. Esto fue inesperado considerando que se realizó el análisis de una estructura de una cepa patogénica (P3441) a una atenuada, y que se esperaría una estructura más inestable. El cálculo combinado $\Delta\Delta G_{pred}$ fue de -0,318 (estabilizador), con una confianza de predicción (C_{pred}) de 0,865.

Con el fin de explorar el efecto de estas diferencias y confirmar lo obtenido con el servidor MAESTROweb, se procedió a realizar el análisis con el servidor DynaMut para predecir mutaciones que desestabilizan la estructura. Los resultados se muestran en la Tabla 16 e indican que la mayoría de las diferencias observadas entre P3441 y C#1 son estabilizadoras. Los valores y magnitud de los resultados también varían tanto respecto a lo obtenido con MAESTRO como entre sí. La mayor estabilización predicha resultó para la mutación S1698F con un valor de 0,806, la más alta de los valores obtenidos. Esta diferencia es única para C#1. La mayor desestabilización predicha ocurrió para V2040A con un valor de -0,887, siendo una diferencia de C#1 y XJ13 respecto a P3441. Con respecto a la flexibilidad calculada con $\Delta\Delta S_{vib}$. La mayoría de los valores fueron cercanos a 0 o menores a 0,2, lo que sugiere un aumento reducido o poco significativo en la flexibilidad estructural.

Tabla 15. Cambio de estabilidad ($\Delta\Delta G_{\text{pred.}}$) calculada como mutación independiente de un residuo en la estructura para la proteína L de P3441 a un residuo en otra cepa (X).

Mutación (P >> C#1,XJ13)	$\Delta\Delta G_{\text{pred.}}$	$C_{\text{pred.}}$
H76Y	-0,121	0,817
D265K	-0,201	0,809
M415A	-0,232	0,795
D462N	-0,309	0,787
L936P	-0,126	0,819
K1059R	-0,369	0,771
R1156K	-0,251	0,788
C1344F	-0,254	0,796
K1582R	-0,234	0,806
Q1668K	-0,181	0,817
S1698F	-0,304	0,799
V1700I	-0,145	0,804
I1883V	-0,282	0,793
N2005S	-0,041	0,853
V2040A	-0,054	0,845

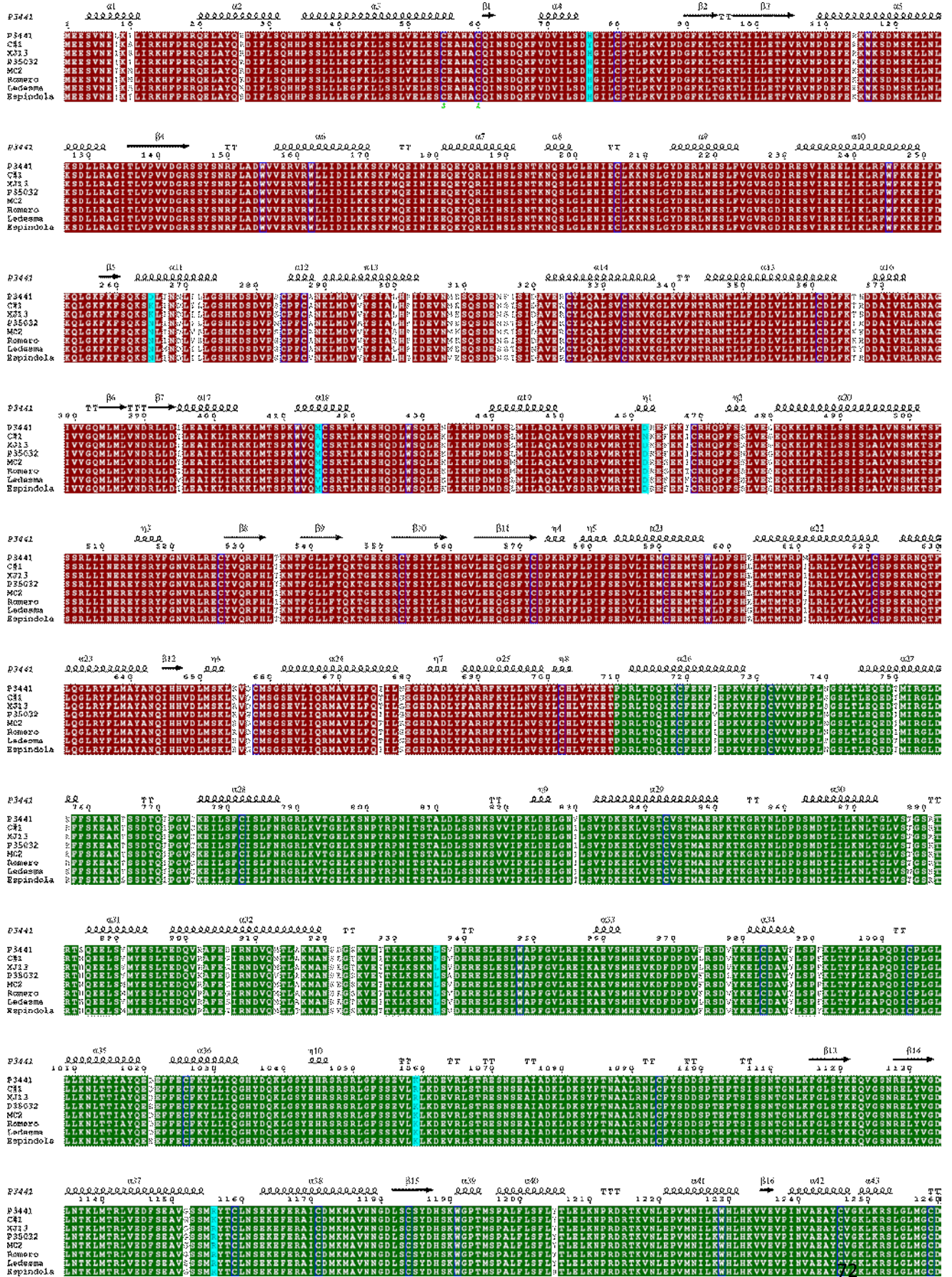
Valores de $\Delta\Delta G_{\text{pred.}} > 0$ se consideran desestabilizadores. $C_{\text{pred.}}$ Es un valor de confianza de la predicción, siendo $C_{\text{pred.}} = 0$ de mínima confianza e $C_{\text{pred.}} = 1$ máxima confianza de predicción.

Tabla 16. Cambio de estabilidad ($\Delta\Delta G_{\text{pred. DynaMut}}$) y entropía vibracional ($\Delta\Delta S_{\text{vib}}$) calculada como mutación independiente de un residuo de la proteína L de P3441 a un residuo de una cepa (X).

Mutación (P > X)	$\Delta\Delta G_{\text{pred. DynaMut}}$	$\Delta\Delta S_{\text{vib}}$
H76Y (X = C#1)	0,238	-0,019
D265K (X = C#1,XJ13)	0,363	-0,029
M415A (X = C#1)	0,197	0,231
D462N (X = C#1)	0,015	0,023
L936P (X = C#1)	-0,372	0,184
K1059R (X = C#1,XJ13)	0,295	0,053
R1156K (X = C#1,)	0,238	0,092
C1344F (X = C#1,XJ13)	-0,134	-0,295
K1582R (X = C#1,XJ13)	-0,402	0,637
Q1668K (X = C#1,XJ13)	0,429	0,09
S1698F (X = C#1)	0,806	-0,379
V1700I (X = C#1,XJ13)	0,36	-0,225
I1883V (X = C#1)	-0,225	0,152
N2005S (X = C#1,XJ13)	0,675	0,16
V2040A (X = C#1,XJ13)	-0,887	0,915

$\Delta\Delta G_{\text{pred. DynaMut}}$ corresponde a la diferencia de energía libre entre el modelo salvaje y el mutante calculado por el servidor DynaMut, el cual define valores negativos (< 0) como mutación desestabilizadora. $\Delta\Delta S_{\text{vib}}$ es la diferencia de energía de entropía vibracional entre el modelo salvaje y el mutante, con valores positivos (> 0) para mutaciones en las cuales aumenta la flexibilidad molecular.

A



P3441 $\alpha 44$ $\beta 17$ $\alpha 45$ $\beta 18$ $\beta 19$ $\eta 11$ $\alpha 46$ $\beta 20$ $\beta 21$ $\beta 22$
 1279 1289 1299 1309 1319 1329 1339 1349 1359 1369 1379 1389

P3441 C11 XJ13 P35032 MC2 Romero Ledesma Espindola

P3441 $\alpha 47$ $\alpha 48$ $\alpha 49$ TT $\alpha 50$ $\alpha 51$
 1400 1410 1420 1430 1440 1450 1460 1470 1480 1490 1500 1510

P3441 C11 XJ13 P35032 MC2 Romero Ledesma Espindola

P3441 $\alpha 52$ $\alpha 53$ $\alpha 54$ $\beta 23$ TT $\alpha 55$ $\alpha 56$ TT $\alpha 57$ $\beta 24$ $\beta 25$
 1520 1530 1540 1550 1560 1570 1580 1590 1600 1610 1620 1630

P3441 C11 XJ13 P35032 MC2 Romero Ledesma Espindola

P3441 $\alpha 58$ $\eta 12$ $\beta 26$ $\alpha 59$ $\eta 13$ TT $\beta 27$ $\alpha 60$ $\alpha 61$ $\eta 14$
 1640 1650 1660 1670 1680 1690 1700 1710 1720 1730 1740 1750 1760

P3441 C11 XJ13 P35032 MC2 Romero Ledesma Espindola

P3441 $\alpha 62$ $\alpha 63$ $\beta 28$ $\alpha 64$ TT $\alpha 65$ $\beta 29$ $\eta 15$ TT $\alpha 66$
 1770 1780 1790 1800 1810 1820 1830 1840 1850 1860 1870 1880 1890 1900

P3441 C11 XJ13 P35032 MC2 Romero Ledesma Espindola

P3441 TT $\alpha 67$ $\alpha 68$ $\alpha 69$ $\beta 30$ $\beta 31$ $\beta 32$ $\alpha 70$
 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990 2000 2010

P3441 C11 XJ13 P35032 MC2 Romero Ledesma Espindola

P3441 $\alpha 71$ TT $\beta 33$ $\beta 34$ $\beta 35$ $\alpha 72$ $\alpha 73$ $\alpha 74$
 2020 2030 2040 2050 2060 2070 2080 2090 2100 2110 2120 2130 2140

P3441 C11 XJ13 P35032 MC2 Romero Ledesma Espindola

P3441 $\alpha 74$ $\alpha 75$
 2150 2160 2170 2180 2190 2200 2210

P3441 C11 XJ13 P35032 MC2 Romero Ledesma Espindola

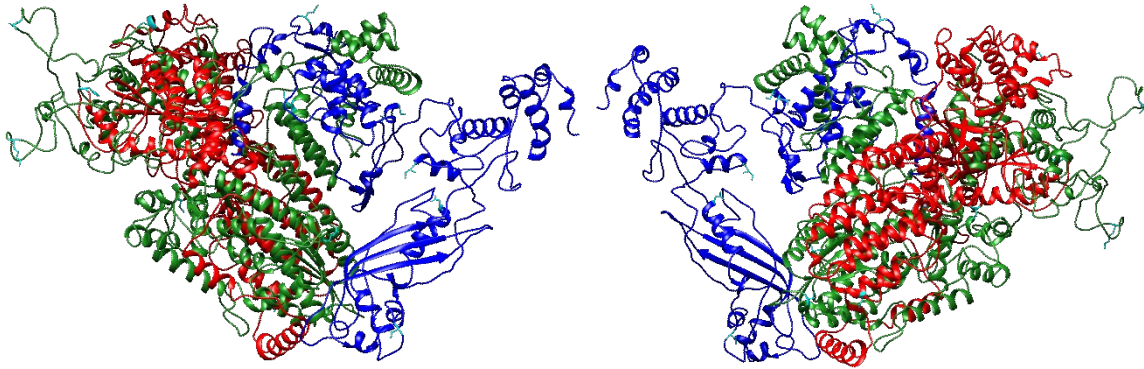
B

Figura 5.2. (A) Alineamiento de secuencias de aminoácidos para la proteína L de la cepa P3441, C#1 y otras cepas patogénicas de JUNV. Las equivalencias se encuentran marcadas en color, con las diferencias presentes en blanco o cian, y la estructura secundaria de la proteína puede observarse por encima del alineamiento, con α indicando alfa hélices, β lámina beta, η hélices 3_{10} , y T para giro beta. (B) Estructura tridimensional de la proteína JUNV L obtenida por ColabFold (ver materiales y métodos). Cada imagen corresponde a lados opuestos de la proteína. Se muestran el dominio N terminal similar a PA de Influenza (rojo), C terminal similar a PB2 de influenza (azul) y el dominio central asociado a la actividad polimerasa (verde). En cian se muestran los sitios con diferencias de interés y los cationes Zn coordinados por la proteína están representados como esferas grises.

Las figuras 5.3, 5.4, 5.5, 5.6 y 5.7 muestran las mutaciones L936P, C1344F, K1582R, I1883V y V2040A (respectivamente) y los cambios generados en las interacciones con los residuos cercanos.

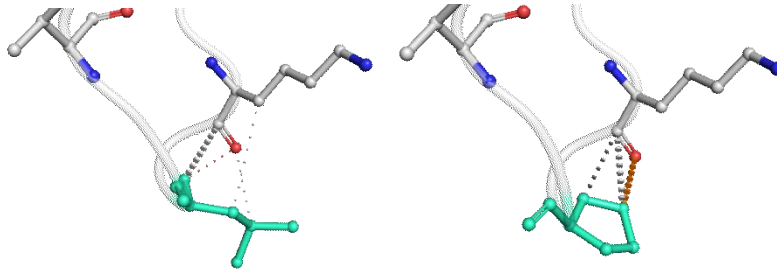


Figura 5.3. Esquema tridimensional del residuo 936 de JUNV L. A la izquierda se encuentra la estructura para P3441 y a la derecha se encuentra la estructura para C#1. En verde se ve la leucina (izquierda) y prolina (derecha) correspondiente a cada cepa. En gris se ven los carbonos (C), en rojo los oxígenos (O) y en azul los nitrógenos (N). En amarillo se muestran átomos de azufre (S).

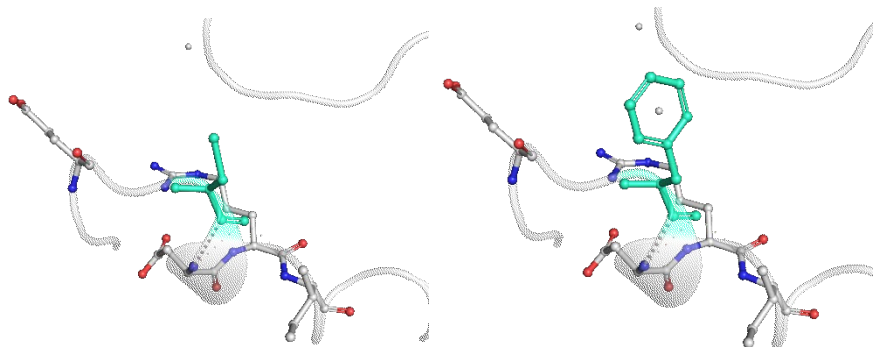


Figura 5.4. Esquema tridimensional del residuo 1344 de JUNV I. A la izquierda se encuentra la estructura para P3441 y a la derecha se encuentra la estructura para C#1. En verde se ve la cisteína (izquierda) y fenilalanina (derecha) correspondiente a cada cepa. En gris se ven los carbonos (C), en rojo los oxígenos (O) y en azul los nitrógenos (N).

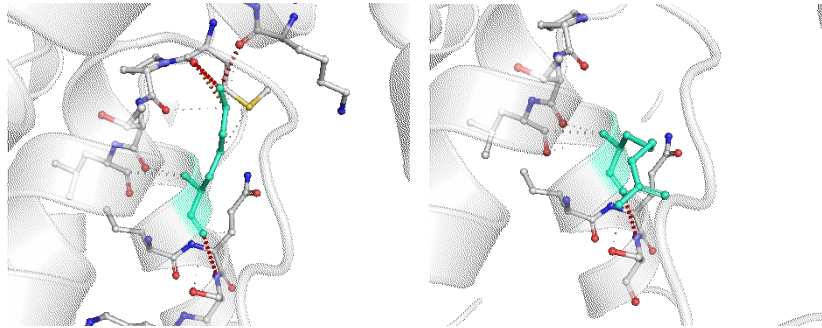


Figura 5.5. Esquema tridimensional del residuo 1582 de JUNV L. A la izquierda se encuentra la estructura para P3441 y a la derecha se encuentra la estructura para C#1. En verde se ve la lisina (izquierda) y arginina (derecha) correspondiente a cada cepa. En gris se ven los carbonos (C), en rojo los oxígenos (O) y en azul los nitrógenos (N). En amarillo se muestran átomos de azufre (S).

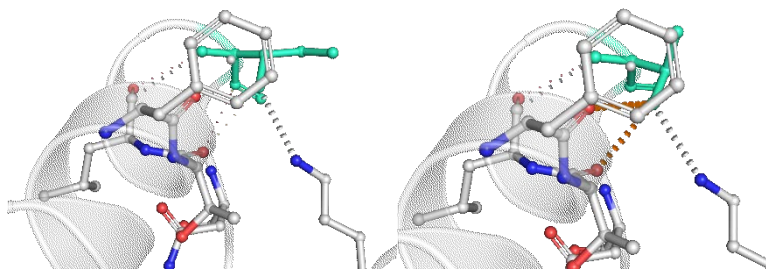


Figura 5.6. Esquema tridimensional del residuo 1883 de JUNV L. A la izquierda se encuentra la estructura para P3441 y a la derecha se encuentra la estructura para C#1. En verde se ve la isoleucina (izquierda) y valina (derecha) correspondiente a cada cepa. En gris se ven los carbonos (C), en rojo los oxígenos (O) y en azul los nitrógenos (N). En amarillo se muestran átomos de azufre (S).

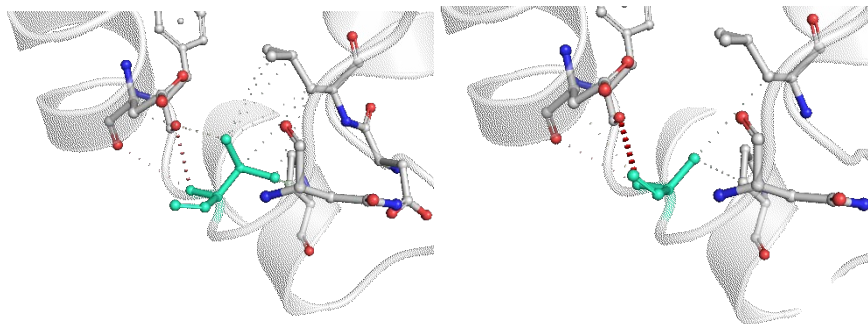


Figura 5.7. Esquema tridimensional del residuo 2040 de JUNV L. A la izquierda se encuentra la estructura para P3441 y a la derecha se encuentra la estructura para C#1. En verde se ve la valina (izquierda) y alanina (derecha) correspondiente a cada cepa. En gris se ven los carbonos (C), en rojo los oxígenos (O) y en azul los nitrógenos (N). En amarillo se muestran átomos de azufre (S).

5.2.2 Generación de un vector de expresión eucariota de L

Como fue descrito en el capítulo 1, se desarrolló una estrategia para subclonar la polimerasa viral en un vector de expresión eucariota con un fluoróforo C terminal. Tercios de la secuencia de L fueron clonados en el vector pGEM-T para después subclonar en conjunto en el vector de expresión

eucariota pEGFP-N3 de forma de obtener el marco de lectura completo de la proteína L. Después de varios intentos infructuosos, se decidió contratar los servicios de clonado de una empresa.

Inicialmente se contrató el servicio de clonado de WatsonBio Sciences (EE.UU), previamente utilizado para el clonado de la proteína N de P3441. Después de dos intentos, la empresa no logró completar el pedido. GenScript Biotech Corporation (Hong Kong), nuestra segunda alternativa, se rehusó a realizar el trabajo. Finalmente, se contrataron los servicios de la empresa Synbio Technologies (EE.UU).

Aunque extendiéndose del período estipulado debido a dificultades diversas, finalmente el plásmido pEGFP-L de P3441 fue completado (Figura 5.8). El producto entregado se confirmó por PCR. No obstante, la suspensión de pEGFP-L demostró tener una cantidad menor a la esperada. Luego de varios intentos infructuosos de transformar el plásmido, se encargó una mayor cantidad a la misma empresa con un grado de purificación que eliminara endotoxinas potenciales. Aunque la empresa demoró ante complicaciones en amplificar el plásmido, finalmente fue entregado un nuevo tubo de pEGFP-L liofilizado y se pudo verificar la presencia del plásmido en gel de agarosa.

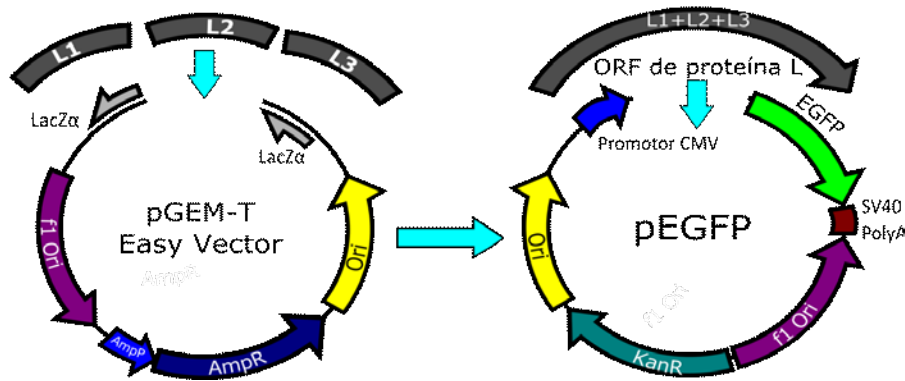


Figura 5.8. Esquema del clonado del marco de lectura para la proteína L de JUNV. Se partió de un vector que permitía la clonado directo de fragmentos amplificados de PCR para poder realizar la posterior digestión y ligación en el vector de expresión eucariota.

Luego se procedió con el estudio de su distribución celular transfectando células A549. En el primer ensayo no se observó la expresión de la proteína L en los 3 días que duro el estudio. Se sospechó que el considerable largo del marco de lectura desfavorecía su expresión y por tanto se extendió el período de observación hasta 6 días antes de fijar las células. En esa instancia se pudieron observar puntos en el citoplasma para pEGFP-L, lo cual coincide con lo esperado para L en la forma focos de replicación, mientras que pEGFP mostró fluorescencia distribuida en toda la célula, concentrándose principalmente en el núcleo (Figura 5.9).

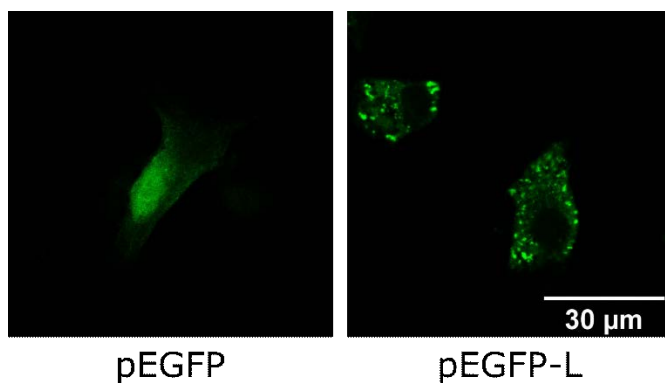


Figura 5.9. Imágenes de microscopía confocal de células A549 transfectadas con el plásmido pEGFP vacío o con el ORF de L de P3441. En verde se observa la fluorescencia correspondiente a EGFP, tanto la proteína sola (izquierda) o como parte de proteína recombinante con JUNV L de P3441 (derecha).

5.3 Discusión

Por primera vez se pudo conocer la secuencia de la polimerasa L de la cepa P3441 de JUNV. Esta secuencia fue posteriormente comparada con aquellas de otras cepas y se pudieron identificar diferencias distribuidas a lo largo de L que resultarían relevantes por la ubicación en los dominios estructurales de la polimerasa. El mayor tamaño de la proteína L explica que se hayan encontrado mayor número de diferencias entre cepas con respecto a las proteínas anteriores. Si bien el porcentaje de similitud aminoacídica fue muy elevado, resultó ligeramente menor a las otras proteínas quizás por el mencionado rol en la selección evolutiva (295, 296).

La proteína L de JUNV fue posteriormente modelada para estudiar la posición e impacto de las diferencias entre las cepas. Se utilizaron servidores y algoritmos para determinar cuáles de estas diferencias podrían tener el mayor impacto dentro de la estructura de L de P3441 respecto a L de la cepa C#1, estudiándose como mutaciones de la cepa P3441 a C#1. No resultó sorprendente que la mutación L936P se predijera como desestabilizadora ya que la rigidez del residuo prolina suele alterar estructuras preexistentes y forzar giros. Sin embargo, fue curioso que se acompañara de un aumento de la flexibilidad estructural. Esta mutación se encuentra en un *loop* expuesto y sin estructuras cercanas, de manera que no pareciera impactar directamente en la estructura terciaria de la proteína. Esta posición aislada podría explicar la existencia de un residuo que afecte la estructura secundaria de la proteína sin consecuencias. Fue esperable que otras diferencias entre las cepas se localizaran en posiciones cercanas.

En el caso de C1344F, la desestabilización calculada se puede explicar por el cambio de un aminoácido polar e hidrófilo por su grupo tiol (-SH) para la cisteína a otro aminoácido con una cadena lateral más voluminosa con un anillo bencénico que confiere hidrofobicidad en una fenilalanina. Las interacciones con cadenas cercanas limitan el movimiento de la cadena lateral de la fenilalanina en mayor medida que la cadena lateral de la cisteína, lo cual explica la reducción de flexibilidad predicha. En el caso de K1582R, tanto la lisina de P3441 como la arginina de C#1 y XJ13 cuentan con cadenas laterales protonables que pueden funcionar como bases y que podrían explicar la tolerancia de este cambio en la proteína. No obstante, la lisina está más asociada a la formación de enlaces estabilizadores de estructuras terciarias y cuaternarias por tener mayor densidad de carga que la arginina, justificando la desestabilización de esta mutación. La mutación I1883V es similar a K1582 en el sentido que se mantienen las propiedades químicas en el sitio, con la isoleucina de P3441 y la

valina de C#1 siendo similarmente hidrofóbicas. No obstante, la isoleucina cuenta con una cadena lateral más grande que facilita aproximarse a cadenas laterales cercanas para interactuar y estabilizar la estructura terciaria. Similarmente, la mutación V2040A mantiene similitud con I1883V, ya que tanto la valina de P3441 y la alanina de C#1 cuentan con cadenas laterales alifáticas hidrofóbicas. La diferencia entre estos residuos es de un grupo metilo en la cadena lateral de la valina que le confiere mayor tamaño respecto a la alanina.

En conjunto, las diferencias mostraron que la estructura de C#1 resultaba más estable respecto a la de P3441 con excepción en cinco residuos. Si bien se podría asumir que la estabilidad es un atributo que favorece a la proteína, este no sería necesariamente el caso. Considerando que se ha observado por ensayos de cristalografía que la polimerasa L de mammarenavirus se acomoda a su sustrato de ARN (297, 298), es posible que mayor rigidez dificulten los procesos necesarios para el funcionamiento óptimo de la proteína. Resta establecer en futuros estudios si esta mayor estabilidad se traduce en menor número de errores de copia y/o eficiencia de replicación. También es posible que las cinco diferencias que se detectaron como desestabilizadoras con DynaMut impacten sobre el funcionamiento de la polimerasa L, particularmente la diferencia en el residuo 1588 que tiene una ubicación dentro de un dominio definido de L llamado "pulgar" (258, 298).

Lograda la subclonado de la proteína L en un vector de expresión eucariota con fluorescencia, se demostró que esta construcción expresaba la proteína recombinante. Esta herramienta servirá en proyectos futuros para profundizar en las predicciones y observaciones realizadas a lo largo de este capítulo, estudiar su interacción con otros componentes virales y eventualmente, su función.

Capítulo 6. Sesgo de uso de codones

6.1 Introducción

El código genético está compuesto por 64 codones, secuencias de 3 nucleótidos (o bases) en una molécula de ARNm que corresponden a un aminoácido específico o señal de terminación (o STOP) durante la síntesis proteica. De esos codones, 61 traducen en 20 aminoácidos de manera que varios codones codifican para un mismo aminoácido (con las excepciones de metionina y triptofano). El uso de estos codones sinónimos no es uniforme o equitativo. De hecho, no es el mismo para todos los organismos. Esta peculiaridad se ha denominado sesgo de uso de codones (299).

El sesgo de uso de codones se debe a diferentes razones: factores transcripcionales, contenido GC (guanina-citosina), traducción, los niveles de expresión de genes y motivos de estructura secundaria del ARNm (300, 301). Una combinación de procesos de selección (como la disponibilidad de los diferentes ARNt propios de un organismo) y mutaciones (causadas de diferentes formas como mutaciones espontáneas por error de replicación) definen la evolución de estos sesgos (302), siendo la selección natural y la presión de mutaciones las causantes principales de los patrones observados (303, 304). Se define como presión de mutaciones a la tendencia o influencia de procesos mutacionales distintivos que introducen tipos específicos de mutaciones o sesgos en la secuencia de un organismo. Por ejemplo, se ha determinado que existe una mayor frecuencia de transiciones (cambios entre adenina y guanina o entre citosina y timina) respecto a transversiones (cambio entre las pirimidinas timina y citosina y las purinas adenina y guanina) (305, 306), y se ha podido explicar cómo un sesgo propio de las polimerasas (307-309). Por su parte, la selección natural es el proceso evolutivo donde se perpetúan rasgos en una población cuando favorecen la supervivencia y/o reproducción de los organismos en cuestión, o desaparecen cuando los mismos reducen la supervivencia y/o capacidad de reproducirse del organismo.

Dentro de un mismo genoma se pueden llegar a encontrar un amplio rango de distribuciones de uso de codones (310), lo cual se ha explicado como una herramienta para controlar los niveles de expresión y el plegamiento de las proteínas (300, 301, 311). De hecho, dentro de diferentes regiones para un mismo marco de lectura (ORF) se han observado sesgos diferentes. Cuanto más se conoce del sesgo de uso de codones, más se descubre su rol en la maquinaria traduccional de cada organismo (312).

Aunque se ha propuesto que la presión de mutaciones es el principal determinante del sesgo de uso de codones (313), también se propuso que este no es el caso para los virus de ARN o ADN (314). Para comenzar, los genomas virales requieren de la maquinaria celular de su hospedante para replicarse y, por tanto, la evolución de los virus está fuertemente asociada a presiones de selección específicas del hospedante como también aquella dirigida por la respuesta inmune (315, 316). En otras palabras, el estudio del sesgo de uso de codones puede brindar información sobre la evolución viral.

El sesgo de uso de codones se ha estudiado en varios virus. Por ejemplo, en el virus Zika se observó un sesgo reducido y una mayor influencia de la presión mutacional respecto a presiones de selección, inclusive si factores geográficos y procesos evolutivos tuviesen cierto impacto en dicho sesgo (317). Dentro de los virus de la familia *Hantaviridae*, género del orden *Bunyavirales* donde también se encuentra la familia *Arenaviridae*, se analizó el sesgo de uso de codones del virus Hantaan (HTNV) usando su genoma completo y se descubrió que la selección natural impuesta por el reservorio

natural y hospedante tenía más influencia en este fenómeno respecto a sesgos mutacionales (318). Adicionalmente, se estudiaron para diferentes especies de la familia que producían fiebre hemorrágica con síndrome renal, que incluía a HTNV y los virus Dobrava-Belgrade (DOBV), Seoul (SEOV) y Puumala (PUUV). En este caso se analizó el sesgo de codones para la proteína de nucleocápside de cada virus únicamente (319). De este último estudio se concluyó que entre estas cuatro especies virales había una preferencia por codones terminados en adenosina o uracilo y la presión de mutación tuvo mayor fuerza en la formación del patrón de uso de codones observado (319).

Como se describió previamente, el estudio del sesgo de uso de codones puede aplicarse en el estudio evolutivo de una especie, incluyendo el estudio particular para una especie virus cuando encuentra un nuevo hospedante. Gracias al extensivo seguimiento y secuenciación de SARS-CoV-2 durante la pandemia entre los años 2019 y 2021, se pudo realizar un estudio extensivo en su uso de codones desde su diseminación mundial entre los humanos y se observó como el patrón de uso de codones de este virus se alejaba de aquel del hospedante humano en el tiempo (320).

Dentro de los mammarenavirus (o arenavirus en general), ha habido pocos estudios respecto al sesgo de uso de codones. Por ejemplo, los patrones de uso de codones fueron explorados para SABV y CHAPV (321) pero la mayoría de los estudios encontrados se enfocaban principalmente en la optimización o deoptimización de proteínas virales para la atenuar expresión en potenciales candidatos vacunales (322, 323). Aunque se ha encontrado información del sesgo de uso de codones de los arenavirus, este ha sido en su generalidad sin incluir las diferencias individuales entre géneros, mucho menos apreciando potenciales diferencias entre proteínas y clados dentro de un mismo género (324).

Con el fin de completar la caracterización molecular de la cepa P3441 e identificar diferencias claves con otras cepas, se estudió el sesgo de uso de codones en JUNV.

6.2 Resultados

6.2.1 Composición

Para comenzar se utilizó el servidor CAIcal para determinar la composición porcentual de guanina y citosina tanto en la secuencia completa (GC%) como en la tercera posición de los codones (GC3%) para las proteínas de JUNV utilizando las secuencias de las cepas comparadas en este trabajo. Los resultados se muestran en la Figura 6.1 donde se observan diferencias significativas entre las proteínas. En lo que respecta la comparación entre cepas, la cepa MC2 presentó el mayor valor de GC% y GC3% en GPC, N y Z, mientras que C#1 tuvo el valor más elevado de GC% en L y P35032 el valor de GC3% en la polimerasa. La cepa P3441, por otro lado, se mantuvo con valores promedios, menos el caso de Z donde mostró el valor de GC% más bajo entre cepas. Aparte de L, C#1 no se distinguió de las otras cepas.

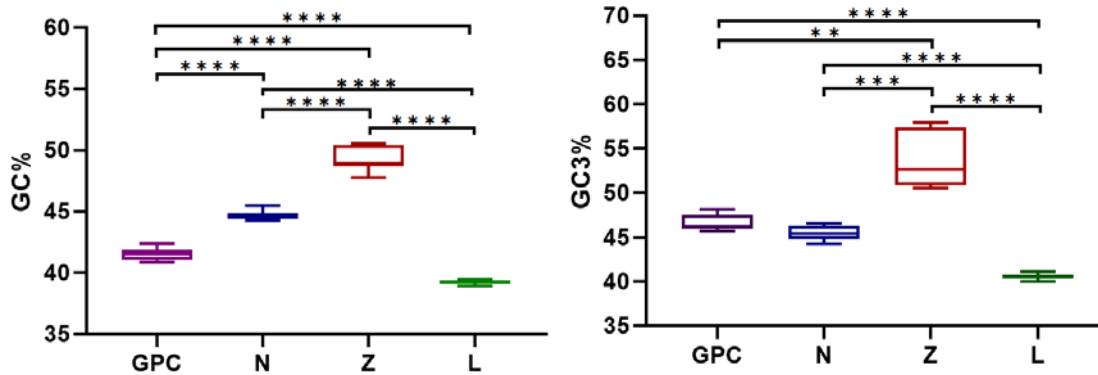


Figura 6.1. Composición porcentual de guanina-citosina total (GC%) (izquierda) y en la tercera posición de los codones (GC3%) (derecha) de la secuencia para las diferentes proteínas de JUNV. El análisis estadístico se hizo por ANOVA unidireccional bajo el método de Brown-Forsythe, aplicando el Tukey's test para comparaciones múltiples (**: $p < 0,005$, ***: $p < 0,0005$, ****: $p < 0,0001$). Se indicaron únicamente dónde hubo diferencias significativas.

Para la proteína Z se puede ver una mayor desviación estándar comparada a las otras proteínas debido principalmente a su menor tamaño. Con 95 codones (incluyendo el codón STOP), la secuencia ofrece mayor desviación cuando ocurre alguna modificación comparada con, por ejemplo, la polimerasa L, la cual muestra la menor desviación. Únicamente GPC y N no mostraron diferencias significativas para GC3%.

6.2.2 Análisis del uso relativo de codones sinónimos

El uso relativo de codones sinónimos (RSCU) es un número que representa la frecuencia en el que se encuentra un codón respecto a todos los codones que codifican para el mismo aminoácido para una secuencia dada. Por ejemplo, un RSCU = 1 para cualquier codón significaría que no hay sesgo (o que este codón no se encuentra más o menos presente en la secuencia que sus sinónimos), mientras que valores mayores a 1 representan un sesgo hacia dicho codón y un valor inferior a 1 se denomina con sesgo negativo, o cuyo uso es menos frecuente respecto a codones sinónimos (325). Se han denominado en trabajos previos un valor menor a 0,6 como subrepresentado, mientras que valores de RSCU > 1,6 como codones sobrerrepresentados (319, 326).

Los RSCU para las proteínas de todas las cepas fueron determinados por CAIcal. Ante la cantidad de datos a representar, los resultados se encuentran en el capítulo Anexo RSCU. Las tablas en dicho anexo han sido coloreadas en base a sus valores, cuya relación está representada en la escala de colores al final de cada tabla. Esta coloración permite visualizar diferencias entre cepas y codones del RSCU, como también facilita la comparación entre proteínas.

En lo que respecta a las diferencias entre proteínas, el patrón de uso de codones fue mayoritariamente similar, pero con claras diferencias en casos puntuales. Por ejemplo, CGT de arginina tuvo RSCU próximos a 1 en GPC, lo que indica falta de sesgo o preferencia por este codón, pero en otras proteínas tuvo valores más próximos a 0, lo que indica una preferencia contra su utilización. Similarmente, TGT tuvo RSCU ≈ 1 para GPC y $< 0,6$ para N y Z pero $> 1,4$ en L.

Siguiendo con la comparación entre proteínas y estableciendo dónde hay codones sub- o sobrerrepresentados (RSCU $< 0,6$ y $> 1,6$, respectivamente), se identificaron diferencias entre los valores de las proteínas GPC, L y NP. Se excluyó en este análisis Z por carecer de la suficiente cantidad

de codones para tener significancia estadística respecto a las otras proteínas en este análisis. Los resultados se resumen en la Tabla 17.

Tabla 17. Codones sub- y sobrerrepresentados en las proteínas GPC, N y L cuando 6 o más cepas tenían RSCU > 1,6 y < 0,6 respectivamente. *

Proteína	RSCU > 1,6	RSCU < 0,6
GPC	TTG (L), CCA (P), ACA (T), AGA (R), AGG (R)	TCG (S), CCG (P), ACG (T), GCG (A), GAG (E), CGC (R), CGA (R), CGG (R), GGG (G)
N	GTT (V), CCA (P), GCA (A), AGA (R), AGG (R)	CTA (L), GTA (V), TCG (S), CCC (P), CCG (P), ACG (T), GCG (A), TGT (C), CGT (R), CGC (R), CGA (R), CGG (R)
L	TTG (L), GTT (V), ACA (T), AGA (R), AGG (R)	CTC (L), GTA (V), TCG (S), CCG (P), ACG (T), GCG (A), TGC (C), CGT (R), CGC (R), CGA (R), CGG (R)

*En paréntesis se muestra el aminoácido al cual codifica el codón bajo el sistema de nomenclatura de una letra.

Respecto a los codones identificados, hay más codones subrepresentados que sobrerrepresentados. No hay un sesgo definido en la terminación de los codones identificados, pero para GPC se observa que 4 de 5 codones sobrerrepresentados terminan en A y 7 de 9 terminan en G para los subrepresentados. Similarmente para N, 4 de 5 codones sobrerrepresentados son de terminación AT, aunque para los subrepresentados sólo 7 de 12 tienen terminación GC.

Codones AGA y AGG del aminoácido arginina están sobrerrepresentados en las tres proteínas, mientras que CGC, CGA y CGG de arginina, GCG de alanina, ACG de treonina, CCG de prolina y TCG de serina están subrepresentados. Al mismo tiempo hay diferencias entre las proteínas. Por ejemplo, TTG de leucina está sobrerrepresentado en las proteínas GPC y L y GTA de valina está subrepresentado en N y L.

La diferencia de RSCU entre cepas difiere entre proteínas como también por codón. Por ejemplo, para el codón CTA (leucina) en GPC, C#1 tiene mayor sesgo negativo con un RSCU de 0,517 respecto a P3441, donde casi no tiene sesgo con un valor de 0,931. En la proteína L se observó menor variabilidad entre las cepas. Aparte de eso, en la generalidad de los casos se observó una tendencia a sesgos similares entre cepas. En N y GPC, se notó que MC2 parecía distinguirse de las otras cepas respecto a las demás proteínas.

6.2.3 Número efectivo de codones

El número efectivo de codones (ENC) expresa la preferencia por uso de codones, siendo un valor entre 20 a 61. Valores elevados reflejan una menor preferencia por codones sinónimos específicos, y viceversa. En la figura 6.2 se observa el análisis comparándose el ENC entre proteínas juntando los valores para las cepas utilizadas hasta el momento. Mientras N, Z y L parecen no contar diferencias significativas entre ellas, GPC demostró tener ENC significativamente mayor a las otras proteínas, lo que significa un menor sesgo. Entre las cepas, P3441 tuvo el mayor ENC para GPC (55,925), contrario a C#1 el cual tuvo el menor valor (52,917). Para N, P3441 tuvo el segundo mayor valor de ENC (51,58, segundo a Romero con 51,69) y C#1 nuevamente tuvo el menor valor (50,58). En Z, tanto C#1 como P tuvieron valores medios (51,28 y 50,01, respectivamente), diferentes a P35032 con 55,34. Curiosamente para L, C#1 y P ambos tuvieron los mayores valores (50,26 y 50,23, respectivamente), pero la diferencia con el menor valor fue mucho más estrecha entre cepas. Aun así, todos estos valores muestran un sesgo de uso de codones bastante reducido para JUNV.

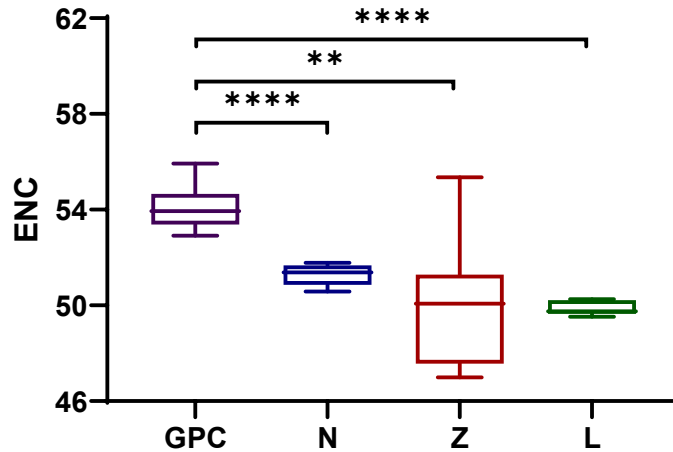


Figura 6.2. El número efectivo de codones (ENC) de la secuencia para las diferentes proteínas de JUNV. El análisis estadístico se hizo por ANOVA unidireccional bajo el método de Brown-Forsythe, aplicando el *Tukey's test* para comparaciones múltiples (**: $p < 0,005$, ****: $p < 0,0001$). Se indicaron únicamente dónde hubo diferencias significativas.

El gráfico ENC-GC3 ayuda a determinar el efecto de la composición de nucleótidos en el patrón de uso de codones sinónimos. Se utilizó el programa Rstudio con el lenguaje de programación R para armar el gráfico ENC-GC3. En este gráfico se comparan los valores con la línea estándar, o la curva que representa la situación dónde el ENC tiene relación directa con GC3%. La posición de los puntos relativos a la línea estándar en el gráfico ENC-GC3s da una idea de la presencia de un sesgo, dónde un mayor sesgo es reflejado como una a mayor distancia a dicha línea. Acorde a lo graficado para las proteínas de cada cepa, se pueden ver los resultados agrupados en cúmulos acorde a la proteína en la figura 6.3. Esto refleja una diferencia reducida entre cepas. Considerando los resultados anteriores, no sorprende que Z sea un cúmulo más disperso, pero su mayor distancia de la línea estándar sugiere un sesgo mayor al de las otras proteínas. En contraste, L presenta un sesgo visible, pero al encontrarse cerca de la línea estándar no es un sesgo apreciable como la proteína N. GPC cuenta con ENC más elevado pero tiene la misma distancia a la línea estándar que L.

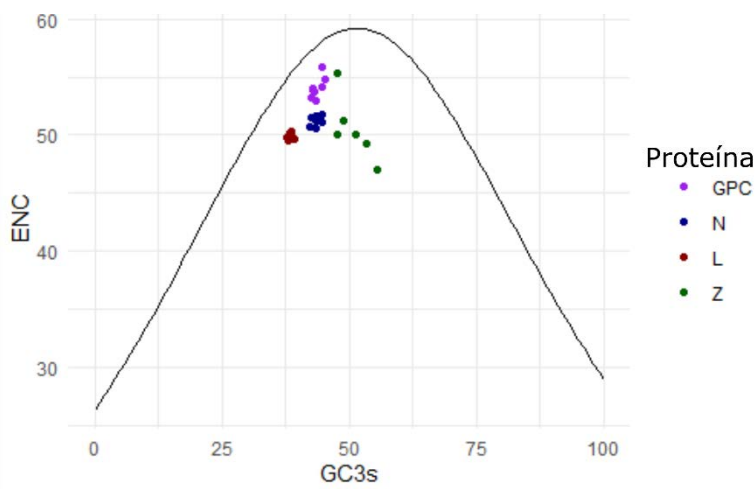


Figura 6.3. Gráfico ENC-GC3s. Las proteínas GPC (violeta), N (azul), L (rojo) y Z (verde) para cada cepa fueron graficadas en base a su ENC y GC3s%. La curva negra representa los valores de ENC esperados cuando ocurre relación directa con GC3s.

A manera de estudiar la posible influencia de la presión de selección, se graficó la relación entre la composición de guanina-citosina en la primera y segunda posición (GC12) respecto a aquella en tercera posición (GC3) en lo que se llama gráfico de neutralidad. Dicho gráfico se muestra en la figura 6.4. La pendiente de cada línea de regresión indica la medida la cual la presión de mutación explica los patrones de codones observados, comparándose con la recta que representa el escenario teórico donde la presión de mutación es el único factor que influye en el sesgo de uso de codones, o en otras palabras GC3 es igual a GC12. Los resultados muestran que el sesgo de uso de codones observado puede ser explicado mayoritariamente por presión de selección, siendo el efecto de presión de mutación más alta para GPC y N (con pendientes de 0,17 y 0,15, respectivamente) y más baja para L y Z (con 0,096 y -0,38, respectivamente). En otras palabras, al menos el 83% del sesgo de uso de codones puede ser explicado por selección.

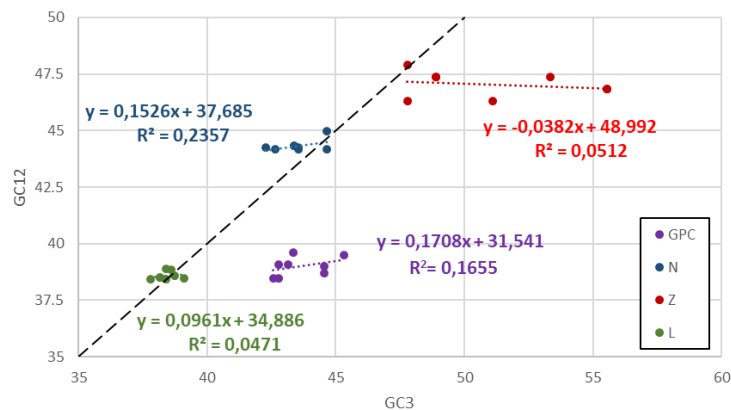


Figura 6.4. Gráfico GC12-GC3s. Las proteínas GPC (violeta), N (azul), L (rojo) y Z (verde) para cada cepa fueron graficadas en base a su GC12% y GC3s%. La recta negra representa los valores de ENC esperados cuando ocurre relación directa con GC3s.

6.2.4 Adaptación y deoptimización relativa de codones

El índice de adaptación de codones (CAI) predice el nivel de expresión génica relativo y mide la adaptación de los genes virales a un organismo de referencia, como por ejemplo el reservorio natural o hospedante. Su valor puede caer entre 0 a 1, representando la medida desde codones carentes de adaptación a una situación de adaptación completa (327). Se calcularon los CAI para las proteínas de JUNV de las diferentes cepas. Como no se cuenta con secuencias de *Calomys musculus*, y para continuar estudiando el impacto de JUNV en el ser humano, se utilizó a *Homo sapiens* como patrón de uso de codones sinónimos. Debido a las diferencias que ocurren entre tipos celulares y tejidos, este estudio se desglosó en dos partes: el cálculo de CAI para células del sistema inmune y para tejidos donde se han detectado efecto o presencia del virus. Para todos los casos se utilizaron las secuencias de las proteínas de mayor expresión génica para las células dadas por la base de datos *Human Protein Atlas*, con las cuales se obtuvieron el patrón de referencia para calcular el CAI en el servidor CAIcal.

Como se puede observar en la Figura 6.5, el CAI para las proteínas virales varía entre células del sistema inmune. Para todas las células estudiadas, la proteína L tenía el menor CAI mientras que la proteína Z tiene el mayor CAI. No se observaron diferencias significativas entre GPC y N dentro de cada célula del sistema inmune. Se observó CAI mayores para las células B, y menores para las células dendríticas, lo que sugiere una mejor adaptación de las proteínas virales en las células B.

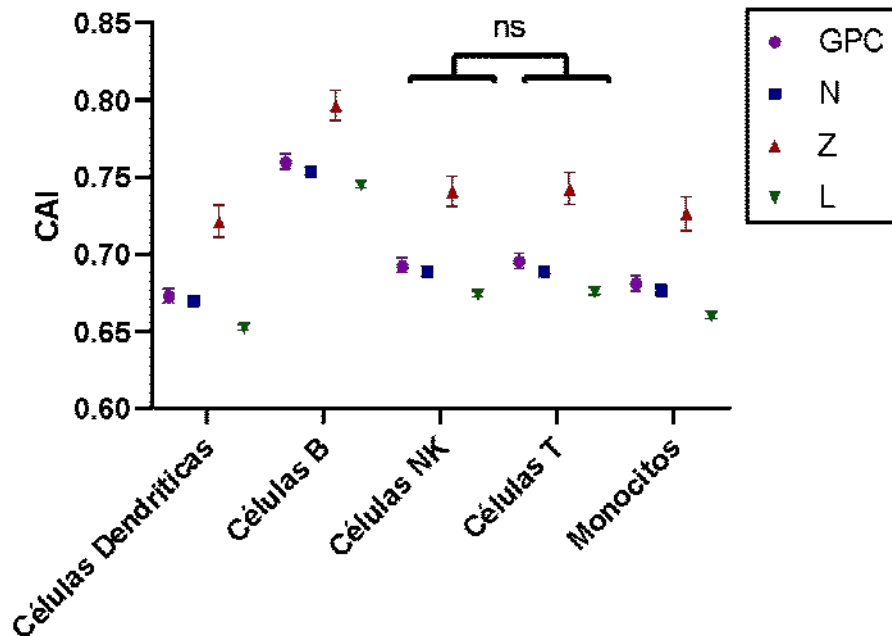


Figura 6.5. El Índice de Adaptación de codones (CAI) en diferentes tipos celulares del sistema inmune para las secuencias de las proteínas de JUNV. Se realizó un ANOVA bidireccional con el test de Tukey para comparación múltiple con el cual se determinaron diferencias significativas entre los resultados para los diferentes tipos celulares con un valor $p < 0,0001$. La única excepción fue la diferencia no significativa entre las células NK y células T y se muestra en el gráfico.

Respecto a las diferencias entre las cepas mostradas en la Tabla 18, MC2 tuvo el mayor CAI en GPC para casi todas las células, y P35032 el menor, con P3441 similar a P35032. Contrario a GPC, en las proteínas N y L la cepa P35032 tuvo el mayor valor de CAI para todas las células estudiadas, con la cepa Ledesma teniendo el menor CAI para N y XJ13 para L (excepto en las células B, donde la cepa Espíndola tuvo el menor valor). Para Z, la cepa Espíndola tuvo el mayor valor de CAI y P3441 el menor. Un patrón interesante observado cuando se compararon los resultados entre C#1 y P3441 fue que el CAI de C#1 fue mayor que los de P3441 en todas las células estudiadas para las proteínas GPC, N y Z, pero no para L. En la mayoría de los casos, P3441 tuvo un CAI menor que la de otras cepas patogénicas como Romero, Ledesma y Espíndola.

Tabla 18. Índice de adaptación de codones para las diferentes secuencias proteicas de las cepas de JUNV en base a diferentes tipos celulares del sistema inmune humano.

Proteína	Cepa	Dendríticas	Células B	Células NK	Monocitos	Células T
GPC	Candid1	0,677	0,764	0,696	0,685	0,699
	P3441	0,669	0,754	0,688	0,677	0,691
	XJ13	0,67	0,757	0,689	0,677	0,692
	P35032	0,667	0,754	0,686	0,675	0,69
	MC2	0,682	0,768	0,7	0,69	0,704
	Romero	0,675	0,763	0,696	0,684	0,699
	Ledesma	0,672	0,758	0,692	0,679	0,694
	Espindola	0,674	0,762	0,693	0,682	0,696
N	Candid1	0,671	0,755	0,69	0,678	0,69
	P3441	0,668	0,753	0,686	0,675	0,687
	XJ13	0,672	0,756	0,691	0,679	0,691
	P35032	0,675	0,758	0,693	0,683	0,694
	MC2	0,67	0,752	0,689	0,676	0,688
	Romero	0,671	0,755	0,69	0,678	0,69
	Ledesma	0,665	0,749	0,684	0,672	0,684
	Espindola	0,668	0,753	0,686	0,674	0,687
L	Candid1	0,652	0,744	0,674	0,66	0,675
	P3441	0,653	0,746	0,675	0,662	0,677
	XJ13	0,65	0,743	0,672	0,658	0,673
	P35032	0,656	0,748	0,678	0,665	0,68
	MC2	0,653	0,747	0,674	0,661	0,677
	Romero	0,654	0,747	0,675	0,661	0,677
	Ledesma	0,652	0,746	0,674	0,66	0,676
	Espindola	0,65	0,742	0,672	0,658	0,673
Z	Candid1	0,723	0,801	0,744	0,727	0,744
	P3441	0,7	0,78	0,72	0,704	0,722
	XJ13	0,723	0,801	0,744	0,727	0,744
	P35032	0,723	0,802	0,743	0,731	0,747
	MC2	0,727	0,797	0,745	0,732	0,748
	Romero	0,727	0,797	0,745	0,732	0,748
	Ledesma	0,712	0,785	0,733	0,718	0,733
	Espindola	0,734	0,811	0,752	0,741	0,755

En la Figura 6.6 se pueden ver los resultados para el mismo análisis en tejidos humanos. La mayor adaptación se observó para riñón, seguido de bazo e hígado, con valores de CAI mayor que para las células del sistema inmune. Los menores valores de CAI fueron observados para el tejido endotelial.

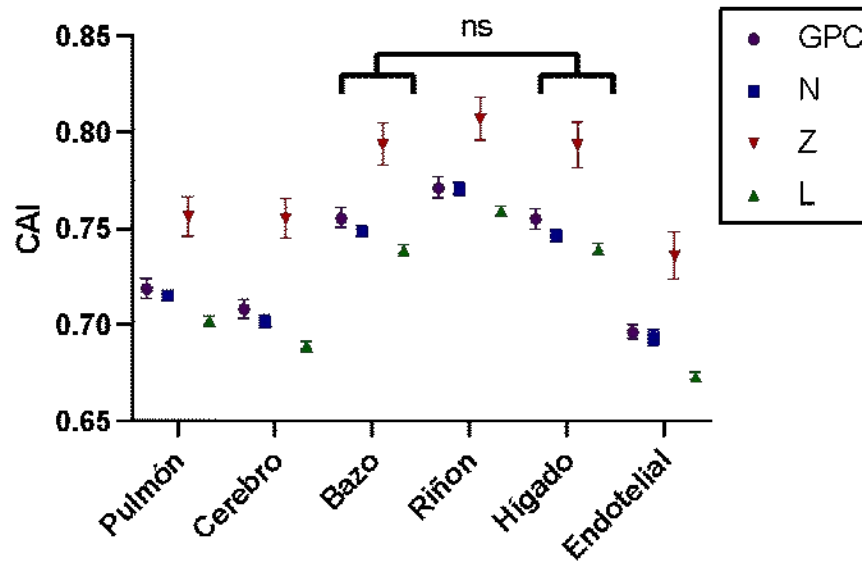


Figura 6.6. El Índice de Adaptación de codones (CAI) en diferentes tejidos para las secuencias de las proteínas de JUNV. Se realizó un ANOVA bidireccional con el test de Tukey para comparación múltiple con el cual se determinaron diferencias significativas entre los resultados para los diferentes tejidos con un valor $p < 0,0001$. La única excepción fue la diferencia no significativa entre los resultados de bazo e hígado y se muestra en el gráfico.

Respecto a los valores obtenidos para las diferentes cepas, se puede observar en la Tabla 19 que MC2 presentó el mayor valor en GPC, P35032 en N y L y Espíndola en Z. Por otro lado, se observó que el menor valor en GPC correspondió a P35032 en pulmón y cerebro, y P3441 en los otros tejidos. Para las otras proteínas, la cepa Ledesma mostró el menor CAI en N para todos los tejidos menos hígado, mientras que en L fue la cepa XJ13 la que presentó el menor CAI calculado. Nuevamente, el menor CAI para Z en todos los tejidos correspondió a P3441. Además, se mantuvo lo observado previamente donde C#1 tuvo mayor CAI que P3441 en GPC, N y Z, pero no en L.

Similar al CAI, el Índice de deoptimización relativa de codón (RCDI) estima la tasa de traducción de una secuencia dada respecto a una tabla de uso de codones de referencia. A diferencia del CAI, se calculó el RCDI con una tabla que reflejara el patrón de uso de codones para todo un organismo. En nuestro caso, no se cuenta con suficiente información para utilizar *Calomys musculinus* y nuevamente se utilizó la tabla de uso de codones para humanos, cómo se ha realizado para otras proteínas virales (328).

Como se puede ver en la Figura 6.7, Z tiene el mayor valor de RCDI, sugiriendo que es el menos optimizado para el humano. GPC, NP y L, por otro lado, parecen estar más optimizados y sin diferencias significativas entre ellas. Comparando entre cepas acorde a lo presentado en la Tabla 20, C#1 tiene el mayor valor de RCDI para Z (1,586), pero más cercano al promedio para las demás proteínas, mientras que P3441 mantiene el segundo valor más alto en Z, y valores más cercanos a C#1 en las otras proteínas respecto a las otras cepas, con excepción de XJ13 (que también conserva valores próximos a C#1).

Tabla 19. Índice de adaptación de codones para las diferentes secuencias proteicas de las cepas de JUNV en base a diferentes tejidos del ser humano.

Proteína	Cepa	Cerebro	Pulmón	Bazo	Riñón	Hígado	Endotelial
GPC	Candid1	0,712	0,723	0,76	0,777	0,76	0,7
	P3441	0,704	0,714	0,749	0,764	0,749	0,692
	XJ13	0,705	0,716	0,753	0,769	0,751	0,693
	P35032	0,702	0,712	0,75	0,765	0,75	0,693
	MC2	0,717	0,727	0,763	0,778	0,763	0,703
	Romero	0,711	0,722	0,76	0,776	0,76	0,698
	Ledesma	0,707	0,717	0,753	0,769	0,752	0,695
	Espíndola	0,709	0,72	0,757	0,772	0,757	0,697
N	Candid1	0,703	0,716	0,749	0,772	0,747	0,694
	P3441	0,7	0,714	0,748	0,77	0,746	0,692
	XJ13	0,704	0,717	0,751	0,773	0,748	0,696
	P35032	0,707	0,719	0,753	0,775	0,75	0,701
	MC2	0,701	0,714	0,746	0,767	0,743	0,69
	Romero	0,703	0,717	0,75	0,772	0,749	0,696
	Ledesma	0,697	0,711	0,744	0,766	0,743	0,689
	Espíndola	0,7	0,714	0,748	0,771	0,746	0,691
L	Candid1	0,689	0,701	0,738	0,758	0,738	0,673
	P3441	0,69	0,703	0,74	0,76	0,74	0,674
	XJ13	0,687	0,7	0,736	0,756	0,736	0,671
	P35032	0,693	0,706	0,743	0,763	0,744	0,677
	MC2	0,69	0,703	0,74	0,761	0,741	0,674
	Romero	0,691	0,704	0,741	0,761	0,741	0,675
	Ledesma	0,689	0,702	0,739	0,76	0,74	0,673
	Espíndola	0,687	0,7	0,737	0,756	0,737	0,672
Z	Candid1	0,757	0,758	0,797	0,812	0,796	0,734
	P3441	0,735	0,737	0,775	0,79	0,775	0,713
	XJ13	0,757	0,758	0,797	0,812	0,796	0,734
	P35032	0,758	0,759	0,802	0,815	0,807	0,741
	MC2	0,76	0,761	0,794	0,806	0,792	0,743
	Romero	0,76	0,761	0,794	0,806	0,792	0,743
	Ledesma	0,746	0,746	0,781	0,793	0,779	0,726
	Espíndola	0,769	0,77	0,809	0,823	0,81	0,754

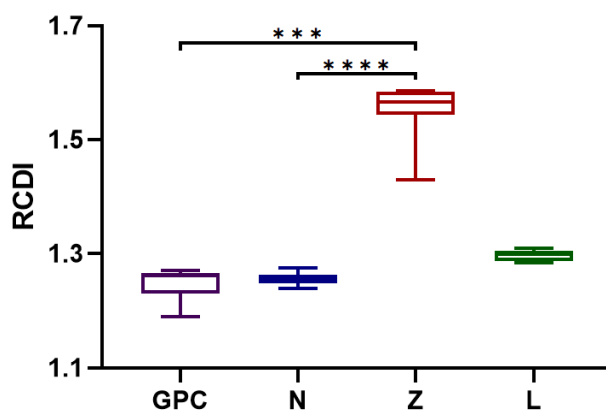


Figura 6.7. Comparación de RCDI entre las proteínas de JUNV usando valores para las cepas estudiadas. El análisis estadístico se hizo por medio de Kruskal-Wallis para datos no-paramétricos, aplicando el Dunn's test para comparaciones múltiples (***: $p < 0,005$, ****: $p < 0,0001$). Sólo se indicó dónde hubo diferencias significativas.

Tabla 20. Índice de desoptimización relativa de codones para las diferentes secuencias proteicas de las cepas de JUNV.

Cepa	GPC	N	Z	L
Candid1	1,243	1,254	1,586	1,285
P3441	1,226	1,257	1,579	1,287
XJ13	1,26	1,257	1,586	1,288
P35032	1,267	1,253	1,43	1,303
MC2	1,19	1,24	1,567	1,298
Romero	1,271	1,246	1,567	1,3
Ledesma	1,264	1,266	1,559	1,306
Espíndola	1,266	1,275	1,539	1,31
<u>PROMEDIO</u>	1,248375	1,256	1,551625	1,297125

6.2.5 Ensayo de correspondencia

El análisis de correspondencia es una técnica estadística multivariable para la identificación de variaciones en el sesgo de codones entre proteínas de interés (316). Se realizó este ensayo para los RSCU calculados. A manera de referencia y para comparar, se agregaron los RSCU promedio de humano proveniente del cálculo de 75.000 secuencias codificantes del genoma humano (el mismo usado para obtener la tabla de referencia en el cálculo del RCDI). Además, se incorporó la secuencia de ARNm para el receptor de transferrina 1 de *Calomys musculus*, siendo éste el único ARNm en base de datos para el ratón maicero.

Cómo se puede ver en la Figura 6.8 izquierda, las secuencias para las cepas forman cúmulos en base a la proteína, lo cual indica un sesgo de codones similar entre cepas. Hay una alta dispersión para Z, pero eso coincide con los resultados hasta el momento. No obstante, es curioso que Z de P3441 tenga una mayor aproximación a Z de Espíndola que a otras, lo cual sugiere una fuerte relación entre ambas cepas. Los valores del ratón maicero se encuentran en un ángulo agudo respecto al promedio para el cúmulo de la proteína L, lo cual sugiere que hay asociación positiva. Por el otro lado, el cúmulo de L está completamente opuesto al promedio humano, lo cual indica una asociación completamente negativa. Z muestra una tendencia opuesta a L, dónde el cúmulo de Z se encuentra próxima al humano pero opuesta a *Calomys musculus*. La distancia de los puntos del origen indica la importancia de las variables en nuestros datos, indicando que los cúmulos de GPC, Z y N parecen tener mayor peso que los datos de humano y L. GPC y N se encuentran en cúmulos separados y en ángulos grandes respecto del origen y los otros cúmulos, sugiriendo poca relación con estos y todavía menor entre ellos. En la figura 6.8 derecha se puede ver la posiciones de los codones para el mismo análisis de correspondencia. Comparando estas respecto a las posiciones de las secuencias, se puede ver proximidad entre los codones AGG, ATC, AAG y CAG por la proteína Z, CAT y TGT a GPC, AGA a N y TTC y TCT a L.

Las dimensiones en el ensayo de correspondencia indican las diferentes fuentes de variación entre un conjunto de puntos de datos multivariados. En este caso, el eje X (o Dimensión 1) representó casi la mitad de las variaciones (48,8%), mientras que el eje Y (Dimensión 2) representó el 16%. En conjunto, las dos primeras dimensiones explican el 64,8% de la variabilidad total de los datos. Esto indica que estas dos dimensiones capturan una parte significativa de la estructura de los datos.

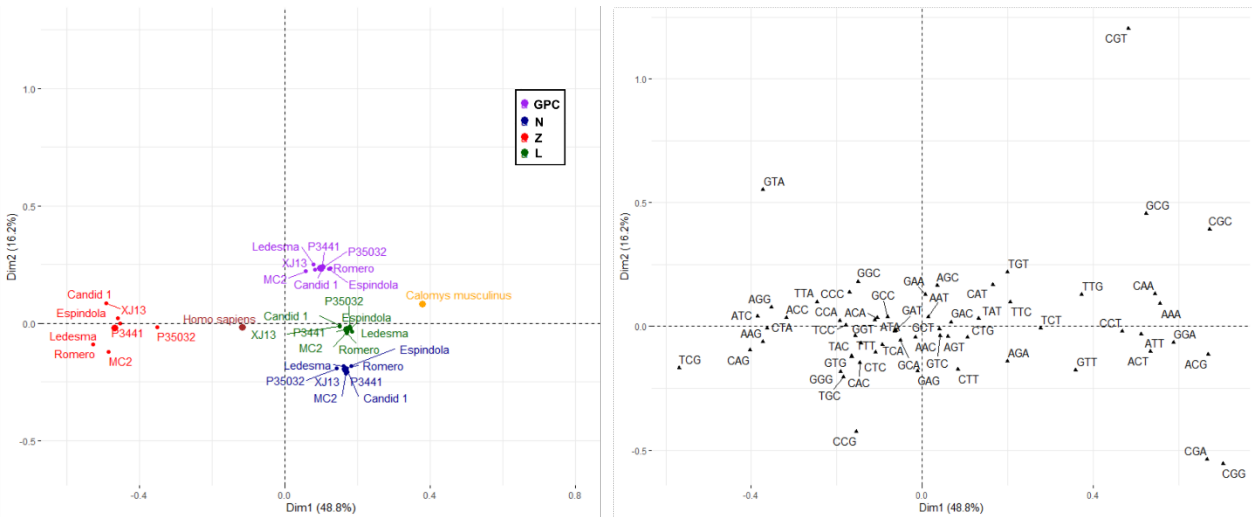


Figura 6.8. Análisis de correspondencia de los valores de uso relativo de codones sinónimos (RSCU). Izquierda: gráfico para las secuencias de las diferentes proteínas de JUNV (violeta para GPC, azul para N, rojo para Z, verde para L) y para cada cepa. Se agregaron a la comparación valores de RSCU para *Homo sapiens* (humano) y *Calomys musculus* (reservorio natural de JUNV) en marrón y naranja respectivamente. Derecha: La posición de cada codón para el mismo análisis se encuentra en negro.

6.2.6 Ensayo de correlación

La correlación por rango de Spearman (*Spearman's rank correlation*) es una medida no paramétrica que no asume una distribución particular, a diferencia de la correlación de Pearson, que asume una distribución normal y una relación lineal. El coeficiente de correlación puede variar de -1 a 1, indicando el nivel de la relación monótona entre variables. Es decir, valores próximos a 1 sugieren una relación fuertemente positiva (cuando una variable aumenta, la otra también, y viceversa) y cuando los valores son próximos a -1, la relación es negativa, en la cual el aumento de una variable corresponde a la disminución de la otra y viceversa. Valores más cercanos a 0 indican una relación débil o inexistente. Con el fin de determinar la existencia de correlación entre diferentes variables para JUNV, se realizó para cada proteína un análisis de correlación para todas las cepas. Dentro de las variables estudiadas se eligieron la composición de nucleótidos adenina (A), citosina (C), uracilo (U) y guanina (G) totales o en tercera posición de los codones (A3, C3, U3 y G3), como también la composición de guanina y citosina en primera (GC1), segunda (GC2) o tercera (GC3) posición, el promedio en las primeras dos posiciones (GC12) para codones usados, y el promedio general de hidrofobicidad (GRAVY). Se eligió el rango de Spearman después que algunas variables (por ejemplo, GC2) no pasaran el test de Wilk-Shapiro para verificar una potencial distribución normal. Los resultados para GPC, N, L y Z se muestran en las figuras 6.9, 6.10, 6.11 y 6.12, respectivamente.

Para GPC se observaron elevadas correlaciones, tanto negativas como positivas, algo que se visualiza en la intensidad de color en la parte inferior izquierda de la Figura 6.9. GC3 en particular parece tener fuerte correlación con las otras variables, incluyendo el contenido total de nucleótidos A, C y U. El ENC parece tener una correlación negativa elevada para A3 (-0,87) y positiva para G3 (0,72). En ambos casos se puede ver en la Tabla 20 que su valor p fue menor a 0,05, lo que demuestra que estas relaciones entre ENC y A3 y G3 fueron estadísticamente significativas. Una correlación relevante fue entre GC y los contenidos de A, C y U en tercera posición con valores -0,73, 0,82 y -0,88 (los tres con valores p < 0,05).

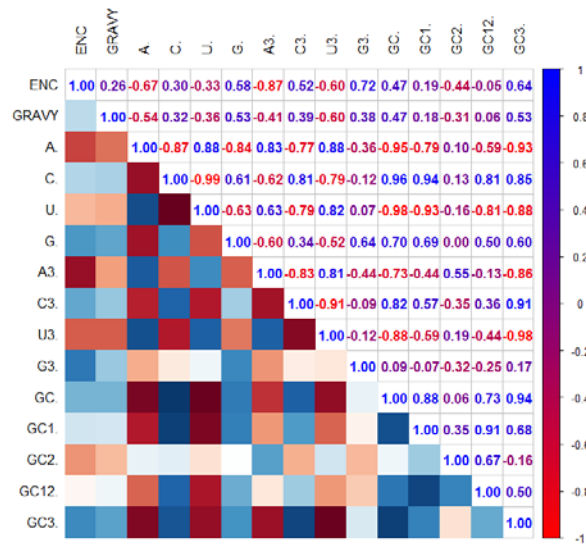


Figura 6.9. Correlación por método de Spearman entre componentes de la secuencia, ENC y otros para GPC de las cepas estudiadas. Azul oscuro indica correlación positiva mientras que rojo oscuro indica correlación negativa. Los valores más elevados indican una correlación más significativa.

Tabla 20. Valores p del análisis de correlación por rango de Spearman correspondiente a la Figura 6.9.

	ENC	GRAVY	A%	C%	U%	G%	A3%	C3%	U3%	G3%	GC%	GC1%	GC2%	GC12%	GC3%
ENC	NA	ns	ns	ns	ns	ns	**	ns	ns	*	ns	ns	ns	ns	ns
GRAVY	0.53640873	NA	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
A%	0.06640157	0.1650407	NA	**	**	**	**	*	**	ns	***	*	ns	ns	***
C%	0.47126053	0.4346394	0.00534729	NA	***	ns	ns	*	*	ns	***	**	ns	*	**
U%	0.43168839	0.3789641	0.00355753	5.49E-07	NA	ns	ns	*	*	ns	***	***	ns	*	**
G%	0.13430119	0.1787495	0.00828137	1.07E-01	9.16E-02	NA	ns	ns	ns	ns	ns	ns	ns	ns	ns
A3%	0.00542395	0.3099587	0.00992362	9.94E-02	9.68E-02	0.11944972	NA	**	*	ns	*	ns	ns	ns	**
C3%	0.18839393	0.3455111	0.02407017	1.57E-02	1.90E-02	0.40747399	0.00992362	NA	**	ns	*	ns	ns	ns	**
U3%	0.11396448	0.1139645	0.00412958	1.86E-02	1.33E-02	0.18468087	0.01394477	0.00145735	NA	ns	**	ns	ns	ns	***
G3%	0.04608687	0.3588778	0.37825871	7.85E-01	8.74E-01	0.08495136	0.27412813	0.82842246	7.72E-01	NA	ns	ns	ns	ns	ns
GC%	0.24326992	0.2432699	0.00038893	1.15E-04	1.48E-05	0.05133336	0.04142857	0.01304869	3.49E-03	0.82944758	NA	**	ns	*	***
GC1%	0.64738445	0.6684275	0.01895632	5.31E-04	7.17E-04	0.05846141	0.27741933	0.14269128	1.27E-01	0.87384754	4.06E-03	NA	ns	**	ns
GC2%	0.27412812	0.4473269	0.82177195	7.65E-01	7.06E-01	1	0.16193823	0.39441143	6.50E-01	0.43856909	8.81E-01	0.39441143	NA	ns	ns
GC12%	0.90977184	0.887313	0.12738767	1.57E-02	1.46E-02	0.21022848	0.76765967	0.38139905	2.76E-01	0.54788179	4.09E-02	0.00145735	0.06938451	NA	ns
GC3%	0.08833209	0.1765098	0.00092642	6.88E-03	4.13E-03	0.1186415	0.00672848	0.00145735	1.51E-05	0.69504606	5.31E-04	0.06524951	0.70615939	0.20703125	NA

En la figura 6.10 se muestran las correlaciones para N, las cuales fueron similares a GPC. Hay con claras diferencias en la medida de estas, visibles con claridad ante la diferencia de intensidad de colores. De hecho, el ENC no parece tener correlación significativa con alguna variable, mostrando valores entre -0,29 para U3 y 0,49 para GC2. De la misma manera se puede ver en la Tabla 21 que hay menos cantidad de correlaciones calculadas que tienen una significancia estadística (valores p < 0,05). GC y GC3 tienen un valor de correlación positiva (y valor p < 0,01) similar a GPC, esta correlación sugiere que el contenido de nucleótidos puede influenciar la tercera posición del codón, un fenómeno generalmente asociado a presiones mutacionales (302).

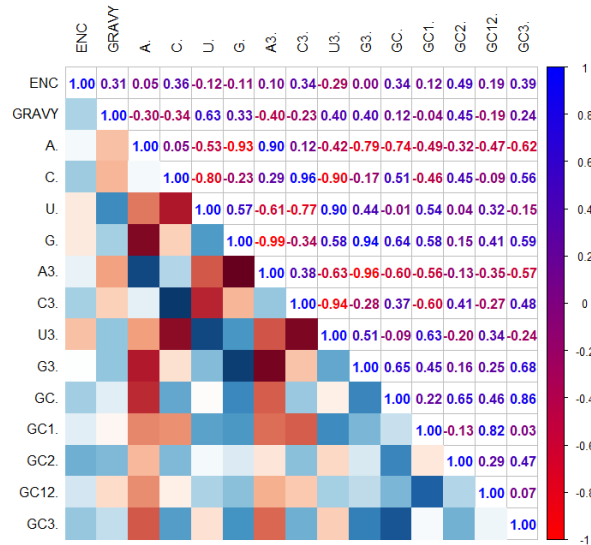


Figura 6.10. Correlación por método de Spearman entre componentes de la secuencia, ENC y otros para N de las cepas estudiadas. Azul oscuro indica correlación positiva mientras que rojo oscuro indica correlación negativa. Los valores más elevados indican una correlación más significativa.

Tabla 21. Valores p del análisis de correlación por rango de Spearman correspondiente a la Figura 6.10.

	ENC	GRAVY	A%	C%	U%	G%	A3%	C3%	U3%	G3%	GC%	GC1%	GC2%	GC12%	GC3%
ENC	NA	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
GRAVY	0.4527988	NA	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
A%	0.9097718	0.47504474	NA	ns	ns	***	**	ns	ns	*	*	ns	ns	ns	ns
C%	0.3893849	0.41679159	0.90977184	NA	*	ns	ns	ns	**	ns	ns	ns	ns	ns	ns
U%	0.777583	0.09649265	0.17930937	0.01654072	NA	ns	ns	**	**	ns	ns	ns	ns	ns	ns
G%	0.7945268	0.41960249	0.00097196	0.57827909	0.14182865	NA	****	ns	ns	**	ns	ns	ns	ns	ns
A3%	0.8181772	0.31960728	0.00222912	0.48161781	0.10566062	6.13E-01	NA	ns	ns	***	ns	ns	ns	ns	ns
C3%	0.4137853	0.58320069	0.78471031	0.00011421	0.02550171	4.17E-01	3.49E-01	NA	***	ns	ns	ns	ns	ns	ns
U3%	0.4787261	0.33256311	0.30503345	0.00258811	0.00221554	1.30E-01	9.48E-02	0.00056708	NA	ns	ns	ns	ns	ns	ns
G3%	1	0.32935526	0.01862891	0.69146535	0.27558952	4.05E-01	1.90E-01	0.49408785	0.194066	NA	ns	ns	ns	ns	ns
GC%	0.4076337	0.77219948	0.03551213	0.19422342	0.97699228	8.64E-02	1.16E-01	0.36640929	0.83575376	0.08075989	NA	ns	ns	ns	**
GC1%	0.7762571	0.92054848	0.22010184	0.25394134	0.16767276	1.34E-01	1.53E-01	0.11304951	0.09584891	0.25783394	0.59681884	NA	ns	*	ns
GC2%	0.222872	0.26366268	0.43499022	0.26687619	0.91580556	7.20E-01	7.57E-01	0.30840052	0.62779628	0.70403644	0.07836978	0.76034113	NA	ns	ns
GC12%	0.6604121	0.65846695	0.24118584	0.83865833	0.4443601	3.08E-01	3.89E-01	0.51969602	0.41347231	0.54255594	0.25644444	0.01288358	0.48353055	NA	ns
GC3%	0.3422935	0.56046587	0.10130847	0.15093553	0.71855359	1.21E-01	1.39E-01	0.22352705	0.56071721	0.06191606	0.00652784	0.94251965	0.2346468	0.87067345	NA

La Figura 6.11 muestra las correlaciones para Z mientras que la Tabla 22 muestra los valores p para las correlaciones realizadas. Se observaron diferencias con las primeras dos proteínas. Mientras en N y GPC se observó correlación ligeramente positiva entre ENC y GC3%, para Z se observa una correlación altamente negativa, igual que la correlación ENC vs GRAVY. En contraposición se observa una alta correlación positiva entre ENC y U3% (valor p < 0,0001), la cual es baja y/o negativa en las otras proteínas. De todas formas, para la mayoría de las otras correlaciones se mantiene un patrón similar a aquellas de GPC y N. De hecho, se asemeja bastante más al de GPC que N en algunos casos. Notablemente, N y Z parecen tener elevada correlación positiva entre GC% y GC3%.

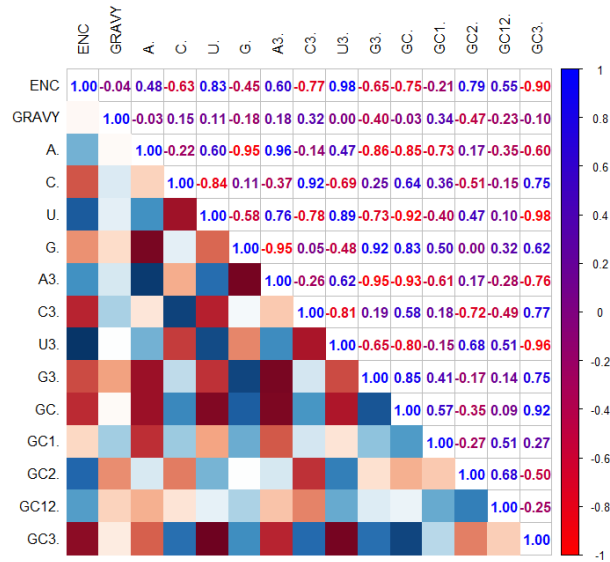


Figura 6.10. Correlación por método de Spearman entre componentes de la secuencia, ENC y GRAVY para Z de las cepas estudiadas. Azul oscuro indica correlación positiva mientras que rojo oscuro indica correlación negativa. Los valores más elevados indican una correlación más significativa.

Tabla 22. Valores p del análisis de correlación por rango de Spearman correspondiente a la Figura 6.10.

	ENC	GRAVY	A%	C%	U%	G%	A3%	C3%	U3%	G3%	GC%	GC1%	GC2%	GC12%	GC3%
ENC	NA	ns	ns	ns	*	ns	ns	*	****	ns	*	ns	*	ns	**
GRAVY	9.30E-01	NA	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
A%	2.30E-02	0.9531409	NA	ns	ns	***	***	ns	ns	ns	**	**	*	ns	ns
C%	9.74E-02	0.7178042	0.60033664	NA	**	ns	ns	**	ns	ns	ns	ns	ns	ns	*
U%	1.03E-02	0.7882334	0.115804	0.00892411	NA	ns	*	*	**	*	**	ns	ns	ns	****
G%	2.60E-01	0.6620879	0.00031293	0.79057132	1.32E-01	NA	***	ns	ns	**	*	ns	ns	ns	ns
A3%	1.16E-01	0.6726217	0.00020393	0.3734379	2.72E-02	0.00023252	NA	ns	ns	***	**	ns	ns	ns	*
C3%	2.68E-02	0.4355752	0.74191322	0.00119277	2.24E-02	0.90834319	0.52768215	NA	*	ns	ns	ns	*	ns	*
U3%	3.70E-02	1	0.23790444	0.05705467	2.92E-03	0.224585	0.09978427	0.01435795	NA	ns	*	ns	ns	ns	***
G3%	8.08E-02	0.3261187	0.0064143	0.54775241	4.18E-02	0.00135041	0.0003038	0.65934212	7.84E-02	NA	**	ns	ns	ns	*
GC%	3.22E-02	0.9506422	0.0072598	0.08563674	1.13E-03	0.01099493	0.0006740	0.13132512	1.67E-02	0.0072598	NA	ns	ns	ns	**
GC1%	6.21E-01	0.4029979	0.03844659	0.38064302	3.29E-01	0.21067224	0.10549762	0.66425416	7.29E-01	0.31835687	0.14402587	NA	ns	ns	ns
GC2%	1.89E-02	0.2424759	0.69090481	0.19299654	2.42E-01	1	0.68524499	0.04279039	6.10E-02	0.69090481	0.38976227	0.52223644	NA	ns	ns
GC12%	1.58E-01	0.5848883	0.39211146	0.72753765	8.10E-01	0.4426437	0.49853197	0.21344699	2.00E-01	0.73248712	0.83928477	0.19964244	0.06098133	NA	ns
GC3%	2.62E-03	0.8137319	0.11679034	0.03039227	3.00E-02	0.10268456	0.02825739	0.02533162	2.04E-04	0.03100851	0.00135041	0.50980378	0.20288445	0.55813551	NA

Finalmente, para la proteína L se observan ciertas particularidades en la Figura 6.12, aunque mantiene un patrón similar a las otras proteínas. GC12% y GRAVY tiene una correlación positiva, más que en GPC. Además, el ENC de L presenta mayor correlación con C3 en comparación con GPC, N y Z. El contenido total de adenina y guanidina tienen una correlación con el contenido GC en la tercera posición del codón, de la misma manera que GC y GC3 tienen una correlación positiva (en todos los casos con valores p menor a 0,05). Como se puede ver en la Tabla 23, donde se muestran todos los valores p calculados para las correlaciones hechas para la proteína L, que para el caso de ENC no hubo ninguna correlación con las otras variables que haya resultado estadísticamente significativa.

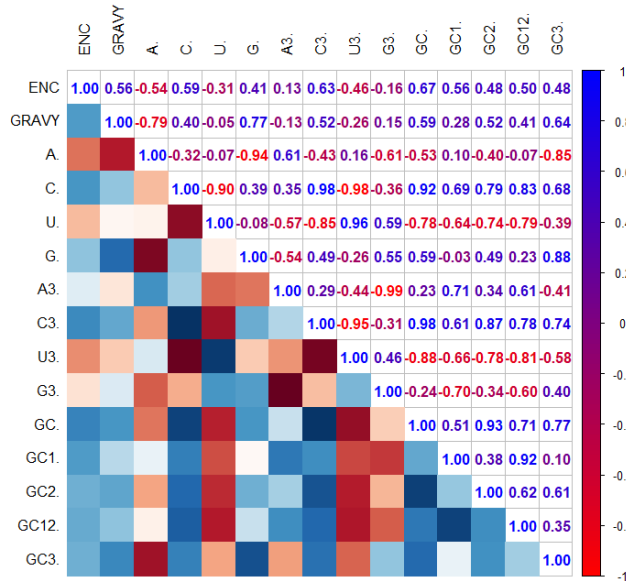


Figura 6.11. Correlación por método de Spearman entre componentes de la secuencia, ENC y otros para L de las cepas estudiadas. Azul oscuro indica correlación positiva mientras que rojo oscuro indica correlación negativa. Los valores más elevados indican una correlación más significativa.

Tabla 23. Valores p del análisis de correlación por rango de Spearman correspondiente a la Figura 6.11.

	ENC	GRAVY	A%	C%	U%	G%	A3%	C3%	U3%	G3%	GC%	GC1%	GC2%	GC12%	GC3%
ENC	NA	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
GRAVY	0.14633153	NA	*	ns	ns	*	ns	ns	ns	ns	ns	ns	ns	ns	ns
A%	0.1650407	0.01862891	NA	ns	ns	***	ns	ns	ns	ns	ns	ns	ns	ns	**
C%	0.12619899	0.32935526	0.44704254	NA	**	ns	ns	****	****	ns	**	ns	*	*	ns
U%	0.45279876	0.90977837	0.8753676	2.48E-03	NA	ns	ns	**	***	ns	*	ns	*	*	ns
G%	0.31670918	0.02507627	0.0093108	3.37E-01	0.84261571	NA	ns	ns	ns	ns	ns	ns	ns	ns	**
A3%	0.75583264	0.75440043	0.11122997	3.96E-01	0.13822916	0.17076032	NA	ns	ns	****	ns	*	ns	ns	ns
C3%	0.09116266	0.19109865	0.2849714	1.50E-05	0.00799252	0.21676597	4.79E-01	NA	***	ns	****	ns	**	*	*
U3%	0.24732326	0.5377582	0.70417993	1.50E-05	0.00018954	0.5377582	2.73E-01	3.01E-04	NA	ns	**	ns	*	*	ns
G3%	0.71107259	0.72022205	0.10848428	3.76E-01	0.12535215	0.15645757	5.49E-01	4.57E-01	2.55E-01	NA	ns	ns	ns	ns	ns
GC%	0.08308532	0.12619899	0.17650983	1.11E-03	0.02286733	0.12619899	5.88E-01	3.44E-05	4.10E-03	5.65E-01	NA	ns	***	ns	*
GC1%	0.14633153	0.50640757	0.8193151	5.99E-02	0.08445209	0.94355825	4.81E-02	1.06E-01	7.33E-02	5.18E-02	1.92E-01	NA	ns	**	ns
GC2%	0.22324513	0.18210949	0.32777526	2.06E-02	0.03431808	0.22006235	4.17E-01	5.02E-03	2.17E-02	4.14E-01	6.99E-04	0.34740952	NA	ns	ns
GC12%	0.21617063	0.31670918	0.86491956	1.14E-02	0.01954454	0.58784456	1.08E-01	2.22E-02	1.59E-02	1.14E-01	5.76E-02	0.00111082	0.10227853	NA	ns
GC3%	0.22651056	0.08583096	0.0078742	6.42E-02	0.33421464	0.00405573	3.10E-01	3.55E-02	1.32E-01	3.31E-01	2.51E-02	0.8193151	0.11025548	0.39620768	NA

El GRAVY no parece tener una correlación fuerte en las proteínas, con cierta excepción en L. Curiosamente se ven diferentes intensidades para GC1% y GC2% entre las proteínas. Entre las proteínas se observó una elevada correlación entre GC y GC3, con valores de correlación de 0,73, 0,92, 0,96 y 0,84 para L, Z, N y GPC, respectivamente. Por otro lado, la relación ENC vs GC3 para las proteínas GPC, N y L tuvo valores de 0,48, 0,39 y 0,64 respectivamente. Curiosamente, en Z se observó un valor de -0,90.

6.3 Discusión

En este capítulo se exploró el sesgo de uso de codones en JUNV, enfocándose en cada proteína y viendo si era posible identificar alguna diferencia entre cepas, haciendo mayor énfasis en la cepa P3441. Se exploraron diferentes índices y medidas utilizadas en varios estudios, tales como el ENC, CAI y RCDI. Dentro de los resultados, hubo diferencias significativas entre las proteínas y variación entre las cepas. Hubo diferencias esperables del porcentaje total de GC%, debido a las diferentes

funciones y composición de aminoácidos de cada proteína. Aun así, se observó para las proteínas L, GPC y N un menor porcentaje de contenido GC respecto a AT, lo cual se ha visto para otros virus con genoma de ARN (329-331). No obstante, los resultados de GC3% mostraron diferencias respecto a GC% para el caso de GPC. Inclusive si eran proteínas para una misma especie, y siendo únicamente dos segmentos genómicos, parecen haber diferencias entre las proteínas en el uso de codones.

El análisis del uso relativo de codones sinónimos permitió ver un patrón común entre las proteínas, aunque hubo diferencias notables en los valores de RSCU. Para prolina, treonina y alanina se observó en las proteínas de JUNV una mayor frecuencia de uso de codones terminados en adenina y frecuencia menor a codones sinónimos para codones terminados en guanina. Curiosamente, se observó para argenina un RSCU elevado para codones con adenina en primera posición pero RSCU casi nulo para codones con citosina en primera posición. Algunas diferencias fueron encontradas entre cepas, pero en su generalidad se observaron más similitudes que diferencias. C#1 y P3441 no se distinguieron significativamente de las otras cepas. Los resultados sugirieron una preferencia por codones con terminación con A y lo contrario para aquellos codones terminados con G particularmente para GPC.

Al parecer JUNV no tiene un sesgo de codones apreciable ya que para las cuatro proteínas fueron igual o mayor a 50. Este valor, respecto a 61, sugiere bajo sesgo, lo cual es un fenómeno que se ha observado para otros virus (332). Por ejemplo, los ENC calculados para los virus de la fiebre hemorrágica de Crimea-Congo y Hepatitis C fueron de 52,34 (329) y 52,62 (333), respectivamente. El mayor sesgo para L podría explicarse por su mayor tamaño. Se ha hipotetizado que el uso de codones raros (y por ende mayor sesgo) podría servir para reducir los niveles de expresión y la velocidad de plegamiento de proteínas largas hasta una velocidad óptima (334). De esta misma manera, podría también explicarse el menor sesgo de Z, la cual es una proteína pequeña.

El gráfico de neutralidad es una herramienta gráfica que permite analizar la relación entre el sesgo mutacional y la selección natural en el uso de codones (335). Acorde a los resultados, parecería que en JUNV la polimerasa L tiene un patrón de uso de codones explicado mayoritariamente por otros factores que la presión de mutaciones, similar a la polimerasa Z. GPC y N, en cambio, mostraron líneas de regresión de mayor pendiente que indican que ambas tienen un patrón de uso explicado parcialmente por presión mutacional. Aun así, parecería que para las cuatro proteínas de JUNV el sesgo de uso de codones se explica principalmente con factores como la selección natural u otras presiones de selección.

El RCDI agregó el concepto que las proteínas contaban con codones levemente desoptimizados respecto a *Homo sapiens*, con la proteína Z mostrando el nivel de desoptimización más elevado. Contrario a lo esperado, el cálculo del CAI, que da idea de la adaptación del uso de codones respecto a un organismo de referencia, dio valores que sugieren un grado de adaptación, particularmente en riñón, bazo y hígado con valores alrededor de 0,75. Esto se ha observado en otros virus que infectan humanos. Por ejemplo, se observó en el virus Hantaan (HTNV) un alto nivel de adaptación en *Homo sapiens* alcanzando valores de CAI de 0,793 (318). El mayor CAI para todos los casos correspondió para la proteína Z, lo cual contradice lo obtenido para el RCDI. No obstante, es posible que sea un resultado causado por el reducido largo de la secuencia para esta proteína, dónde hay menos proteínas para obtener un sesgo de uso de codones confiable.

Los tejidos y células de *Homo sapiens* mostraron diferencias en CAI para JUNV, aunque se mantuvo la relación entre proteínas. En las células inmunes se observó un mayor CAI en las células B y lo opuesto para las células dendríticas. Por su parte, en los tejidos estudiados, el mayor CAI obtenido correspondió a riñón, seguido de hígado y bazo, y el menor CAI calculado fue del tejido endotelial. Considerando que las células dendríticas son un blanco de JUNV (336, 337) y el efecto de la infección de JUNV en las células endoteliales (338-340), es curioso que en estos casos se haya obtenido los menores CAI, aunque esto podría sugerir una estrategia de replicación de JUNV en humano donde el virus encuentra beneficio en mantener una menor adaptación en el uso de codones para esas células. De hecho, se ha sugerido para Hantavirus que una menor adaptación permite una menor replicación sin inducir enfermedades severas en sus hospedantes naturales y así mantenerse diseminado (341). Es decir, un nivel de adaptación reducido podría favorecer la evasión a la respuesta inmune.

El estudio de correspondencia permitió visualizar la relación entre las proteínas, cepas e inclusive las proteínas con diferentes hospedantes. L mostró tener la mejor asociación con *Calomys musculinus*, lo cual puede explicarse por la necesidad de una proteína tan larga de plegarse óptimamente dentro de su reservorio natural y asegurar una traducción eficiente. En algunos virus se ha observado que la semejanza de patrones de uso de codones con aquellos sus hospedantes favorecen una traducción eficiente (328). Esto podría agregar a lo observado del ENC, dónde la polimerasa cuenta con el mayor sesgo de uso de codones entre las proteínas.

Las correlaciones realizadas sirven para determinar relaciones posibles entre variables. Por ejemplo, la relación entre GC y GC3 podrían indicar la presencia de presión de mutación de darse una correlación elevada (302), como ocurrió para las proteínas de JUNV. Por otro lado, la relación entre ENC y GC3 hace alusión a la posible presión de selección en el uso de codones. Para las proteínas GPC, N y L esta correlación fue menor a la relación GC y GC3, aunque una correlación positiva coincide con lo observado para diversos hantavirus cuya correlación fue positiva y alta (319). Curiosamente, en Z se observó una correlación altamente negativa entre ENC y GC3. No obstante, cabe aclarar que únicamente la correlación entre ENC y GC3 en Z fue estadísticamente significativa. Otras correlaciones que pueden agregar información son ENC vs GC y GRAVY vs GC3 y/o algún nucleótido en tercera posición. Para el caso de estudio de este capítulo, no se observaron correlaciones significativas, con excepción en Z para ENC vs GC con un valor negativo, lo cual agrega a una potencial influencia en la presión de mutaciones.

Acorde a lo observado, las proteínas parecen compartir generalidades en su sesgo, pero aun así mantienen un grado de diferencia entre ellas, tal como fue mencionado para los RSCU, posiciones en diferentes gráficos informativos entre ellos para el análisis correspondencia y en las correlaciones. Adicionalmente, es posible que haya un sesgo en las secuencias de GPC para JUNV hacia mayor contenido A y menor contenido de G en su tercera posición. No hubo suficiente distancia en ningún valor para P3441 o C#1 como para distinguirlas. Habría que realizar más estudios modificando codones para poder asociar patogenicidad al uso de codones. No obstante, se logró profundizar en la caracterización molecular de JUNV.

Discusión general y conclusiones

Una manera de expandir el conocimiento de JUNV ha sido el estudio comparativo entre cepas. Se ha realizado para determinar los diferentes fenotipos de infección para aislamientos de JUNV (127, 132) y para estudiar la evolución de JUNV en la cepa vacunal (342). Fue gracias a estos estudios que se pudo determinar los mayores factores de atenuación de C#1 (39). En nuestro laboratorio hemos trabajado tanto con la cepa vacunal (C#1) como la cepa patogénica P3441. Estas dos cepas han demostrado fenotipos diferentes en diversos modelos de infección, incluyendo la polarización M1/M2 en macrófagos según la patogenicidad de la cepa (53), y diferentes respuestas en células dendríticas plasmacitoides humanas (52). Hasta la fecha se han podido identificar sitios claves que explican la atenuación en la cepa C#1, pero aún queda por definir sitios que determinen la patogenicidad de P3441 y ayuden a profundizar la razón por las cuales se dan tales diferencias entre las cepas. De ahí la razón de este estudio.

En este trabajo se logró obtener una nueva secuencia codificante de un JUNV patogénico, agregando al limitado número de cepas secuenciadas completamente para este virus. Adicionalmente se logró el clonado completo de los genes de una cepa de JUNV. Sin embargo, no se pudieron obtener las secuencias de las regiones intergénicas de los segmentos S y L quedando para estudios futuros. Aun así, esta secuencia de fenotipo conocido y estudiado en diversos modelos de infección quedará disponible en base de datos. A diferencia de la vasta mayoría de los aislamientos presentes en base de datos, las secuencias de la cepa P3441 cuentan con la secuencia de la polimerasa completa.

Se cumplieron exitosamente la mayoría de los objetivos propuestos inicialmente. Una vez obtenida la secuencia codificante de la cepa P3441 de JUNV, se realizaron análisis comparativos con otras cepas de distinta patogenicidad y atenuación cuya secuencia completa es conocida. La mayoría de los trabajos realizados identifican meras diferencias entre cepas, intentando localizar diferencias antes de proceder con ensayos de mutagénesis (39, 343) o se utilizan para ensayos de seguimiento y expansión geográfica del área endémica (18, 344, 345). En este trabajo se realizó por primera vez una aproximación de todas las estructuras proteicas seguida del análisis del efecto de mutaciones puntuales comparándose entre C#1 y P3441, haciendo uso de herramientas modernas bioinformáticas como AlphaFold y servidores DynaMut y MAESTROweb. Avances en la informática han permitido que herramientas como AlphaFold no sólo puedan usarse con facilidad, sino que sus resultados tengan una importante capacidad predictiva (346-349). Adicionalmente, ha habido nuevas estructuras tridimensionales obtenidas por cristalografía para proteínas de mammarenavirus (aunque incompletas y con sus limitaciones) (235, 265, 267, 350, 351). En conjunto, esto ha hecho propicio la integración de estos avances para este trabajo.

Según el análisis bioinformático en este trabajo, las diferencias observadas entre las cepas de JUNV C#1 y P3441 no son producto de mutaciones que cambian radicalmente la estructura de las proteínas. No obstante, es probable que haya diferencias que puedan afectar parcialmente un dominio o motivo funcional, y potencialmente la interacción de la proteína con otras. Acorde al análisis bioinformático de este trabajo, el residuo 64 de la proteína Z es una de dichas diferencias relevantes. Coincidentemente, la diferencia V64G en esta proteína es un factor de atenuación menor al compararse C#1 con la cepa Romero (343), aunque en aquel estudio no se logró confirmar cómo ocurre dicha atenuación. Este trabajo predice que el residuo 64 afecta la estructura de la proteína Z dentro del dominio RING, el cual se sabe tiene un rol importante en la interacción con las otras

proteínas virales como consigo misma en otros mammarenavirus (137, 352, 353). De hecho, este resultado coincide con una interacción propuesta por ensayos de cristalografía entre Z y L, al menos en uno de los sitios teóricos de interacción (265), en la cual se propuso que V64 pudiese formar parte de esta interacción necesaria, y la ausencia de cadena lateral de G64 en C#1 implicaría una reducción en la eficiencia de la proteína de Z en la cepa vacunal de bloquear efectivamente la replicación viral y proceder con el ensamblado.

La proteína N es un factor esencial en la replicación viral e interviene en mayor medida que Z en la actividad evasora de la respuesta de interferón. Nuestro análisis predice que los residuos 268 y 546 es donde más se distinguen estructuralmente C#1 y P3441. Además, se encontraron aminoácidos únicos para P3441, como I123. No obstante, la diferencia en la posición 476 podría afectar la interacción entre Z y N sin necesariamente impactar en la estabilidad estructural de la proteína N, considerando la observación que la interacción entre Z y N para TCRV se ha propuesto de ocurrir entre los residuos 461 y 489 (354). Esta observación gana relevancia cuando se considera la proximidad evolutiva entre JUNV y TCRV.

Las glicoproteínas son claves en el contacto y mecanismos de ingreso del virión a la célula como en la posterior fusión y liberación del contenido viral en el citoplasma. En este caso, nuestro análisis predice que N75 en G1 y F180 de G2 distinguen estructuralmente las cepas C#1 y P3441. Coincidentemente, se ha demostrado que la atenuación de C#1 es debido a I427 en GPC (G2) (355). En este trabajo se agrega que dicha atenuación tiene un origen estructural que impacta en la función intrínseca de la proteína.

La proteína L permite la replicación del genoma como además la generación de ARNm para la traducción de otras proteínas. Aquí, el número de diferencias fue mayor lo cual fue esperado dado el mayor tamaño de la proteína e incluyeron residuos únicos para P3441 respecto a las otras cepas: D265, M611, K655, N679, V830, S885, D922, F987, G1452, H1664, S1751, V1909, V1940 y P1944. Estos resultados ameritan comenzar futuros estudios con proteínas quiméricas que intercambien segmentos más que con mutaciones puntuales. Los residuos L936, C1344, K1582, I1883 y V2040 fueron predichos como aquellos que podrían desestabilizar la estructura de la proteína en C#1 respecto P3441. Acorde a las estructuras generadas, L936 y C1344 caen en regiones externas sin estructura, mientras que se encuentra K1582 en lo que se ha considerado el “pulgar” sobre el cuál la polimerasa retiene el ARN acorde a estudios estructurales (258, 298) y se encuentra próximo a uno de los aminoácidos conservados en mammarenavirus que se ha propuesto en la interacción Z-L (356). Respecto a I1883 y V2040, ambos residuos son parte del dominio C-terminal sobre regiones sin estructura. De hecho, la región C-terminal de L parece establecer un dominio aparte cuya función se desconoce. Esta región no ha podido ser confirmada o estudiada por otros investigadores ante las limitaciones de la cristalografía de rayos X (258, 265, 298), algo que este trabajo intentó superar con la utilización de AlphaFold2.

En este trabajo no se pudo obtener información concluyente de factores de patogenicidad que definen el fenotipo de infección de P3441. Diferencias puntuales entre cepas varían sin un patrón identificable. Por ejemplo, la cepa MC2 posee una alta similitud con la cepa Romero para la proteína L pero comparte similitudes con C#1 en el dominio N terminal en la proteína N. Entre las cepas estudiadas, se observó un porcentaje de identidad aminoacídica para todas las proteínas mayor o igual al 96%. Dicho esto, de este trabajo surgen las siguientes observaciones relevantes:

- La proteína L mostró en promedio la mayor diferencia porcentual entre todas las cepas.
- Dentro de GPC, la mayoría de las diferencias entre cepas radica en G1 y casi ninguna en G2.
- Dentro de la proteína Z, la mayoría de las diferencias caen en el dominio N terminal.
- En la proteína N las diferencias parecen tener una dispersión relativamente uniforme a excepción de la ausencia de diferencias en su región entre dominios.

Claramente, para obtener conclusiones definidas se necesitará un mayor número de secuencias y así realizar estudios estadísticos con una población de datos apropiada. Desafortunadamente, existen más secuencias de las glicoproteínas y nucleoproteínas (en su mayoría parciales), pero faltan de la polimerasa L, lo cual este trabajo ha ayudado a remediar.

Los servidores utilizados para determinar el impacto puntual de las diferencias observadas también arrojaron resultados contradictorios, demostrando las limitaciones de los análisis bioinformáticos aun cuando se empleen múltiples herramientas bioinformáticas. También debe considerarse que estas predicciones están realizadas en base a modelos que son a su vez predicciones estructurales. Adicionalmente, es convención considerar valores mayores a 1 y menores a -1 de $\Delta\Delta G$ para asegurar que una mutación teórica sea desestabilizante o estabilizante respectivamente, lo cual no fue el caso para la mayoría de las predicciones en este trabajo. Por lo tanto, es imperioso obtener verificación por medio de estudios estructurales no predictivos incluyendo experimentos de mutagénesis dirigida seguidos de estudios funcionales de las proteínas o de virus quiméricos generados por virología reversa. No obstante, cómo primera medida de potenciales sitios a estudiar, este trabajo logró conseguir un número trabajable de diferencias puntuales entre cepas para explorar y que podría replicarse para otras especies. De hecho, sería de interés determinar si se repiten patrones similares de diferencia en, por ejemplo, LASV donde hay un mayor número de secuencias en base de datos que para JUNV. También es de interés explorar otras propiedades de las proteínas que varíen con la secuencia aminoacídica, como cambios en propiedades químicas de superficie o el impacto de mutaciones sobre las propiedades de las cadenas laterales de los residuos vecinos.

Hay varias herramientas bioinformáticas disponibles para analizar proteínas, como aquellas utilizadas en este trabajo y algunas de las cuales se intentaron utilizar pero cuyos resultados fueron poco concluyentes en el mejor de los casos o directamente cuestionables. Una de ellas fue HADDOCK. Este programa permite predecir y analizar interacciones biomoleculares, particularmente entre proteínas (357, 358). HADDOCK requiere información de potenciales sitios de interacción, y se realizaron pruebas con la estructura de JUNV Z y lo observado para LASV Z con su interacción con eIF4e (136). Los resultados, aunque interesantes, no fueron concluyentes y se debe buscar un refinamiento con dinámica molecular para los modelos tridimensionales de las proteínas obtenidas a fin de obtener conclusiones razonables. NAMD es otro programa que se utilizó para agregar información en este trabajo. El programa sirve para el estudio de dinámica molecular y potencialmente poder refinar los modelos conseguidos con AlphaFold (359). No obstante, este software requería una capacidad de procesamiento muy superior a lo disponible y de las pocas repeticiones logradas no hubo resultados concluyentes para reportar en este trabajo. Se requerirá mayor tiempo, especialización y dedicación con estos programas para lograr conseguir información relevante.

Finalmente, se realizó un estudio de sesgo de uso de codones para JUNV para ver si se podían ver diferencias entre las diferentes cepas y profundizar en la caracterización molecular de P3441, además de explorar este fenómeno biológico en JUNV. Esta es la primera vez que se realiza un estudio de

este tipo y profundidad en JUNV. Se ha realizado previamente en SABV y CHAPV (321) con resultados algo similares. En ese trabajo, no obstante, no se consideró la posibilidad de que ocurrieran diferencias significativas en el sesgo de uso de codones entre las proteínas, la cual probablemente se debió al número reducido de secuencias en base de datos para cada virus.

En el caso de JUNV, se logró determinar diferencias significativas entre las proteínas. La proteína L mostró tener el mayor sesgo de uso de codones, opuesto a la proteína Z, inclusive si el sesgo de codones de JUNV es bastante reducido respecto a otras especies. Los valores de ENC fueron en promedio por encima de 50 para las cuatro proteínas, mientras que otros virus que producen fiebres hemorrágicas dentro del orden de los *Bunyavirales* tienen ENC por debajo de 42 para la proteína de nucleocápside (319). Los resultados para JUNV coinciden con lo observado para SABV y CHPV que tuvieron valores promedio de ENC de $50,144 \pm 2,07$ y $46,2375 \pm 6,038$ respectivamente (321). Además se puede llegar a la conclusión que JUNV tiene una influencia mayoritariamente de presión de selección y otros factores respecto a la presión mutacional, opuesto a lo observado para SABV y CHPV (321), aunque cómo se observó de aquel trabajo no se utilizaron varios aislamientos. Se requieren más trabajos para profundizar el impacto de los patrones de uso de codones en la patogenicidad de JUNV.

Siguiendo con lo alcanzado en este trabajo, se lograron clonar las proteínas de la cepa P3441 en vectores de expresión eucariota. Se demostró la expresión y localización celular de cada una de las proteínas recombinantes. Pese a estar incluido en el Plan de Tesis, se decidió no realizar ensayos de ELISA con sueros de pacientes convalecientes debido a los bajos niveles de transfección observados. En su lugar, se verificó la antigenicidad de las proteínas recombinantes utilizando anticuerpos específicos para las proteínas Z, N y GPC restando la caracterización de L para cuando se logre generar un antisuero específico. Desafortunadamente, aunque se contaba con antisueros de Z, los anticuerpos secundarios para revelar este antisuero coincidían en fluorescencia con EGFP, y las alternativas no funcionaron o no se pudieron conseguir para este objetivo.

Ya se han generado replicones de JUNV para la cepa Romero y C#1 (360). Este trabajo agrega nuevas herramientas moleculares disponibles para otra cepa patogénica de JUNV. El hecho que las proteínas separadas se encuentran clonadas con fluoróforo permite el estudio individual de las proteínas, cómo fue explorado para Z y N. Esto resultará útil en estudios futuros de mutagénesis dirigida para precisar las diferencias entre cepas identificadas en este trabajo. El clonado (y el logro) más relevante de este trabajo fue aquella para la proteína L. Este plásmido servirá como base para estudios con otras finalidades, como por ejemplo subclonar en un vector de expresión de proteínas para su purificación y posterior generación de anticuerpos. Queda profundizar en la caracterización del producto de este plásmido, quedando para trabajos futuros determinar si conserva actividad de síntesis de ARN.

Una consideración para este trabajo es el modelo de estudio utilizado. Aunque la línea celular A549 resultó útil para la realización de diversos estudios empleando microscopía confocal, resultó poseer un porcentaje de eficiencia en la transfección relativamente menor comparada con otras como la HEK293 (resultados no mostrados). El motivo por el cual no se utilizó esa línea fue porque resultó menos útil en estudios de imágenes. Sin embargo, para futuros estudios que requieran una eficiencia de transfección relativamente alta se contemplará emplear la línea HEK293 u otras con eficiencia equivalente.

En conclusión, este trabajo fue exitoso en el estudio del modelo patogénico de P3441 principalmente por la generación de vectores que expresan las proteínas de P3441. Estos plásmidos podrán emplearse en nuevas herramientas futuras como un antisuero contra la proteína L, realizar análisis comparativos con proteínas de otras cepas como C#1 y utilizarse en una gran variedad de ensayos para profundizar el estudio del fenotipo de infección de P3441. Este trabajo sienta las bases para futuros análisis, tanto en el campo computacional teórico como experimental.

Materiales y Métodos

Células

La línea celular BHK-21, clon 15, derivada de riñón de hámster (*Mesocricetus auratus*), la línea celular A549 derivadas de cáncer con fenotipo de epitelio alveolar humano tipo II y la línea celular Vero-76 derivada de fibroblastos de riñón de mono verde africano (*Chlorocebus sp*) originalmente provenientes del *American Type Culture Collection* (ATCC), fueron obtenidas de diversos medios locales. Las células se cultivaron en estufa a 37°C con 5% CO₂ utilizando medio esencial mínimo (MEM) (Gibco, Thermo Fisher Scientific) suplementado con 10% de suero fetal bovino (SFB)(Internegocios).

Virus

La cepa Candid 1 del virus Junín fue gentilmente cedida por el Dr. Mario Lozano (UNQ) y proviene de una muestra de preparado vacunal amplificado en línea celular Vero-76. La cepa P3441 del virus Junín fue provista gentilmente por la Dra. A. Ambrosio del Instituto Nacional de Enfermedades Virales Humanas “Dr. Julio I. Maiztegui”, ANLIS, Pergamino, Argentina. La cepa originalmente fue aislada de un paciente de Fiebre Hemorrágica Argentina.

Para generar *stocks* virales, se utilizaron monocapas celulares de células BHK crecidas desde el día anterior en una confluencia del 70-80% en frascos T-175 e infectadas a una multiplicidad de infección de 1 (MOI=1). Las células fueron incubadas con 3 ml de una solución conteniendo el virus durante 1 hora a 37°C. Luego el inóculo fue retirado y se agregaron 25ml de MEM con 5% SFB. El cultivo fue observado diariamente usando un microscopio invertido a luz visible y se mantuvo 3 o 4 días, dependiendo del estado de la monocapa y la cantidad de efecto citopático observado. Posteriormente se cosecharon los sobrenadantes, se verificó el título, y se alicuotó en crioviales mantenidos a -80°C.

Ensayo de unidades formadoras de placas

Para los ensayos de titulación por unidades formadoras de placas (UFP), se sembraron células Vero en placas de 24 o 48 pocillos a una confluencia del 70%. Las muestras se diluyeron al décimo de forma seriada y se utilizaron 100µl de inóculo para infectar las monocapas por 1 hora a 37°C. Luego, se descartó el inóculo y reemplazó con 0,75ml de medio compuesto por MEM, metilcelulosa 0,8% y 2% SFB. Las células se incubaron a 37°C durante 6 días, y luego se fijaron con una solución de paraformaldehído (PFA) al 10% (p/v) durante 15 min y se tiñeron con una solución conteniendo cristal violeta 1% (p/v) en etanol 96° durante 30 min a temperatura ambiente (62) previo al conteo de placas. El título viral, definido como el número de placas o calvas por ml de solución viral (UFP/ml), fue calculado según la siguiente fórmula:

$$\text{Título} \left(\frac{\text{UFP}}{\text{ml}} \right) = \frac{\text{N}^{\circ} \text{ de placas contadas}}{\text{dilución del inóculo} \times \text{volumen del inóculo}}$$

Transcripción reversa y reacción en cadena de la polimerasa

Se obtuvo ARN viral cosechando células BHK infectadas con P3441 a 3 dpi en un frasco T75 a 100% de confluencia en 1ml de *TransZol Up* (TransGen) según las indicaciones del fabricante. La calidad y cantidad de ARN fue verificada por la medición de absorbancia con el espectrofotómetro Nanodrop

ND-1000 (*NanoDrop Technologies*). La transcripción reversa (RT) se llevó a cabo utilizando la transcriptasa de MMLV Transcripta (Productos Bio-Lógicos) según el protocolo del fabricante. La reacción se realizó con oligonucleótidos aleatorios (Thermo Fisher Scientific), siguiendo las temperaturas y tiempos recomendados por el fabricante.

Las reacciones de PCR se realizaron con 0,25 µg de ADNc y 1 U de polimerasa *Taq Pegasus* (Productos Bio-Lógicos) según el protocolo del fabricante. Los oligonucleótidos utilizados se muestran en la Tabla 26 y 27 del anexo para el clonado y secuenciación respectivamente. Los productos de PCR fueron verificados en geles de 1,5% de agarosa tenidos con bromuro de etidio (0,25µg/ml) y las bandas del tamaño esperado fueron cortadas y el ADN purificado usando kit comercial *PURO Gel Extraction* (Productos Bio-Lógicos) según el protocolo del fabricante.

Secuenciación

Para la secuenciación de fragmentos amplificados por RT-PCR del genoma viral, se eligió el servicio de secuenciación por electroforesis capilar basado en el método de Sanger de la empresa *Macrogen Inc.* (República de Corea). Los *primers* utilizados se diseñaron usando el servidor PrimerBlast (www.ncbi.nlm.nih.gov/tools/primer-blast/) a partir de las secuencias en base de datos con número de acceso FJ969422 y AY819707 para el segmento S y L de JUNV C#1, respectivamente. Los *primers* diseñados fueron recopilados y utilizados en PCR *In Silico* usando el programa UGene y el servidor Benchling. Las temperaturas de fusión fueron verificadas también usando el programa Snapgene Viewer.

Desfosfatación

Para evitar la recirculación de plásmidos linealizados, se trataron 2µg de ADN resuspendido después de precipitación alcohólica en 40µl de 10mM Tris-HCl (pH 8.0) para utilizar 0,5U de la enzima fosfatasa alcalina de intestino de carnero (*Promega*) siguiendo las instrucciones de fabricante. Al final del ensayo, se purificó el ADN por extracción con fenol-cloroformo (ver más adelante).

Clonado

Los productos de la RT-PCR con la secuencia codificante de las proteínas virales fueron clonados en el vector pGEM-T (*pGEM-T Easy Vector*, Promega) usando ligasa T4 (New England Biolabs) según las instrucciones del fabricante. El producto de la ligación fue transformado por electroporación en células electro-competentes *E. coli* TOP10 y las células fueron colocadas en placas de Petri plásticas (*Guangzhou Jet Bio-Filtration Co., Ltd.*, República de Corea) con medio Luria (LB) sólido conteniendo ampicilina 100 µg/ml y X-GAL 20mg/ml (Productos Bio-Lógicos) según el protocolo del fabricante. Las placas fueron incubadas aproximadamente 16 horas a 37°C en estufa. Las colonias crecidas con coloración blanca fueron repicadas en una nueva placa (que se dejó en estufa 16 horas) y se utilizaron para realizar un ensayo de *colony-PCR*. Este ensayo consistió en levantar la colonia con la punta de un escarbadiante estéril en condiciones de esterilidad para ligeramente picar en una nueva placa y el remanente sumergirse en 10µl de agua estéril en tubo eppendorf de 0,5ml. El tubo se llevó a baño con agua en ebullición, incubándose por 5 min. Posteriormente, se utilizó 1µl de la solución como molde para PCR siguiendo instrucciones de fabricante, la polimerasa y el método de visualización por electroforesis en gel de agarosa como fue descrito previamente.

Las colonias positivas en la nueva placa fueron repicadas en 5ml de LB líquido con 100µg/ml de ampicilina y se incubaron en un agitador por 16 horas. Crecido el cultivo, se procedió con una

minipreparación de plásmidos por el método de lisis alcalina (*minipreps*, ver inciso siguiente). Una vez obtenidos los plásmidos y purificados por extracción con fenol-cloroformo, aproximadamente 2 µg de plásmido fueron digeridos con las enzimas respectivas acorde a los sitios de digestión de los extremos agregados con los cebadores de los productos de PCR clonados (Tabla 26 y 28). La mezcla de reacción fue preparada acorde a las enzimas utilizadas siguiendo las instrucciones del fabricante. La digestión enzimática fue monitoreada sembrando alícuotas en gel de agarosa con bromuro de etidio. Los productos de digestión de tamaño esperado fueron escindidos del gel de agarosa y purificados usando kit comercial *PURO Gel Extraction* (Productos Bio-Lógicos) según el protocolo del fabricante. Posteriormente, se subclonaron en plásmidos previamente digeridos y purificados utilizando la ligasa T4 (*New England Biolabs*) para unir el vector y los insertos correspondientes. El producto de la reacción se transformó en células electro-competentes y las células fueron sembradas en cajas de Petri 90mm con LB sólido conteniendo kanamicina (50µg/ml).

Minipreparación de plásmidos por el método de lisis alcalina

Para la obtención de plásmidos, se crecieron cultivos de LB con antibiótico correspondiente por 16 horas con agitación (180rpm) a 37°C. Se inocularon con repique de una colonia o 25µl de solución glicerol 50% de un cultivo previo. De ser la primera vez de realizarse la *miniprep*, se alicuotaron 0,5ml de cultivo y se le agregaron 0,5ml de glicerol puro para guardarse a -20°C.

En la *miniprep*, se centrifugaron 1,5 ml de cultivo a 14000rpm por 30 segundos. Se descartó el sobrenadante y se repitió el proceso dos veces. Posteriormente se resuspendió el *pellet* concentrado de bacterias transformadas en 200 µl de solución mM glucosa, 25 mM Tris-HCl (pH 8.0), 10 mM EDTA (pH 8.0). Después se agregaron 400 µl de solución 0,2 M NaOH y 1% p/v SDS preparada en el momento y se mezcló por inversión suavemente hasta alcanzar la clarificación del medio. Finalmente se agregó 300µl de solución 3M Acetato de potasio y 120,45 g/l ácido acético glacial. Se centrifugó 14000rpm durante 5 minutos y se recuperó el sobrenadante, repitiéndose de seguir teniendo precipitado blanquecino en sobrenadante claro. Seguidamente se agregaron 4 µl de RNAsa A (Productos Bio-Lógicos) y se incubó el tubo en estufa a 37°C por 60 minutos. Transcurrido el período de incubación con RNAsa se agregaron 0,8 volúmenes de Isopropanol y se incubó 30 min a -20°C, después de lo cual se centrifugó a 14000rpm durante 20 minutos a 4 °C. Se eliminó el sobrenadante y el *pellet* se resuspendió en 50 µL de agua bidestilada a 55°C.

Purificación de ADN por extracción con fenol-cloroformo

Posterior a la precipitación alcohólica se utilizó 1 µl de la suspensión obtenida para una corrida electroforética en gel de agarosa 0,8% para verificar la presencia y cantidad de plásmido obtenido. De ser un resultado favorable, se procedió con la purificación con fenol-cloroformo para tener ADN sin remanentes de proteínas o lipopolisacáridos. Para esto se llevó la suspensión de plásmido a 200 µl de 25 mM Tris-HCl y 10 mM EDTA (pH 8.0), se agregó un volumen de fenol:cloroformo:alcohol isoamílico (25:24:1) y se mezcló intensamente por 30 segundos en vortex. Seguidamente se centrifugó a 14000rpm 5 minutos TA y se recuperó la fase acuosa (superior). Se agregaron 180 µl de la solución Tris-HCl y EDTA al remanente de la fase orgánica, se repitió el mezclado y centrifugación previa para recuperar un segundo volumen de fase acuosa para llevar con la primera. Se agregó un volumen de cloroformo:alcohol isoamílico (24:1), se mezcló intensamente en vortex por 30 segundos y centrifugó de la misma manera a los pasos anteriores. Finalmente se recuperó la fase acuosa, se agregó 3M de acetato de sodio hasta una concentración final de 0,75M y se realizó una resuspensión alcohólica idéntica a la descripta para la *miniprep*.

Transfección

Células A549 crecidas en cubreobjetos de vidrios de 12mm insertos en placa de 24 pocillos al 90% de confluencia fueron transfectadas con 2µg de ADN plasmídico utilizando 1,5µl de Lipofectamina 3000 (Thermo Fisher Scientific) siguiendo las instrucciones del fabricante. A las 24hs se observó la placa en un microscopio invertido Olympus CK2 con luz visible. En el caso de transfecciones con plásmidos conteniendo fluoróforos, fueron monitoreados utilizando un microscopio de fluorescencia invertido *Eclipse* (Nikon).

Ensayos de inmunofluorescencia

Células A549 crecidas en cubreobjetos de vidrios de 12mm insertos en placa de 24 pocillos al 90% de confluencia fueron cosechadas a los tiempos postransfección o posinfección correspondientes, y se fijaron con una solución de PFA 4% en PBS 30 min a TA. En ciertos casos, se permeabilizaron junto a la fijación usando PFA 4% y Tritón X-100 0,5% en PBS 1X por 30 min a TA. Después de tres lavados con PBS, los vidrios se guardaron con 1ml de PBS a 4°C hasta el día de su utilización. Para los ensayos de inmunofluorescencia, las monocapas fueron bloqueadas con 30µl de solución de Tween-20 0,05% y SFB 5% en PBS por 30 min. Luego, la solución de bloqueo se reemplazó con 30µl de solución de anticuerpo primario y se incubó a TA por 2 horas. Se lavaron los vidrios un mínimo de 4 veces con PBS antes de agregar 30µl de solución del anticuerpo secundario correspondiente y se incubaron por 1 hora en oscuridad a TA. Después de dos lavados con PBS, se incubaron por 15 min con 30µl de DAPI o Verde Metilo (colorantes nucleares) acorde a los fluoróforos que se deseaban observar y el microscopio confocal a utilizarse. Se lavaron los vidrios dos veces con PBS y un último lavado en agua destilada antes de montar en un portaobjeto de vidrio con 4µl de mowiol (*MilliporeSigma*). Los anticuerpos y diluciones utilizadas se encuentran en la Tabla 24 y 25 del anexo para anticuerpos primarios y secundarios respectivamente.

Microscopía confocal

Para la visualización de los ensayos de inmunofluorescencia se emplearon los siguientes equipos: Leica TCS SP5, con el apoyo de la Plataforma de Microscopía Avanzada, Facultad de Exactas, UNLP o Olympus FV1000, con el apoyo del Laboratorio de Análisis de Imágenes, Facultad de Ciencias Veterinarias, UNLP. Generalmente, se utilizó el objetivo 60X con aceite de inmersión.

Secuencias

Las secuencias de diferentes cepas de JUNV fueron obtenidas de la base de datos de nucleótidos de NCBI (361). Las secuencias de las proteínas de mayor expresión en tejidos y células de *Homo sapiens* fueron identificados por medio de la página web *Human Protein Atlas* (<https://www.proteinatlas.org/>) para después ser descargados de base de datos de NCBI por medio de códigos caseros con el lenguaje de programación Python (Python Software Foundation, <https://www.python.org/>) usando el programa Pycharm Community Edition 2023.3.3 (Jetbrains s.r.o, <https://www.jetbrains.com/pycharm/>).

Programas y servicios informáticos

Los análisis de imágenes se realizaron utilizando el programa Fiji (362) derivado del programa libre ImageJ (NIH, EU). La visualización de archivos PDB para el estudio de estructuras tridimensionales se realizó utilizando los programas VMD (363) y PyMol (*The PyMOL Molecular Graphics System*, Version 2.0 Schrödinger). Imágenes y visualización adicionales de estructuras proteicas fueron realizadas

utilizando Chimera 1.17.1, software desarrollado por *Resource for Biocomputing, Visualization and Informatics* de la Universidad de California, San Francisco, con el apoyo de NIH P41-GM103311 (364). Los alineamientos de secuencias nucleotídicas y aminoácidas y PCR *In Silico* se realizaron con el programa Ugene (365). El diseño de oligonucleótidos y estrategias de clonado se realizó empleando el programa *Snapgene Viewer* (www.snapgene.com). Para la construcción de árboles filogenéticos se utilizó el programa MegaX (366) y el servidor NGPhyML (367).

Las estructuras tridimensionales de las proteínas virales se obtuvieron por medio del servidor ColabFold (181) utilizando el programa de predicción de estructuras AlphaFold2 (368). La excepción fue la proteína L donde se utilizó el servidor SWISS-MODEL (369, 370).

El refinado de los modelos obtenidos se realizó por medio de ReFOLD3, un servidor web que incluye un conjunto de programas y métodos de refinamiento de estructuras tridimensionales de proteínas por medio de restricciones graduales basadas en calidad local predicha y contacto entre residuos. [<https://www.reading.ac.uk/bioinf/ReFOLD/> (182)].

Los modelos fueron verificados por medio de PROCHECK (371) y VERIFY3D (372) utilizando el servidor web SAVESv6.0 del *Institute for Genomics and Proteomics* (IGP) dentro del Departamento de Ciencia de Energía en conjunto con la Universidad de California, Los Angeles (UCLA) (<https://saves.mbi.ucla.edu/>).

Las imágenes de alineamientos presentados fueron obtenidas con el servidor ESPript 3.0 accesible en <https://esprpt.ibcp.fr/ESPript/cgi-bin/ESPript.cgi> (373).

El efecto de mutaciones en las estructuras fue predicho usando los servidores MAESTROweb (374) y DynaMut (184).

El texto fue generado utilizando Office Professional Plus 2013 (Microsoft) y las figuras fueron generadas utilizando Inkscape 1.0.2-1 accesible en <https://www.inkscape.org>. Para el manejo de las citas bibliográficas se utilizó el programa EndNote X9. Para ciertas imágenes se utilizó el programa Photoshop CC 2017 (Adobe).

En el estudio de codones, se utilizó el servidor CAIcal (375) para los cálculos de composición, RSCU y RCDI. Para los estudios de correspondencia y correlación, el cálculo de ENC, el gráfico ENC-GC3 y generación de tabla de uso de codones en células y tejidos de *Homo sapiens*, se escribieron *scripts* en el lenguaje de programación R R 4.3.2 (376), editándose por medio del programa RStudio build 369 23.12.0 (Posit Software, PBC). Se usaron los paquetes como FactoMiner (377) para ensayos de correlación y corrplot (378) para correlaciones.

Análisis Estadístico

Para realizar el análisis estadístico se utilizó el programa GraphPad Prism 8. Para determinar si los datos obtenidos tenían distribución normal se aplicó la prueba de Shapiro-Wilk. Para comparaciones individuales se usó la prueba t de student no apareada o el test de U de Mann-Whitney, de acuerdo con la normalidad o no de la distribución de los datos. El ANOVA seguido del test de Tukey o el de Kruskal-Wallis seguido del test de Dunn fueron usados para comparaciones múltiples acorde a datos que se demostraran distribución normal o lo contrario, respectivamente. En todos los análisis estadísticos, la significancia fue aceptada en $p < 0,05$.

Anexo Materiales y métodos

Tabla 24. Anticuerpos primarios utilizados para inmunofluorescencia.

Anticuerpo (Marcador)	Especie	Dilución de uso	Código Entrez-Gene
PML (E-11):sc-377390	Ratón	1/50	
JUNV N (Clon QB06-AE05)	Ratón	1/50	NR-2576
JUNV N (Clon SA02-BG12)	Ratón	1/50	NR-49274
JUNV N (Clon MA03-BE06)	Ratón	1/50	NR-41860
JUNV N (Clon NA05-AG12)	Ratón	1/50	NR-48834
JUNV GPC (Clon GB03-BE08)	Ratón	1/50	NR-43227
JUNV GPC	Ratón	1/50	NR-43776
JUNV GPC	Ratón	1/50	NR-48833
JUNV GPC	Ratón	1/50	NR-43775
GM130 [EP892Y]	Conejo	1/100	Ab52649
IRF-3 (D6I4C)	Conejo	1/100	#11904 3661
IRF-1 (D5E4)	Conejo	1/100	#8478 3659
MAVS (D5A9E)	Conejo	1/100	#24930 57506
JUNV Z	Conejo	1/10	

Tabla 25. Anticuerpos secundarios utilizados para inmunofluorescencia.

Fluoróforo	Especie/Anti-Especie	Dilución de uso	Código y/o N°Catálogo	Fabricante
AF647	Cabra/Ratón	1/2000	#A-21241; RRID = AB_2535810	Invitrogen
FITC	Cabra/Ratón	1/100	A01318	Genscript
FITC	Cabra/Conejo	1/200	A10158	Genscript

Tabla 26. Oligonucleótidos para clonado de vectores y proteínas recombinantes.

Nombre	Secuencia	Sitio de restricción
JUNV L KpnI fwd	CCTCTGTAGACGGTACCCT	GGTAC ν C
JUNV L KpnI rev	CGGGTCTAGAGGGTACCG	GGTAC ν C
JUNV BstEII fwd	CAGGCGATCAGGTGTCTC	G ν GTNACC
JUNV BstEII rev	GAAGGGGAAGATGCGGATC	G ν GTNACC
JUNVPL+BamH1fwd	GGATCCGCCAATGTCTTTGATCTCGACTTTGTG	G ν GATCC
JP3441LC-tNheI	GGTTATCGCTAGCCTACCA	G ν CTAGC
JUNVNPNheI1fwd	GGAAGCTAGCATGGCACACTCCAAGGAG	G ν CTAGC
JUNVNPKpnIrev	GAGAGGTACCAGCAGTGCATAGGCTGCC	GGTAC ν C
JUNVNP+KpnI ν P3r	GGTACCTGCATAGGCTGCCTTCG	GGTAC ν C
JP3441GPCf	ATGGGGCAGTTCATTAGCTTCAT	
JP3441GPCr	GGTCTTAGTGCTCTACGC	

Tabla 27. Oligonucleótidos utilizados para la secuenciación.

Designación	Nombre de compra	Secuencia
1	JUNV_L_AY819707_1_fw	CGCACAGTGGATCCTAGGCG
2	JUNV_L_AY819707_770_rev	AGTCAAGGCAATTGGAGCACT
3	JUNV_L_AY819707_651_fw	AGGACAAAGTCCTCTTGCTGAA
4	JUNV_L_AY819707_1541_rev	GTCAACATCGTGTCTCAAATCCT
5	JUNV_L_AY819707_1441_fw	ACAAGGTAAAATCTGTGAGTTCAGC
6	JUNV_L_AY819707_2248_rev	GTTGCCTCAGGCTTCATTGG
7	JUNV_L_AY819707_2112_fw	GGCTTTCTCTCACAAATCCACAA
8	JUNV_L_AY819707_2891_rev	TGTGTCTGCTGCTCTACACAA
9	JUNV_L_AY819707_2708_fw	AATAGCTCCAAACGGAGTCTCTC
10	JUNV_L_AY819707_3528_rev	ATGACCACTCCAAATGGGGAC
11	JUNV_L_AY819707_3379_fw	CAACGACCTTATGAAGATGCC
12	JUNV_L_AY819707_4178_rev	TGATCCAGATGTGCTCAGGTC
13	JUNV_L_AY819707_4076_fw	GATGTCCTGCGGTGCTTCAA
14	JUNV_L_AY819707_4876_rev	CCCCCTCTAAGTGGGAGTCT
15	JUNV_L_AY819707_4673_fw	AGGCCTGTAGGGATTGCTTT
16	JUNV_L_AY819707_5445_rev	CAGGTGAGAAGTCAAGGTGCT
17	JUNV_L_AY819707_5925_fw	CAGCTGGTCATTTCTCACAC
18	JUNV_L_AY819707_6103_rev	CAAGCCCTTTCCGTGTGTAAC
19	JUNV_L_AY819707_5999_fw	GCATCATCATGCGTTTTGAAAAGA
20	JUNV_L_AY819707_6679_rev	ACGTTGGTTCCTGTGGTTGA
21	JUNV_L_AY819707_6408_fw	ACCCCAACAAAGAGGGATTCA
22	JUNV_L_AY819707_7114_rev	CGCCTAGGATCCACTGTGC
23	JUNV_S_FJ969442_1_fw	CACAGTGGATCCTAGGCGAT
24	JUNV_S_FJ969442_754_rev	AGGCCAACTGGTAGGTTGTG
25	JUNV_S_FJ969442_618_fw	ACACCTCCAAGACTGGTGTGTC
26	JUNV_S_FJ969442_1339_rev	AGTTTTACCCTGCCTGTCCG
27	JUNV_S_FJ969442_1176_fw	TCAACCACACACTTTCAGGACA
28	JUNV_S_FJ969442_1898_rev	ACATCTGAACAAGCCAGGCA
29	JUNV_S_FJ969442_1762_fw	GGTTCCTGAGGCAACTT
30	JUNV_S_FJ969442_2586_rev	GGCTGGTGCATCCATTCTTG
31	JUNV_S_FJ969442_2400_fw	ATCGCCAGAAAGGCAAAGTT
32	JUNV_S_FJ969442_3056_rev	GGAGATCTGGGGAGGGATGA
33	JUNV_S_FJ969442_2724_fw	AAGCCCAAGAGAGGTTAAAGCC
34	JUNV_S_FJ969442_3389_rev	AGTGTCTAACACTTTCAGTACGCA

Tabla 28. Enzimas de Restricción utilizadas para el clonado de proteínas virales

Nombre	Fabricante (Nº de Catálogo)	Buffer de reacción
KpnI-HF	NEB (R3142S)	rCutSmart
BstEII-HF	NEB (R3162S)	rCutSmart
NheI-HF	NEB (R3131S)	rCutSmart
BamHI-HF	NEB (R3136S)	rCutSmart

Anexo RSCU

Tabla RSCU GPC. La frecuencia relativa de uso de codón sinónimo (RSCU) por aminoácido (AA) y cepa de JUNV en GPC.

Codón	AA	C#1	P3441	XJ13	P35032	MC2	Romero	Ledesma	Espindola
TTT	F	0,8	0,839	0,774	0,903	0,839	0,774	0,71	0,839
TTC	F	1,2	1,161	1,226	1,097	1,161	1,226	1,29	1,161
TTA	L	1,138	1,034	1,158	0,915	1,138	1,138	1,241	1,138
TTG	L	1,655	1,862	1,895	2,034	1,241	2,069	1,966	2,069
CTT	L	0,724	0,724	0,842	0,915	0,724	0,828	0,931	0,828
CTC	L	0,724	0,621	0,632	0,508	0,724	0,621	0,517	0,621
CTA	L	0,517	0,931	0,737	0,915	0,724	0,621	0,724	0,621
CTG	L	1,241	0,828	0,737	0,712	1,448	0,724	0,621	0,724
ATT	I	1,222	1,154	1,111	1,038	1,111	1,222	1,111	1,111
ATC	I	1	1,038	1	1,038	1,222	0,889	1,111	1
ATA	I	0,778	0,808	0,889	0,923	0,667	0,889	0,778	0,889
GTT	V	1,091	1,273	1,143	1,143	0,952	1,524	1,273	0,952
GTC	V	0,909	0,727	0,762	0,762	0,952	0,381	0,545	0,762
GTA	V	1,091	1,091	1,143	0,952	1,143	1,143	1,273	0,952
GTG	V	0,909	0,909	0,952	1,143	0,952	0,952	0,909	1,333
TCT	S	1	1,371	1,167	1,333	1,5	1,167	1,5	1,459
TCC	S	1,5	1,2	1,333	1	1,167	1,333	1	1,135
TCA	S	1	1,029	1	1,167	0,833	1	1	0,973
TCG	S	0,167	0,171	0,167	0,167	0,167	0,167	0,167	0,162
AGT	S	1	1,029	1	1	1	1	1	0,973
AGC	S	1,333	1,2	1,333	1,333	1,333	1,333	1,333	1,297
CCT	P	1,636	1,143	1,455	1,333	1,143	1,333	1,143	1,2
CCC	P	0,364	0,762	0,364	0,571	1,143	0,571	0,762	0,6
CCA	P	2	1,905	2	2,095	1,524	2,095	1,905	2
CCG	P	0	0,19	0,182	0	0,19	0	0,19	0,2
ACT	T	1,12	1,333	1,185	0,889	1,037	1,185	1,077	1,185
ACC	T	0,96	0,889	1,037	1,185	1,037	0,889	1,077	0,889
ACA	T	1,92	1,63	1,778	1,778	1,926	1,926	1,846	1,926
ACG	T	0	0,148	0	0,148	0	0	0	0
GCT	A	1,391	1,091	1,455	1,273	1,273	1,143	1,273	1,273
GCC	A	0,348	0,909	0,545	0,909	0,545	0,952	0,727	0,727
GCA	A	1,913	1,455	1,636	1,455	1,818	1,524	1,636	1,636
GCG	A	0,348	0,545	0,364	0,364	0,364	0,381	0,364	0,364
TAT	Y	1,143	1	1,143	1,143	1,067	1,143	1,143	1,143
TAC	Y	0,857	1	0,857	0,857	0,933	0,857	0,857	0,857
CAT	H	0,889	1,111	0,889	1,333	1,059	1,111	1,111	1,333
CAC	H	1,111	0,889	1,111	0,667	0,941	0,889	0,889	0,667

CAA	Q	1,111	1,176	1,111	1,412	1,125	1,176	1,294	1,412
CAG	Q	0,889	0,824	0,889	0,588	0,875	0,824	0,706	0,588
AAT	N	1,062	1,03	1,062	0,938	1,161	1,125	0,938	1,062
AAC	N	0,938	0,97	0,938	1,062	0,839	0,875	1,062	0,938
AAA	K	1,125	1,125	1,125	1,067	1,03	1,062	1,125	1,125
AAG	K	0,875	0,875	0,875	0,933	0,97	0,938	0,875	0,875
GAT	D	0,87	0,783	0,87	0,957	0,667	0,957	0,87	0,87
GAC	D	1,13	1,217	1,13	1,043	1,333	1,043	1,13	1,13
GAA	E	1,5	1,294	1,529	1,444	1,556	1,444	1,294	1,412
GAG	E	0,5	0,706	0,471	0,556	0,444	0,556	0,706	0,588
TGT	C	1,059	1,176	1,059	1,294	1,176	1,176	0,941	1,176
TGC	C	0,941	0,824	0,941	0,706	0,824	0,824	1,059	0,824
CGT	R	1	1	1	0,9	0,667	0,947	1	0,947
CGC	R	0	0	0	0	0,333	0	0	0
CGA	R	0	0	0	0	0	0	0	0
CGG	R	0	0	0	0	0	0	0	0
AGA	R	2	2	2	2,4	2	2,842	2	2,526
AGG	R	3	3	3	2,7	3	2,211	3	2,526
GGT	G	1,429	1,185	1,481	1,077	1,185	1,077	1,037	1,077
GGC	G	1,143	1,481	1,185	1,538	1,481	1,692	1,63	1,538
GGA	G	1,143	0,889	0,889	0,769	0,889	0,769	0,741	0,769
GGG	G	0,286	0,444	0,444	0,615	0,444	0,462	0,593	0,615

RSCU = 0

RSCU = 1

RSCU >= 2

Tabla RSCU N. La frecuencia relativa de uso de codón sinónimo (RSCU) por aminoácido (AA) y cepa de JUNV en N.

Codón	AA	C#1	P3441	XJ13	P35032	MC2	Romero	Ledesma	Espindola
TTT	F	1,294	1,059	0,941	0,941	1,368	0,941	0,941	0,941
TTC	F	0,706	0,941	1,059	1,059	0,632	1,059	1,059	1,059
TTA	L	0,554	0,727	0,738	0,727	0,462	0,818	0,636	0,545
TTG	L	1,569	1,455	1,477	1,545	1,662	1,273	1,545	1,545
CTT	L	1,2	1,364	1,292	1,636	1,385	1,364	1,364	1,545
CTC	L	1,015	0,909	0,923	0,636	0,923	0,909	0,909	0,727
CTA	L	0,462	0,545	0,554	0,636	0,462	0,364	0,636	0,545
CTG	L	1,2	1	1,015	0,818	1,108	1,273	0,909	1,091
ATT	I	1,444	1,5	1,222	1,111	1,5	1,333	1,333	1,385
ATC	I	0,778	0,643	0,889	1	0,692	0,778	0,778	0,808
ATA	I	0,778	0,857	0,889	0,889	0,808	0,889	0,889	0,808
GTT	V	1,889	1,714	1,889	1,556	1,946	1,889	2	1,838

GTC	V	0,667	0,8	0,667	1	0,649	0,667	0,556	0,649
GTA	V	0,222	0,114	0,111	0	0,216	0,111	0,222	0,216
GTG	V	1,222	1,371	1,333	1,444	1,189	1,333	1,222	1,297
TCT	S	1,463	1,61	1,463	1,364	1,256	1,429	1,286	1,714
TCC	S	1,317	1,171	1,171	1,364	1,256	1,286	1,429	1,143
TCA	S	1,171	1,317	1,463	1,227	1,256	1,286	1,143	1,286
TCG	S	0,146	0	0	0,136	0,14	0,143	0,143	0
AGT	S	1,171	1,171	1,024	1,091	1,116	1,143	1	1,143
AGC	S	0,732	0,732	0,878	0,818	0,977	0,714	1	0,714
CCT	P	1,538	1,44	1,538	1,5	1,44	1,333	1,44	1,5
CCC	P	0,308	0,32	0,308	0,333	0,32	0,5	0,32	0,333
CCA	P	1,692	1,76	1,692	1,833	1,6	1,833	1,76	1,833
CCG	P	0,462	0,48	0,462	0,333	0,64	0,333	0,48	0,333
ACT	T	1,793	1,333	1,733	1,733	1,517	1,655	1,6	1,517
ACC	T	0,69	0,933	0,667	0,8	0,966	0,69	0,8	0,828
ACA	T	1,517	1,6	1,6	1,467	1,379	1,517	1,467	1,655
ACG	T	0	0,133	0	0	0,138	0,138	0,133	0
GCT	A	1,241	1,241	1,241	1,143	1,29	1,2	1,241	1,241
GCC	A	0,69	0,69	0,69	0,857	0,645	0,8	0,69	0,69
GCA	A	1,793	1,931	1,931	1,857	1,677	1,867	2,069	1,931
GCG	A	0,276	0,138	0,138	0,143	0,387	0,133	0	0,138
TAT	Y	1,231	1,077	1,077	1,333	1,091	0,923	0,923	1,231
TAC	Y	0,769	0,923	0,923	0,667	0,909	1,077	1,077	0,769
CAT	H	0,6	0,6	0,6	0,727	0,727	0,6	0,6	0,6
CAC	H	1,4	1,4	1,4	1,273	1,273	1,4	1,4	1,4
CAA	Q	0,968	1,032	0,903	0,839	0,933	1,097	1,097	1,032
CAG	Q	1,032	0,968	1,097	1,161	1,067	0,903	0,903	0,968
AAT	N	0,846	0,923	0,846	0,72	0,88	0,923	1,04	0,889
AAC	N	1,154	1,077	1,154	1,28	1,12	1,077	0,96	1,111
AAA	K	1	0,976	0,977	0,927	0,95	0,952	1,073	1,024
AAG	K	1	1,024	1,023	1,073	1,05	1,048	0,927	0,976
GAT	D	1,053	1	0,919	1	1,053	0,947	0,895	0,947
GAC	D	0,947	1	1,081	1	0,947	1,053	1,105	1,053
GAA	E	0,938	1,097	1,062	0,968	0,966	1,032	1,032	1,032
GAG	E	1,062	0,903	0,938	1,032	1,034	0,968	0,968	0,968
TGT	C	0,286	0,571	0,571	0,571	0,286	0,571	0,571	0,571
TGC	C	1,714	1,429	1,429	1,429	1,714	1,429	1,429	1,429
CGT	R	0	0	0	0	0	0	0	0
CGC	R	0	0	0	0	0	0	0	0
CGA	R	0,353	0,343	0,364	0,343	0,333	0,353	0,343	0,343
CGG	R	0,176	0,171	0,182	0,171	0,333	0,176	0,171	0,171

AGA	R	3,529	3,771	3,636	3,6	3,5	3,529	3,429	3,6
AGG	R	1,941	1,714	1,818	1,886	1,833	1,941	2,057	1,886
GGT	G	1,268	1,171	1,268	1,171	1,048	1,171	0,976	1,268
GGC	G	0,78	0,878	0,78	0,878	0,952	0,878	1,073	0,78
GGA	G	1,171	1,073	1,073	1,171	1,048	1,268	1,171	1,171
GGG	G	0,78	0,878	0,878	0,78	0,952	0,683	0,78	0,78

RSCU = 0
RSCU = 1
RSCU >= 2

Tabla RSCU L. La frecuencia relativa de uso de codón sinónimo (RSCU) por aminoácido (AA) y cepa de JUNV en L.

Codón	AA	C#1	P3441	XJ13	P35032	MC2	Romero	Ledesma	Espindola
TTT	F	1,345	1,333	1,356	1,19	1,345	1,345	1,362	1,265
TTC	F	0,655	0,667	0,644	0,81	0,655	0,655	0,638	0,735
TTA	L	1,256	1,213	1,23	1,2	1,2	1,2	1,23	1,23
TTG	L	1,646	1,755	1,64	1,907	1,789	1,789	1,813	1,835
CTT	L	0,866	0,888	0,863	0,857	0,829	0,829	0,863	0,863
CTC	L	0,477	0,455	0,475	0,45	0,48	0,48	0,453	0,453
CTA	L	0,953	0,823	0,993	0,771	0,916	0,916	0,777	0,906
CTG	L	0,801	0,866	0,799	0,814	0,785	0,785	0,863	0,712
ATT	I	1,312	1,252	1,326	1,181	1,333	1,333	1,427	1,282
ATC	I	0,867	0,898	0,837	0,969	0,81	0,81	0,726	0,847
ATA	I	0,82	0,85	0,837	0,85	0,857	0,857	0,847	0,87
GTT	V	1,565	1,634	1,54	1,8	1,61	1,61	1,659	1,63
GTC	V	0,745	0,707	0,745	0,575	0,732	0,732	0,707	0,667
GTA	V	0,547	0,537	0,571	0,45	0,463	0,463	0,512	0,494
GTG	V	1,143	1,122	1,143	1,175	1,195	1,195	1,122	1,21
TCT	S	1,493	1,478	1,515	1,401	1,58	1,588	1,567	1,419
TCC	S	0,806	0,828	0,802	1,005	0,79	0,794	0,798	0,916
TCA	S	1,433	1,448	1,426	1,462	1,434	1,441	1,478	1,448
TCG	S	0,149	0,148	0,149	0,152	0,146	0,147	0,118	0,148
AGT	S	1,015	1,123	1,01	1,036	1,112	1,088	1,123	1,094
AGC	S	1,104	0,975	1,099	0,944	0,937	0,941	0,916	0,975
CCT	P	1,333	1,455	1,354	1,6	1,415	1,415	1,375	1,562
CCC	P	0,848	0,788	0,862	0,615	0,8	0,8	0,812	0,688
CCA	P	1,636	1,576	1,6	1,538	1,662	1,662	1,5	1,562
CCG	P	0,182	0,182	0,185	0,246	0,123	0,123	0,312	0,188
ACT	T	0,825	0,884	0,866	0,97	0,8	0,8	0,898	0,968
ACC	T	0,866	0,884	0,825	0,808	0,926	0,926	0,898	0,8
ACA	T	2,227	2,063	2,227	2,061	2,105	2,105	2,082	2,021

ACG	T	0,082	0,168	0,082	0,162	0,168	0,168	0,122	0,211
GCT	A	1,463	1,59	1,481	1,463	1,5	1,5	1,418	1,59
GCC	A	0,683	0,564	0,691	0,78	0,7	0,7	0,759	0,615
GCA	A	1,561	1,59	1,58	1,61	1,7	1,7	1,671	1,641
GCG	A	0,293	0,256	0,247	0,146	0,1	0,1	0,152	0,154
TAT	Y	1,284	1,169	1,273	1,284	1,273	1,273	1,273	1,292
TAC	Y	0,716	0,831	0,727	0,716	0,727	0,727	0,727	0,708
CAT	H	1,2	1,19	1,22	1,333	1,268	1,286	1,286	1,273
CAC	H	0,8	0,81	0,78	0,667	0,732	0,714	0,714	0,727
CAA	Q	1,128	1,205	1,128	1,256	1,154	1,154	1,215	1,282
CAG	Q	0,872	0,795	0,872	0,744	0,846	0,846	0,785	0,718
AAT	N	1,26	1,301	1,273	1,314	1,288	1,288	1,231	1,314
AAC	N	0,74	0,699	0,727	0,686	0,712	0,712	0,769	0,686
AAA	K	1,117	1,053	1,098	1,084	1,107	1,107	1,101	1,078
AAG	K	0,883	0,947	0,902	0,916	0,893	0,893	0,899	0,922
GAT	D	1,191	1,232	1,182	1,191	1,191	1,191	1,235	1,17
GAC	D	0,809	0,768	0,818	0,809	0,809	0,809	0,765	0,83
GAA	E	1,091	1,098	1,091	1,123	1,104	1,104	1,143	1,052
GAG	E	0,909	0,902	0,909	0,877	0,896	0,896	0,857	0,948
TGT	C	1,429	1,439	1,429	1,544	1,474	1,483	1,509	1,474
TGC	C	0,571	0,561	0,571	0,456	0,526	0,517	0,491	0,526
CGT	R	0,153	0,154	0,151	0,155	0,157	0,105	0,158	0,153
CGC	R	0,102	0,103	0,101	0,103	0,052	0,053	0,053	0,102
CGA	R	0,407	0,359	0,403	0,362	0,47	0,474	0,421	0,407
CGG	R	0,153	0,205	0,151	0,155	0,209	0,211	0,211	0,102
AGA	R	3,254	3,385	3,277	3,103	3,391	3,421	3,368	3,305
AGG	R	1,932	1,795	1,916	2,121	1,722	1,737	1,789	1,932
GGT	G	1,109	1,02	1,149	1,01	1,072	1,072	1,072	1,2
GGC	G	0,673	0,653	0,634	0,687	0,66	0,66	0,577	0,56
GGA	G	1,109	1,102	1,109	1,091	1,072	1,072	1,113	1,04
GGG	G	1,109	1,224	1,109	1,212	1,196	1,196	1,237	1,2

RSCU = 0
RSCU = 1
RSCU >= 2

Tabla RSCU Z. La frecuencia relativa de uso de codón sinónimo (RSCU) por aminoácido (AA) y cepa de JUNV en Z.

Codón	AA	C#1	P3441	XJ13	P35032	MC2	Romero	Ledesma	Espindola
TTT	F	2	2	2	1	1	1	2	1
TTC	F	0	0	0	1	1	1	0	1
TTA	L	1,714	1,714	1,714	0,857	1,714	1,714	1,714	1,714

TTG	L	0,857	0,857	0,857	0,857	0	0	0	0
CTT	L	0,857	0,857	0,857	0,857	0,857	0,857	0,857	0,857
CTC	L	0,857	0,857	0,857	0,857	0,857	0,857	0,857	0,857
CTA	L	0,857	1,714	0,857	1,714	1,714	1,714	1,714	1,714
CTG	L	0,857	0	0,857	0,857	0,857	0,857	0,857	0,857
ATT	I	0	0	0	0	0	0	0	0,75
ATC	I	2	2	2	2	2	2	2	1,5
ATA	I	1	1	1	1	1	1	1	0,75
GTT	V	0	0,8	0	0,8	0,8	0,8	0,8	0
GTC	V	0	0,8	0	0,8	0,8	0,8	0,8	1
GTA	V	2,667	1,6	2,667	0,8	0	0	0,8	1
GTG	V	1,333	0,8	1,333	1,6	2,4	2,4	1,6	2
TCT	S	0,857	1,714	0,857	0	0,857	0,857	0	0
TCC	S	1,714	0,857	1,714	1,714	0,857	0,857	2	2,571
TCA	S	1,714	1,714	1,714	1,714	1,714	1,714	1	1,714
TCG	S	0	0	0	0	0,857	0,857	1	0
AGT	S	0,857	0,857	0,857	0,857	0,857	0,857	1	0,857
AGC	S	0,857	0,857	0,857	1,714	0,857	0,857	1	0,857
CCT	P	0	0	0	0,444	0,444	0,444	0	0,444
CCC	P	0,889	0,889	0,889	0,889	0,444	0,444	1,2	0,889
CCA	P	2,667	2,667	2,667	2,667	2,667	2,667	2,4	2,667
CCG	P	0,444	0,444	0,444	0	0,444	0,444	0,4	0
ACT	T	0	0	0	0,571	0	0	0	0
ACC	T	1,714	1,714	1,714	1,143	1,714	1,714	1,714	1,714
ACA	T	2,286	2,286	2,286	2,286	2,286	2,286	2,286	2,286
ACG	T	0	0	0	0	0	0	0	0
GCT	A	1,143	1,333	1,143	0,571	1,333	1,333	2	0,667
GCC	A	1,143	0,667	1,143	1,714	0,667	0,667	0	1,333
GCA	A	1,714	2	1,714	1,714	2	2	2	2
GCG	A	0	0	0	0	0	0	0	0
TAT	Y	1,333	0,667	1,333	0,667	0,667	0,667	0,667	0,667
TAC	Y	0,667	1,333	0,667	1,333	1,333	1,333	1,333	1,333
CAT	H	0,667	0,667	0,667	0,667	0,667	0,667	0,667	0,667
CAC	H	1,333	1,333	1,333	1,333	1,333	1,333	1,333	1,333
CAA	Q	0	0	0	0	0	0	0	0
CAG	Q	2	2	2	2	2	2	2	2
AAT	N	1	1	1	1,429	1	1	1	1
AAC	N	1	1	1	0,571	1	1	1	1
AAA	K	0	0	0	0	0	0	0	0
AAG	K	2	2	2	2	2	2	2	2
GAT	D	1,5	1,5	1,5	1,5	0,5	0,5	1	1,5

GAC	D	0,5	0,5	0,5	0,5	1,5	1,5	1	0,5
GAA	E	1	1	1	1	1	1	1	2
GAG	E	1	1	1	1	1	1	1	0
TGT	C	0,667	0,667	0,667	1,111	0,444	0,444	0,444	0,667
TGC	C	1,333	1,333	1,333	0,889	1,556	1,556	1,556	1,333
CGT	R	0	0	0	0	0	0	0	0
CGC	R	0	0	0	0	0	0	0	0
CGA	R	0	0	0	0	0	0	0	0
CGG	R	0	0	0	0	0	0	0	0
AGA	R	2	2	2	3	1	1	2	2
AGG	R	4	4	4	3	5	5	4	4
GGT	G	2	1,333	2	1,333	1,333	1,333	1,333	1,333
GGC	G	1	1,333	1	1,333	1,333	1,333	1,333	1,333
GGA	G	0	0	0	0	0	0	0	0
GGG	G	1	1,333	1	1,333	1,333	1,333	1,333	1,333

RSCU = 0
RSCU = 1
RSCU >= 2

Bibliografía

1. Gomez RM, Jaquenod de Giusti C, Sanchez Vallduvi MM, Frik J, Ferrer MF, Schattner M. Junin virus. A XXI century update. *Microbes and infection / Institut Pasteur*. 2011;13(4):303-11.
2. Arribalzaga RA. Una nueva enfermedad epidémica a germen desconocido. Hipertemia nefrotóxica, leucopénica y enantemática. *Día méd.*. 1955(27):1204-10.
3. Parodi AS, Greenway DJ, Rugiero HR, Frigerio M, De La Barrera JM, Mettler N, et al. [Concerning the epidemic outbreak in Junin]. *Dia Med*. 1958;30(62):2300-1.
4. Piroosky I, Zuccarini J, Molinelli EA. Virosis hemorrágica del noroeste bonaerense: endemo-epidémica, febril, enantemática y leucopenica. *Virosis hemorrágica del noroeste bonaerense: endemo-epidémica, febril, enantemática y leucopenica*1959. p. 114-.
5. Greenway DJ, Rugiero HR, Parodi AS, Frigerio M, Rivero E, De La Barrera JM, et al. Epidemic hemorrhagic fever in Argentina. *Pub Health Rep*. 1959;74:1011-4.
6. Mettler N, Buckley SM, Casals J. Propagation of Junin virus, the etiological agent of Argentinian hemorrhagic fever, in HeLa cell cultures. *Proc Soc Exp Biol Med*. 1961;107:684-8.
7. Rapp F, Buckley SM. Studies with the etiologic agent of Argentinian epidemic hemorrhagic fever (Junin virus). *Am J Pathol*. 1962;40(1):63-75.
8. Sabattini MS, Gonzalez LE. [Direct identification of Junin's virus in infected wild rodents in nature]. *Rev Soc Argent Biol*. 1967;43(5):252-60.
9. Carballal G, Videla CM, Merani MS. Epidemiology of Argentine hemorrhagic fever. *Eur J Epidemiol*. 1988;4(2):259-74.
10. Videla C, Kajon A, Carballal G, Weissenbacher M. *Calomys callidus* as a potential Junin virus reservoir. *J Med Virol*. 1989;27(3):238-43.
11. Rowe WP, Pugh WE, Webb PA, Peters CJ. Serological relationship of the Tacaribe complex of viruses to lymphocytic choriomeningitis virus. *J Virol*. 1970;5(3):289-92.
12. Murphy FA, Webb PA, Johnson KM, Whitfield SG, Chappell WA. Arenoviruses in Vero cells: ultrastructural studies. *J Virol*. 1970;6(4):507-18.
13. Añon MC, Grau O, Segovia ZM, Franzefernandez MT. RNA composition of Junin virus. *J Virol*. 1976;18(3):833-8.
14. Berria MI, Gutman Frugone LF, Girola R, Barrera Oro JG. [Immunological studies with Junin virus. I. Formation of antibodies in guinea pigs inoculated with live viruses]. *Medicina (B Aires)*. 1967;27(2):93-8.
15. Lascano EF, Berria MI. Ultrastructure of Junin virus in mouse whole brain and mouse brain tissue cultures. *J Virol*. 1974;14(4):965-74.
16. De Martinez Segovia ZM, De Mitri MI. Junin virus structural proteins. *J Virol*. 1977;21(2):579-83.
17. Maiztegui J, Feuillade M, Briggiler A. Progressive extension of the endemic area and changing incidence of Argentine Hemorrhagic Fever. *Med Microbiol Immunol*. 1986;175(2-3):149-52.
18. Calderon GE, Provencal MC, Martin ML, Brito Hoyos DM, Garcia JB, Gonzalez-Iltig RE, et al. [Co-circulation of Junin virus and other mammarenaviruses in a geographical area without confirmed cases of Argentine Hemorrhagic Fever]. *Medicina (B Aires)*. 2022;82(3):344-50.
19. Gómez RM, Jaquenod de Giusti C, Sanchez Vallduvi MM, Frik J, Ferrer MF, Schattner M. Junín virus. A XXI century update. *Microbes and infection*. 2011;13(4):303-11.
20. Mills JN, Ellis BA, McKee KJ, Calderon GE, Maiztegui JI, Nelson GO, et al. A longitudinal study of Junin virus activity in the rodent reservoir of Argentine hemorrhagic fever. *American Journal of Tropical Medicine & Hygiene*. 1992;47(6):749-63.

21. Mills JN, Ellis BA, McKee KT, Jr., Calderon GE, Maiztegui JI, Nelson GO, et al. A longitudinal study of Junin virus activity in the rodent reservoir of Argentine hemorrhagic fever. *Am J Trop Med Hyg.* 1992;47(6):749-63.
22. Charrel RN, de Lamballerie X. Zoonotic aspects of arenavirus infections. *Veterinary microbiology.* 2010;140(3-4):213-20.
23. Enria DA, Mills JN, Bausch DG, Shieh W-J, Peters CJ, editors. CHAPTER 68 – Arenavirus Infections 2011.
24. Marta RF, Montero VS, Molinas FC. Systemic disorders in Argentine haemorrhagic fever. *Bull Inst Pasteur.* 1998;96:115-24.
25. Harrison LH, Halsey NA, McKee KT, Jr., Peters CJ, Barrera Oro JG, Briggiler AM, et al. Clinical case definitions for Argentine hemorrhagic fever. *Clin Infect Dis.* 1999;28(5):1091-4.
26. Harrison Lee H, Halsey Neal A, McKee JKelly T, Peters Clarence J, Oro Julio GB, Briggiler Ana M, et al. Clinical Case Definitions for Argentine Hemorrhagic Fever. *Clinical Infectious Diseases.* 1999;28(5):1091-4.
27. Lozano ME, Enria D, Maiztegui JI, Grau O, Romanowski V. Rapid diagnosis of argentine hemorrhagic fever by reverse transcriptase PCR-based assay. *Journal of Clinical Microbiology.* 1995;33(5):1327-32.
28. Levis SC, Saavedra MC, Ceccoli C, Falcoff E, Feuillade MR, Enria DA, et al. Endogenous interferon in Argentine hemorrhagic fever. *J Infect Dis.* 1984;149(3):428-33.
29. Levis SC, Saavedra MC, Ceccoli C, Feuillade MR, Enria DA, Maiztegui JI, et al. Correlation between endogenous interferon and the clinical evolution of patients with Argentine hemorrhagic fever. *J Interferon Res.* 1985;5(3):383-9.
30. Enria DA, Briggiler AM, Sánchez Z. Treatment of Argentine hemorrhagic fever. *Antiviral Research.* 2008;78(1):132-9.
31. Maiztegui JI, Fernandez NJ, de Damilano AJ. Efficacy of immune plasma in treatment of Argentine haemorrhagic fever and association between treatment and a late neurological syndrome. *Lancet.* 1979;2(8154):1216-7.
32. Enria DA, Barrera Oro JG. Junin virus vaccines. *Current topics in microbiology and immunology.* 2002;263:239-61.
33. Enria DA, Ambrosio AM, Briggiler AM, Feuillade MR, Crivelli E, Study Group on Argentine Hemorrhagic Fever V. [Candid#1 vaccine against Argentine hemorrhagic fever produced in Argentina. Immunogenicity and safety]. *Medicina (B Aires).* 2010;70(3):215-22.
34. Kenyon RH, Green DE, Maiztegui JI, Peters CJ. Viral strain dependent differences in experimental Argentine hemorrhagic fever (Junin virus) infection of guinea pigs. *Intervirology.* 1988;29(3):133-43.
35. Yun NE, Linde NS, Dziuba N, Zacks MA, Smith JN, Smith JK, et al. Pathogenesis of XJ and Romero strains of Junin virus in two strains of guinea pigs. *Am J Trop Med Hyg.* 2008;79(2):275-82.
36. Dejean CB, Ayerra BL, Teyssie AR. Interferon response in the guinea pig infected with Junin virus. *J Med Virol.* 1987;23(1):83-91.
37. Webb PA, Justines G, Johnson KM. Infection of wild and laboratory animals with Machupo and Latino viruses. *Bull World Health Organ.* 1975;52(4-6):493-9.
38. Campetella OE, Galassi NV, Sanjuan N, Barrios HA. Susceptible adult murine model for Junin virus. *J Med Virol.* 1988;26(4):443-51.
39. Albarino CG, Bird BH, Chakrabarti AK, Dodd KA, Flint M, Bergeron E, et al. The major determinant of attenuation in mice of the Candid1 vaccine for Argentine hemorrhagic fever is located in the G2 glycoprotein transmembrane domain. *J Virol.* 2011;85(19):10404-8.
40. Kolokoltsova OA, Yun NE, Poussard AL, Smith JK, Smith JN, Salazar M, et al. Mice lacking alpha/beta and gamma interferon receptors are susceptible to junin virus infection. *J Virol.* 2010;84(24):13063-7.

41. McKee KT, Jr., Mahlandt BG, Maiztegui JI, Eddy GA, Peters CJ. Experimental Argentine hemorrhagic fever in rhesus macaques: viral strain-dependent clinical response. *J Infect Dis.* 1985;152(1):218-21.
42. Kenyon RH, McKee KT, Jr., Zack PM, Rippey MK, Vogel AP, York C, et al. Aerosol infection of rhesus macaques with Junin virus. *Intervirology.* 1992;33(1):23-31.
43. Weissenbacher MC, Calello MA, Colillas OJ, Rondinone SN, Frigerio MJ. Argentine hemorrhagic fever: A primate model. *Intervirology.* 1979;11(6):363-5.
44. Bushar G, Sagripanti JL. Method for improving accuracy of virus titration: standardization of plaque assay for Junin virus. *J Virol Methods.* 1990;30(1):99-107.
45. Lascano EF, Berria MI. [Electron microscopy of primary cultures of mouse fibroblasts inoculated with Junin virus]. *Medicina (B Aires).* 1969;29(6):487-93.
46. Carballal G, Rodriguez M, Frigerio MJ, Vasquez C. [Editorial: Ultrastructure of bone marrow and peripheral blood of guinea pigs infected with Junin virus]. *Medicina (B Aires).* 1975;35(4):437-8.
47. Andrews BS, Theofilopoulos AN, Peters CJ, Loskutoff DJ, Brandt WE, Dixon FJ. Replication of dengue and junin viruses in cultured rabbit and human endothelial cells. *Infect Immun.* 1978;20(3):776-81.
48. Gomez RM, Pozner RG, Lazzari MA, D'Atri LP, Negrotto S, Chudzinski-Tavassi AM, et al. Endothelial cell function alteration after Junin virus infection. *Thromb Haemost.* 2003;90(2):326-33.
49. Berria MI, Lascano EF. Astrocyte differentiation induced by Junin virus in rat brain cell cultures. *Acta Neuropathol.* 1985;66(3):233-8.
50. Pozner RG, Collado S, Jaquenod de Giusti C, Ure AE, Biedma ME, Romanowski V, et al. Astrocyte response to Junin virus infection. *Neurosci Lett.* 2008;445(1):31-5.
51. Pozner RG, Ure AE, Jaquenod de Giusti C, D'Atri LP, Italiano JE, Torres O, et al. Junin virus infection of human hematopoietic progenitors impairs in vitro proplatelet formation and platelet release via a bystander effect involving type I IFN signaling. *PLoS Pathog.* 2010;6(4):e1000847.
52. Negrotto S, Mena HA, Ure AE, Jaquenod De Giusti C, Bollati-Fogolin M, Vermeulen EM, et al. Human Plasmacytoid Dendritic Cells Elicited Different Responses after Infection with Pathogenic and Nonpathogenic Junin Virus Strains. *J Virol.* 2015;89(14):7409-13.
53. Ferrer MF, Thomas P, Lopez Ortiz AO, Errasti AE, Charo N, Romanowski V, et al. Junin Virus Triggers Macrophage Activation and Modulates Polarization According to Viral Strain Pathogenicity. *Front Immunol.* 2019;10:2499.
54. Hickerson BT, Westover JB, Wang Z, Lee YM, Gowen BB. Guinea Pig Transferrin Receptor 1 Mediates Cellular Entry of Junin Virus and Other Pathogenic New World Arenaviruses. *J Virol.* 2020;94(4).
55. Radoshitzky SR, Abraham J, Spiropoulou CF, Kuhn JH, Nguyen D, Li W, et al. Transferrin receptor 1 is a cellular receptor for New World haemorrhagic fever arenaviruses. *Nature.* 2007;446(7131):92-6.
56. Kolokoltsova OA, Grant AM, Huang C, Smith JK, Poussard AL, Tian B, et al. RIG-I enhanced interferon independent apoptosis upon Junin virus infection. *PLoS One.* 2014;9(6):e99610.
57. Gaudin R, Kirchhausen T. Superinfection exclusion is absent during acute Junin virus infection of Vero and A549 cells. *Sci Rep.* 2015;5:15990.
58. Huang C, Kolokoltsova OA, Yun NE, Seregin AV, Ronca S, Koma T, et al. Highly Pathogenic New World and Old World Human Arenaviruses Induce Distinct Interferon Responses in Human Cells. *J Virol.* 2015;89(14):7079-88.
59. Perez Vidakovics MLA, Ure AE, Arrias PN, Romanowski V, Gomez RM. Junin virus induces autophagy in human A549 cells. *PLoS One.* 2019;14(6):e0218730.
60. Roldan JS, Candurra NA, Colombo MI, Delgui LR. Junin Virus Promotes Autophagy To Facilitate the Virus Life Cycle. *J Virol.* 2019;93(15).

61. Giovannoni F, Vazquez CA, Thomas P, Gomez RM, Garcia CC. Promyelocytic leukemia protein is a restriction factor for Junin virus independently of Z matrix protein. *Biochem Biophys Res Commun.* 2022;606:168-73.
62. Reuter G, Boros A, Takats K, Matics R, Pankovics P. A novel mammarenavirus (family Arenaviridae) in hedgehogs (*Erinaceus roumanicus*) in Europe. *Arch Virol.* 2023;168(7):174.
63. Radoshitzky SR, Buchmeier MJ, Charrel RN, Gonzalez JJ, Gunther S, Hepojoki J, et al. ICTV Virus Taxonomy Profile: Arenaviridae 2023. *The Journal of general virology.* 2023;104(9).
64. Stenglein MD, Sanders C, Kistler AL, Ruby JG, Franco JY, Reavill DR, et al. Identification, characterization, and in vitro culture of highly divergent arenaviruses from boa constrictors and annulated tree boas: candidate etiological agents for snake inclusion body disease. *mBio.* 2012;3(4):e00180-12.
65. Bodewes R, Kik MJL, Raj VS, Schapendonk CME, Haagmans BL, Smits SL, et al. Detection of novel divergent arenaviruses in boid snakes with inclusion body disease in The Netherlands. *J Gen Virol.* 2013;94(Pt 6):1206-10.
66. Hetzel U, Sironen T, Laurinmaki P, Liljeroos L, Patjas A, Henttonen H, et al. Isolation, identification, and characterization of novel arenaviruses, the etiological agents of boid inclusion body disease. *J Virol.* 2013;87(20):10918-35.
67. Hepojoki J, Hepojoki S, Smura T, Szivovics L, Dervas E, Prahauer B, et al. Characterization of Haartman Institute snake virus-1 (HISV-1) and HISV-like viruses-The representatives of genus Hartmanivirus, family Arenaviridae. *PLoS Pathog.* 2018;14(11):e1007415.
68. Koellhoffer JF, Dai Z, Malashkevich VN, Stenglein MD, Liu Y, Toro R, et al. Structural characterization of the glycoprotein GP2 core domain from the CAS virus, a novel arenavirus-like species. *J Mol Biol.* 2014;426(7):1452-68.
69. Hepojoki J, Salmenpera P, Sironen T, Hetzel U, Korzyukov Y, Kipar A, et al. Arenavirus Coinfections Are Common in Snakes with Boid Inclusion Body Disease. *J Virol.* 2015;89(16):8657-60.
70. Keller S, Hetzel U, Sironen T, Korzyukov Y, Vapalahti O, Kipar A, et al. Co-infecting Reptarenaviruses Can Be Vertically Transmitted in Boa Constrictor. *PLoS Pathog.* 2017;13(1):e1006179.
71. Stenglein MD, Jacobson ER, Chang LW, Sanders C, Hawkins MG, Guzman DS, et al. Widespread recombination, reassortment, and transmission of unbalanced compound viral genotypes in natural arenavirus infections. *PLoS Pathog.* 2015;11(5):e1004900.
72. Shi M, Lin XD, Chen X, Tian JH, Chen LJ, Li K, et al. The evolutionary history of vertebrate RNA viruses. *Nature.* 2018;556(7700):197-202.
73. Gonzalez JP, Emonet S, de Lamballerie X, Charrel R. Arenaviruses. *Curr Top Microbiol Immunol.* 2007;315:253-88.
74. Salazar-Bravo J, Ruedas LA, Yates TL. Mammalian reservoirs of arenaviruses. *Curr Top Microbiol Immunol.* 2002;262:25-63.
75. Li K, Lin XD, Wang W, Shi M, Guo WP, Zhang XH, et al. Isolation and characterization of a novel arenavirus harbored by Rodents and Shrews in Zhejiang province, China. *Virology.* 2015;476:37-42.
76. Wu Z, Du J, Lu L, Yang L, Dong J, Sun L, et al. Detection of Hantaviruses and Arenaviruses in three-toed jerboas from the Inner Mongolia Autonomous Region, China. *Emerg Microbes Infect.* 2018;7(1):35.
77. Radoshitzky SR, Buchmeier MJ, Charrel RN, Gonzalez JJ, Gunther S, Hepojoki J, et al. ICTV Virus Taxonomy Profile: Arenaviridae 2023. *The Journal of general virology.* 2023;104(9).
78. Briese T, Paweska JT, McMullan LK, Hutchison SK, Street C, Palacios G, et al. Genetic detection and characterization of Lujo virus, a new hemorrhagic fever-associated arenavirus from southern Africa. *PLoS Pathog.* 2009;5(5):e1000455.
79. Buckley SM, Casals J, Downs WG. Isolation and antigenic characterization of Lassa virus. *Nature.* 1970;227(5254):174.

80. Smadel JE, Wall MJ. Identification of the Virus of Lymphocytic Choriomeningitis. *J Bacteriol.* 1941;41(4):421-30.
81. Emonet S, Lemasson JJ, Gonzalez JP, de Lamballerie X, Charrel RN. Phylogeny and evolution of old world arenaviruses. *Virology.* 2006;350(2):251-7.
82. Lecompte E, ter Meulen J, Emonet S, Daffis S, Charrel RN. Genetic identification of Kodoko virus, a novel arenavirus of the African pigmy mouse (*Mus Nannomys minutoides*) in West Africa. *Virology.* 2007;364(1):178-83.
83. Gunther S, Hoofd G, Charrel R, Roser C, Becker-Ziaja B, Lloyd G, et al. Mopeia virus-related arenavirus in natal multimammate mice, Morogoro, Tanzania. *Emerg Infect Dis.* 2009;15(12):2008-12.
84. Bowen MD, Peters CJ, Nichol ST. Phylogenetic analysis of the Arenaviridae: patterns of virus evolution and evidence for cospeciation between arenaviruses and their rodent hosts. *Mol Phylogenet Evol.* 1997;8(3):301-16.
85. Hugot JP, Gonzalez JP, Denys C. Evolution of the Old World Arenaviridae and their rodent hosts: generalized host-transfer or association by descent? *Infect Genet Evol.* 2001;1(1):13-20.
86. Bowen MD, Peters CJ, Nichol ST. The phylogeny of New World (Tacaribe complex) arenaviruses. *Virology.* 1996;219(1):285-90.
87. Radoshitzky SR, Bao Y, Buchmeier MJ, Charrel RN, Clawson AN, Clegg CS, et al. Past, present, and future of arenavirus taxonomy. *Archives of virology.* 2015;160(7):1851-74.
88. Delgado S, Erickson BR, Agudo R, Blair PJ, Vallejo E, Albarino CG, et al. Chapare virus, a newly discovered arenavirus isolated from a fatal hemorrhagic fever case in Bolivia. *PLoS Pathog.* 2008;4(4):e1000047.
89. Charrel RN, Feldmann H, Fulhorst CF, Khelifa R, de Chesse R, de Lamballerie X. Phylogeny of New World arenaviruses based on the complete coding sequences of the small genomic segment identified an evolutionary lineage produced by intrasegmental recombination. *Biochem Biophys Res Commun.* 2002;296(5):1118-24.
90. Tortorici MA, Ghiringhelli PD, Lozano ME, Albarino CG, Romanowski V. Zinc-binding properties of Junin virus nucleocapsid protein. *The Journal of general virology.* 2001;82(Pt 1):121-8.
91. Ghiringhelli PD, Rivera-Pomar RV, Lozano ME, Grau O, Romanowski V. Molecular organization of Junin virus S RNA: complete nucleotide sequence, relationship with other members of the Arenaviridae and unusual secondary structures. *J Gen Virol.* 1991;72 (Pt 9):2129-41.
92. Fulhorst CF, Bowen MD, Salas RA, Duno G, Utrera A, Ksiazek TG, et al. Natural rodent host associations of Guanarito and pirital viruses (Family Arenaviridae) in central Venezuela. *Am J Trop Med Hyg.* 1999;61(2):325-30.
93. Knipe DM, Howley P. *Fields Virology.* 6th ed: Wolters Kluwer Health; 2013.
94. Cao W, Henry MD, Borrow P, Yamada H, Elder JH, Ravkov EV, et al. Identification of alpha-dystroglycan as a receptor for lymphocytic choriomeningitis virus and Lassa fever virus. *Science.* 1998;282(5396):2079-81.
95. Spiropoulou CF, Kunz S, Rollin PE, Campbell KP, Oldstone MB. New World arenavirus clade C, but not clade A and B viruses, utilizes alpha-dystroglycan as its major receptor. *J Virol.* 2002;76(10):5140-6.
96. Shimojima M, Stroher U, Ebihara H, Feldmann H, Kawaoka Y. Identification of cell surface molecules involved in dystroglycan-independent Lassa virus cell entry. *J Virol.* 2012;86(4):2067-78.
97. Shimojima M, Kawaoka Y. Cell surface molecules involved in infection mediated by lymphocytic choriomeningitis virus glycoprotein. *J Vet Med Sci.* 2012;74(10):1363-6.
98. Martinez MG, Cordo SM, Candurra NA. Characterization of Junin arenavirus cell entry. *J Gen Virol.* 2007;88(Pt 6):1776-84.

99. Kunz S. Receptor binding and cell entry of Old World arenaviruses reveal novel aspects of virus-host interaction. *Virology*. 2009;387(2):245-9.
100. Grant A, Seregin A, Huang C, Kolokoltsova O, Brasier A, Peters C, et al. Junin virus pathogenesis and virus replication. *Viruses*. 2012;4(10):2317-39.
101. Vela E. Animal models, prophylaxis, and therapeutics for arenavirus infections. *Viruses*. 2012;4(9):1802-29.
102. Abraham J, Kwong JA, Albarino CG, Lu JG, Radoshitzky SR, Salazar-Bravo J, et al. Host-species transferrin receptor 1 orthologs are cellular receptors for nonpathogenic new world clade B arenaviruses. *PLoS Pathog*. 2009;5(4):e1000358.
103. Lavanya M, Cuevas CD, Thomas M, Cherry S, Ross SR. siRNA screen for genes that affect Junin virus entry uncovers voltage-gated calcium channels as a therapeutic target. *Sci Transl Med*. 2013;5(204):204ra131.
104. Jemielity S, Wang JJ, Chan YK, Ahmed AA, Li W, Monahan S, et al. TIM-family proteins promote infection of multiple enveloped viruses through virion-associated phosphatidylserine. *PLoS Pathog*. 2013;9(3):e1003232.
105. Martinez MG, Bialecki MA, Belouzard S, Cordo SM, Candurra NA, Whittaker GR. Utilization of human DC-SIGN and L-SIGN for entry and infection of host cells by the New World arenavirus, Junin virus. *Biochem Biophys Res Commun*. 2013;441(3):612-7.
106. Martinez MG, Forlenza MB, Candurra NA. Involvement of cellular proteins in Junin arenavirus entry. *Biotechnol J*. 2009;4(6):866-70.
107. Rojek JM, Sanchez AB, Nguyen NT, de la Torre JC, Kunz S. Different mechanisms of cell entry by human-pathogenic Old World and New World arenaviruses. *J Virol*. 2008;82(15):7677-87.
108. Martinez MG, Cordo SM, Candurra NA. Involvement of cytoskeleton in Junin virus entry. *Virus Res*. 2008;138(1-2):17-25.
109. Castilla V, Mersich SE, Candurra NA, Damonte EB. The entry of Junin virus into Vero cells. *Arch Virol*. 1994;136(3-4):363-74.
110. Di Simone C, Buchmeier MJ. Kinetics and pH dependence of acid-induced structural changes in the lymphocytic choriomeningitis virus glycoprotein complex. *Virology*. 1995;209(1):3-9.
111. Meyer BJ, Southern PJ. Concurrent sequence analysis of 5' and 3' RNA termini by intramolecular circularization reveals 5' nontemplated bases and 3' terminal heterogeneity for lymphocytic choriomeningitis virus mRNAs. *J Virol*. 1993;67(5):2621-7.
112. Foscaldi S, D'Antuono A, Noval MG, de Prat Gay G, Scolaro L, Lopez N. Regulation of Tacaribe Mammarenavirus Translation: Positive 5' and Negative 3' Elements and Role of Key Cellular Factors. *Journal of virology*. 2017;91(14).
113. Lee KJ, Novella IS, Teng MN, Oldstone MB, de La Torre JC. NP and L proteins of lymphocytic choriomeningitis virus (LCMV) are sufficient for efficient transcription and replication of LCMV genomic RNA analogs. *J Virol*. 2000;74(8):3470-7.
114. Lopez N, Jacamo R, Franze-Fernandez MT. Transcription and RNA replication of tacaribe virus genome and antigenome analogs require N and L proteins: Z protein is an inhibitor of these processes. *J Virol*. 2001;75(24):12241-51.
115. Dalton AJ, Rowe WP, Smith GH, Wilsnack RE, Pugh WE. Morphological and cytochemical studies on lymphocytic choriomeningitis virus. *J Virol*. 1968;2(12):1465-78.
116. Wang W, Zhou Z, Zhang L, Wang S, Xiao G. Structure-function relationship of the mammarenavirus envelope glycoprotein. *Viol Sin*. 2016;31(5):380-94.

117. Lopez N, Franze-Fernandez MT. A single stem-loop structure in Tacaribe arenavirus intergenic region is essential for transcription termination but is not required for a correct initiation of transcription and replication. *Virus Res.* 2007;124(1-2):237-44.
118. Pinschewer DD, Perez M, de la Torre JC. Dual role of the lymphocytic choriomeningitis virus intergenic region in transcription termination and virus propagation. *J Virol.* 2005;79(7):4519-26.
119. Golden JW, Beitzel B, Ladner JT, Mucker EM, Kwilas SA, Palacios G, et al. An attenuated Machupo virus with a disrupted L-segment intergenic region protects Guinea pigs against lethal Guaranito virus infection. *Scientific Reports.* 2017;7(1):1-12.
120. Iwasaki M, Ngo N, Cubitt B, Teijaro JR, de la Torre JC. General Molecular Strategy for Development of Arenavirus Live-Attenuated Vaccines. *Journal of virology.* 2015;89(23):12166-77.
121. Shuzo Urata JY, Juan Carlos de la Torre. The Z Protein of the New World Arenavirus Tacaribe Virus has Bona Fide Budding Activity that does not depend on known late domain motifs. *Journal of Virology.* 2009;Vol.83, No.23; 12651-5.
122. Perez M GD, de la Torre JC. Myristoylation of the RING finger Z protein is essential for arenavirus budding. *J Virol* 78. 2004:11443-8.
123. Juan Cruz Casabona JMLM, Maia Eugenia Loureiro, Guillermo A Gomez, Nora Lopez. The Ring Domain and the L79 Residue of Z Protein are involved in both the Rescue of Nucleocapsids and the Incorporation of glycoproteins into Infectious Chimeric Arenavirus-Like Particles. *Journal of Virology.* 2009;Vol. 83, No. 14; 7029-39.
124. Burri DJ, da Palma JR, Kunz S, Pasquato A. Envelope glycoprotein of arenaviruses. *Viruses.* 2012;4(10):2162-81.
125. Kenyon RH, Green DE, Peters CJ. Effect of immunosuppression on experimental Argentine hemorrhagic fever in guinea pigs. *J Virol.* 1985;53(1):75-80.
126. Weissenbacher MC, Laguens RP, Coto CE. Argentine hemorrhagic fever. *Curr Top Microbiol Immunol.* 1987;134:79-116.
127. Candurra NA, Damonte EB, Coto CE. Antigenic relationships between attenuated and pathogenic strains of Junin virus. *J Med Virol.* 1989;27(2):145-50.
128. Ruggiero HA, Perez Isquierdo F, Milani HA, Barri A, Val A, Maglio F, et al. [Treatment of Argentine hemorrhagic fever with convalescent's plasma. 4433 cases]. *Presse Med.* 1986;15(45):2239-42.
129. de Guerrero LB, Boxaca MC, Malumbres E, Dejean C, Caruso E. Early protection to Junin virus of guinea pig with an attenuated Junin virus strain. *Acta Virol.* 1985;29(4):334-7.
130. Ambrosio A, Saavedra M, Mariani M, Gamboa G, Maiza A. Argentine hemorrhagic fever vaccines. *Hum Vaccin.* 2011;7(6):694-700.
131. McKee KT, Jr., Oro JG, Kuehne AI, Spisso JA, Mahlandt BG. Candid No. 1 Argentine hemorrhagic fever vaccine protects against lethal Junin virus challenge in rhesus macaques. *Intervirology.* 1992;34(3):154-63.
132. Medeot SI, Contigiani MS, Brandan ER, Sabattini MS. Neurovirulence of wild and laboratory Junin virus strains in animal hosts. *J Med Virol.* 1990;32(3):171-82.
133. Maiztegui JI, McKee KT, Jr., Barrera Oro JG, Harrison LH, Gibbs PH, Feuillade MR, et al. Protective efficacy of a live attenuated vaccine against Argentine hemorrhagic fever. AHF Study Group. *J Infect Dis.* 1998;177(2):277-83.
134. McKee KT, Mahlandt BG, Maiztegui JI, Eddy GA, Peters CJ. Experimental Argentine hemorrhagic fever in rhesus macaques: Viral strain-dependent clinical response. *Journal of Infectious Diseases.* 1985;152(1):218-21.
135. Fehling SK, Lennartz F, Strecker T. Multifunctional nature of the arenavirus RING finger protein Z. *Viruses.* 2012;4(11):2973-3011.

136. Volpon L, Osborne MJ, Capul AA, de la Torre JC, Borden KL. Structural characterization of the Z RING-eIF4E complex reveals a distinct mode of control for eIF4E. *Proc Natl Acad Sci U S A.* 2010;107(12):5441-6.
137. Kentsis A, Gordon RE, Borden KL. Self-assembly properties of a model RING domain. *Proc Natl Acad Sci U S A.* 2002;99(2):667-72.
138. Kentsis A, Gordon RE, Borden KL. Control of biochemical reactions through supramolecular RING domain self-assembly. *Proc Natl Acad Sci U S A.* 2002;99(24):15404-9.
139. Kentsis A, Dwyer EC, Perez JM, Sharma M, Chen A, Pan ZQ, et al. The RING domains of the promyelocytic leukemia protein PML and the arenaviral protein Z repress translation by directly inhibiting translation initiation factor eIF4E. *J Mol Biol.* 2001;312(4):609-23.
140. Freed EO. Viral late domains. *J Virol.* 2002;76(10):4679-87.
141. Lemmon SK, Traub LM. Sorting in the endosomal system in yeast and animal cells. *Curr Opin Cell Biol.* 2000;12(4):457-66.
142. Perez M, Greenwald DL, de la Torre JC. Myristoylation of the RING finger Z protein is essential for arenavirus budding. *J Virol.* 2004;78(20):11443-8.
143. Strecker T, Maisa A, Daffis S, Eichler R, Lenz O, Garten W. The role of myristoylation in the membrane association of the Lassa virus matrix protein Z. *Virol J.* 2006;3:93.
144. Farazi TA, Waksman G, Gordon JI. The biology and enzymology of protein N-myristoylation. *J Biol Chem.* 2001;276(43):39501-4.
145. Capul AA, Perez M, Burke E, Kunz S, Buchmeier MJ, de la Torre JC. Arenavirus Z-glycoprotein association requires Z myristoylation but not functional RING or late domains. *J Virol.* 2007;81(17):9451-60.
146. Loureiro ME, Wilda M, Levingston Macleod JM, D'Antuono A, Foscaldi S, Marino Buslje C, et al. Molecular determinants of arenavirus Z protein homo-oligomerization and L polymerase binding. *J Virol.* 2011;85(23):12304-14.
147. Kranzusch PJ, Whelan SP. Arenavirus Z protein controls viral RNA synthesis by locking a polymerase-promoter complex. *Proc Natl Acad Sci U S A.* 2011;108(49):19743-8.
148. Cornu TI, de la Torre JC. RING finger Z protein of lymphocytic choriomeningitis virus (LCMV) inhibits transcription and RNA replication of an LCMV S-segment minigenome. *J Virol.* 2001;75(19):9415-26.
149. Capul AA, de la Torre JC, Buchmeier MJ. Conserved residues in Lassa fever virus Z protein modulate viral infectivity at the level of the ribonucleoprotein. *J Virol.* 2011;85(7):3172-8.
150. Cornu TI, de la Torre JC. Characterization of the arenavirus RING finger Z protein regions required for Z-mediated inhibition of viral RNA synthesis. *J Virol.* 2002;76(13):6678-88.
151. Schlie K, Maisa A, Freiberg F, Groseth A, Strecker T, Garten W. Viral protein determinants of Lassa virus entry and release from polarized epithelial cells. *J Virol.* 2010;84(7):3178-88.
152. Neuman BW, Adair BD, Burns JW, Milligan RA, Buchmeier MJ, Yeager M. Complementarity in the supramolecular design of arenaviruses and retroviruses revealed by electron cryomicroscopy and image analysis. *J Virol.* 2005;79(6):3822-30.
153. Strecker T, Eichler R, Meulen J, Weissenhorn W, Dieter Klenk H, Garten W, et al. Lassa virus Z protein is a matrix protein and sufficient for the release of virus-like particles [corrected]. *J Virol.* 2003;77(19):10700-5.
154. Perez M, Craven RC, de la Torre JC. The small RING finger protein Z drives arenavirus budding: implications for antiviral strategies. *Proc Natl Acad Sci U S A.* 2003;100(22):12978-83.
155. Garrus JE, von Schwedler UK, Pornillos OW, Morham SG, Zavitz KH, Wang HE, et al. Tsg101 and the vacuolar protein sorting pathway are essential for HIV-1 budding. *Cell.* 2001;107(1):55-65.
156. Martin-Serrano J, Zang T, Bieniasz PD. HIV-1 and Ebola virus encode small peptide motifs that recruit Tsg101 to sites of particle assembly to facilitate egress. *Nat Med.* 2001;7(12):1313-9.

157. VerPlank L, Bouamr F, LaGrassa TJ, Agresta B, Kikonyogo A, Leis J, et al. Tsg101, a homologue of ubiquitin-conjugating (E2) enzymes, binds the L domain in HIV type 1 Pr55(Gag). *Proc Natl Acad Sci U S A*. 2001;98(14):7724-9.
158. Martin-Serrano J, Yarovoy A, Perez-Caballero D, Bieniasz PD. Divergent retroviral late-budding domains recruit vacuolar protein sorting factors by using alternative adaptor proteins. *Proc Natl Acad Sci U S A*. 2003;100(21):12414-9.
159. Strack B, Calistri A, Craig S, Popova E, Gottlinger HG. AIP1/ALIX is a binding partner for HIV-1 p6 and EIAV p9 functioning in virus budding. *Cell*. 2003;114(6):689-99.
160. von Schwedler UK, Stuchell M, Muller B, Ward DM, Chung HY, Morita E, et al. The protein network of HIV budding. *Cell*. 2003;114(6):701-13.
161. Urata S, Noda T, Kawaoka Y, Yokosawa H, Yasuda J. Cellular factors required for Lassa virus budding. *J Virol*. 2006;80(8):4191-5.
162. Urata S, Yasuda J, de la Torre JC. The z protein of the new world arenavirus tacaribe virus has bona fide budding activity that does not depend on known late domain motifs. *J Virol*. 2009;83(23):12651-5.
163. Borden KL, Campbell Dwyer EJ, Salvato MS. An arenavirus RING (zinc-binding) protein binds the oncoprotein promyelocyte leukemia protein (PML) and relocates PML nuclear bodies to the cytoplasm. *J Virol*. 1998;72(1):758-66.
164. Borden KL, CampbellDwyer EJ, Carlile GW, Djavani M, Salvato MS. Two RING finger proteins, the oncoprotein PML and the arenavirus Z protein, colocalize with the nuclear fraction of the ribosomal P proteins. *J Virol*. 1998;72(5):3819-26.
165. Campbell Dwyer EJ, Lai H, MacDonald RC, Salvato MS, Borden KL. The lymphocytic choriomeningitis virus RING protein Z associates with eukaryotic initiation factor 4E and selectively represses translation in a RING-dependent manner. *J Virol*. 2000;74(7):3293-300.
166. Topcu Z, Mack DL, Hromas RA, Borden KL. The promyelocytic leukemia protein PML interacts with the proline-rich homeodomain protein PRH: a RING may link hematopoiesis and growth control. *Oncogene*. 1999;18(50):7091-100.
167. Reineke EL, Kao HY. Targeting promyelocytic leukemia protein: a means to regulating PML nuclear bodies. *Int J Biol Sci*. 2009;5(4):366-76.
168. Boccaccio GL, Thomas MG, Garcia CC. Membraneless Organelles and Condensates Orchestrate Innate Immunity Against Viruses. *J Mol Biol*. 2023;435(16):167976.
169. Zhong S, Salomoni P, Ronchetti S, Guo A, Ruggero D, Pandolfi PP. Promyelocytic leukemia protein (PML) and Daxx participate in a novel nuclear pathway for apoptosis. *J Exp Med*. 2000;191(4):631-40.
170. Borden KL, CampbellDwyer EJ, Salvato MS. The promyelocytic leukemia protein PML has a pro-apoptotic activity mediated through its RING domain. *FEBS Lett*. 1997;418(1-2):30-4.
171. Bernardi R, Papa A, Pandolfi PP. Regulation of apoptosis by PML and the PML-NBs. *Oncogene*. 2008;27(48):6299-312.
172. Regad T, Chelbi-Alix MK. Role and fate of PML nuclear bodies in response to interferon and viral infections. *Oncogene*. 2001;20(49):7274-86.
173. Regad T, Saib A, Lallemand-Breitenbach V, Pandolfi PP, de The H, Chelbi-Alix MK. PML mediates the interferon-induced antiviral state against a complex retrovirus via its association with the viral transactivator. *EMBO J*. 2001;20(13):3495-505.
174. Pestova TV, Kolupaeva VG, Lomakin IB, Pilipenko EV, Shatsky IN, Agol VI, et al. Molecular mechanisms of translation initiation in eukaryotes. *Proc Natl Acad Sci U S A*. 2001;98(13):7029-36.
175. Martinez-Sobrido L, Giannakas P, Cubitt B, Garcia-Sastre A, de la Torre JC. Differential inhibition of type I interferon induction by arenavirus nucleoproteins. *J Virol*. 2007;81(22):12696-703.

176. Carnec X, Baize S, Reynard S, Diancourt L, Caro V, Tordo N, et al. Lassa virus nucleoprotein mutants generated by reverse genetics induce a robust type I interferon response in human dendritic cells and macrophages. *J Virol.* 2011;85(22):12093-7.
177. Muller S, Geffers R, Gunther S. Analysis of gene expression in Lassa virus-infected HuH-7 cells. *J Gen Virol.* 2007;88(Pt 5):1568-75.
178. Groseth A, Hoenen T, Weber M, Wolff S, Herwig A, Kaufmann A, et al. Tacaribe virus but not junin virus infection induces cytokine release from primary human monocytes and macrophages. *PLoS Negl Trop Dis.* 2011;5(5):e1137.
179. Fan L, Briese T, Lipkin WI. Z proteins of New World arenaviruses bind RIG-I and interfere with type I interferon induction. *J Virol.* 2010;84(4):1785-91.
180. Xing J, Ly H, Liang Y. The Z proteins of pathogenic but not nonpathogenic arenaviruses inhibit RIG-I-like receptor-dependent interferon production. *J Virol.* 2015;89(5):2944-55.
181. Mirdita M, Schutze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold: making protein folding accessible to all. *Nat Methods.* 2022;19(6):679-82.
182. Adiyaman R, McGuffin LJ. ReFOLD3: refinement of 3D protein models with gradual restraints based on predicted local quality and residue contacts. *Nucleic Acids Res.* 2021;49(W1):W589-W96.
183. Laimer J, Hiebl-Flach J, Lengauer D, Lackner P. MAESTROweb: a web server for structure-based protein stability prediction. *Bioinformatics.* 2016;32(9):1414-6.
184. Rodrigues CH, Pires DE, Ascher DB. DynaMut: predicting the impact of mutations on protein conformation, flexibility and stability. *Nucleic Acids Research.* 2018;46(W1):W350-W5.
185. Pinschewer DD, Perez M, de la Torre JC. Role of the virus nucleoprotein in the regulation of lymphocytic choriomeningitis virus transcription and RNA replication. *J Virol.* 2003;77(6):3882-7.
186. Borrow P, Martinez-Sobrido L, de la Torre JC. Inhibition of the type I interferon antiviral response during arenavirus infection. *Viruses.* 2010;2(11):2443-80.
187. Martinez-Sobrido L, Zuniga EI, Rosario D, Garcia-Sastre A, de la Torre JC. Inhibition of the type I interferon response by the nucleoprotein of the prototypic arenavirus lymphocytic choriomeningitis virus. *J Virol.* 2006;80(18):9192-9.
188. Qi X, Lan S, Wang W, Schelde LM, Dong H, Wallat GD, et al. Cap binding and immune evasion revealed by Lassa nucleoprotein structure. *Nature.* 2010;468(7325):779-83.
189. Brunotte L, Kerber R, Shang W, Hauer F, Hass M, Gabriel M, et al. Structure of the Lassa virus nucleoprotein revealed by X-ray crystallography, small-angle X-ray scattering, and electron microscopy. *J Biol Chem.* 2011;286(44):38748-56.
190. Ortiz-Riano E, Cheng BY, de la Torre JC, Martinez-Sobrido L. Self-association of lymphocytic choriomeningitis virus nucleoprotein is mediated by its N-terminal region and is not required for its anti-interferon function. *J Virol.* 2012;86(6):3307-17.
191. Levingston Macleod JM, D'Antuono A, Loureiro ME, Casabona JC, Gomez GA, Lopez N. Identification of two functional domains within the arenavirus nucleoprotein. *J Virol.* 2011;85(5):2012-23.
192. Ortiz-Riano E, Cheng BY, de la Torre JC, Martinez-Sobrido L. The C-terminal region of lymphocytic choriomeningitis virus nucleoprotein contains distinct and segregable functional domains involved in NP-Z interaction and counteraction of the type I interferon response. *J Virol.* 2011;85(24):13038-48.
193. Martinez-Sobrido L, Emonet S, Giannakas P, Cubitt B, Garcia-Sastre A, de la Torre JC. Identification of amino acid residues critical for the anti-interferon activity of the nucleoprotein of the prototypic arenavirus lymphocytic choriomeningitis virus. *J Virol.* 2009;83(21):11330-40.
194. Pythoud C, Rodrigo WW, Pasqual G, Rothenberger S, Martinez-Sobrido L, de la Torre JC, et al. Arenavirus nucleoprotein targets interferon regulatory factor-activating kinase IKKepsilon. *J Virol.* 2012;86(15):7728-38.

195. Huang Q, Shao J, Lan S, Zhou Y, Xing J, Dong C, et al. In vitro and in vivo characterizations of pichinde viral nucleoprotein exoribonuclease functions. *J Virol*. 2015;89(13):6595-607.
196. Hastie KM, Kimberlin CR, Zandonatti MA, MacRae IJ, Saphire EO. Structure of the Lassa virus nucleoprotein reveals a dsRNA-specific 3' to 5' exonuclease activity essential for immune suppression. *Proc Natl Acad Sci U S A*. 2011;108(6):2396-401.
197. Pyle JD, Whelan SPJ. Isolation of Reconstructed Functional Ribonucleoprotein Complexes of Machupo Virus. *J Virol*. 2021;95(22):e0105421.
198. Hass M, Golnitz U, Muller S, Becker-Ziaja B, Gunther S. Replicon system for Lassa virus. *J Virol*. 2004;78(24):13793-803.
199. Baird NL, York J, Nunberg JH. Arenavirus infection induces discrete cytosolic structures for RNA replication. *J Virol*. 2012;86(20):11301-10.
200. Young PR, Howard CR. Fine structure analysis of Pichinde virus nucleocapsids. *J Gen Virol*. 1983;64(Pt 4):833-42.
201. Ruigrok RW, Crepin T, Kolakofsky D. Nucleoproteins and nucleocapsids of negative-strand RNA viruses. *Curr Opin Microbiol*. 2011;14(4):504-10.
202. Ferron F, Weber F, de la Torre JC, Reguera J. Transcription and replication mechanisms of Bunyaviridae and Arenaviridae L proteins. *Virus Res*. 2017;234:118-34.
203. Knopp KA, Ngo T, Gershon PD, Buchmeier MJ. Single nucleoprotein residue modulates arenavirus replication complex formation. *mBio*. 2015;6(3):e00524-15.
204. Huang C, Mantlo E, Paessler S. Lassa virus NP DEDDh 3'-5' exoribonuclease activity is required for optimal viral RNA replication. *bioRxiv*. 2023.
205. King BR, Hershkowitz D, Eisenhauer PL, Weir ME, Ziegler CM, Russo J, et al. A Map of the Arenavirus Nucleoprotein-Host Protein Interactome Reveals that Junin Virus Selectively Impairs the Antiviral Activity of Double-Stranded RNA-Activated Protein Kinase (PKR). *J Virol*. 2017;91(15).
206. Linero FN, Thomas MG, Boccaccio GL, Scolaro LA. Junin virus infection impairs stress-granule formation in Vero cells treated with arsenite via inhibition of eIF2alpha phosphorylation. *J Gen Virol*. 2011;92(Pt 12):2889-99.
207. Loureiro ME, D'Antuono A, Levingston Macleod JM, Lopez N. Uncovering viral protein-protein interactions and their role in arenavirus life cycle. *Viruses*. 2012;4(9):1651-67.
208. Maeto CA, Knott ME, Linero FN, Ellenberg PC, Scolaro LA, Castilla V. Differential effect of acute and persistent Junin virus infections on the nucleo-cytoplasmic trafficking and expression of heterogeneous nuclear ribonucleoproteins type A and B. *J Gen Virol*. 2011;92(Pt 9):2181-90.
209. Brunetti JE, Scolaro LA, Castilla V. The heterogeneous nuclear ribonucleoprotein K (hnRNP K) is a host factor required for dengue virus and Junin virus multiplication. *Virus Res*. 2015;203:84-91.
210. Geuens T, Bouhy D, Timmerman V. The hnRNP family: insights into their role in health and disease. *Hum Genet*. 2016;135(8):851-67.
211. Thibault PA, Ganesan A, Kalyanamoorthy S, Clarke JWE, Salapa HE, Levin MC. hnRNP A/B Proteins: An Encyclopedic Assessment of Their Roles in Homeostasis and Disease. *Biology (Basel)*. 2021;10(8).
212. Cuevas CD, Lavanya M, Wang E, Ross SR. Junin virus infects mouse cells and induces innate immune responses. *J Virol*. 2011;85(21):11058-68.
213. Mateer EJ, Maruyama J, Card GE, Paessler S, Huang C. Lassa Virus, but Not Highly Pathogenic New World Arenaviruses, Restricts Immunostimulatory Double-Stranded RNA Accumulation during Infection. *J Virol*. 2020;94(9).
214. Moreno H, Kunz S. The Protein Kinase Receptor Modulates the Innate Immune Response against Tacaribe Virus. *Viruses*. 2021;13(7).

215. Jiang X, Huang Q, Wang W, Dong H, Ly H, Liang Y, et al. Structures of arenaviral nucleoproteins with triphosphate dsRNA reveal a unique mechanism of immune suppression. *J Biol Chem.* 2013;288(23):16949-59.
216. Zuo Y, Deutscher MP. Exoribonuclease superfamilies: structural analysis and phylogenetic distribution. *Nucleic Acids Res.* 2001;29(5):1017-26.
217. Mateer EJ, Paessler S, Huang C. Visualization of Double-Stranded RNA Colocalizing With Pattern Recognition Receptors in Arenavirus Infected Cells. *Front Cell Infect Microbiol.* 2018;8:251.
218. Bohn P, Wassmann I, Wendt L, Leske A, Hoenen T, Tews BA, et al. A dsRNA-binding mutant reveals only a minor role of exonuclease activity in interferon antagonism by the arenavirus nucleoprotein. *PLoS Pathog.* 2023;19(1):e1011049.
219. MacArthur MW, Thornton JM. Influence of proline residues on protein conformation. *J Mol Biol.* 1991;218(2):397-412.
220. Breunig SL, Tirrell DA. Incorporation of proline analogs into recombinant proteins expressed in *Escherichia coli*. *Methods Enzymol.* 2021;656:545-71.
221. Schlie K, Maisa A, Lennartz F, Stroher U, Garten W, Strecker T. Characterization of Lassa virus glycoprotein oligomerization and influence of cholesterol on virus replication. *J Virol.* 2010;84(2):983-92.
222. Agnihothram SS, York J, Trahey M, Nunberg JH. Bitopic membrane topology of the stable signal peptide in the tripartite Junin virus GP-C envelope glycoprotein complex. *J Virol.* 2007;81(8):4331-7.
223. Eichler R, Lenz O, Strecker T, Garten W. Signal peptide of Lassa virus glycoprotein GP-C exhibits an unusual length. *FEBS Lett.* 2003;538(1-3):203-6.
224. Buchmeier MJ, Oldstone MB. Protein structure of lymphocytic choriomeningitis virus: evidence for a cell-associated precursor of the virion glycopeptides. *Virology.* 1979;99(1):111-20.
225. Lenz O, ter Meulen J, Klenk HD, Seidah NG, Garten W. The Lassa virus glycoprotein precursor GP-C is proteolytically processed by subtilase SKI-1/S1P. *Proc Natl Acad Sci U S A.* 2001;98(22):12701-5.
226. Beyer WR, Popplau D, Garten W, von Laer D, Lenz O. Endoproteolytic processing of the lymphocytic choriomeningitis virus glycoprotein by the subtilase SKI-1/S1P. *J Virol.* 2003;77(5):2866-72.
227. Rojek JM, Lee AM, Nguyen N, Spiropoulou CF, Kunz S. Site 1 protease is required for proteolytic processing of the glycoproteins of the South American hemorrhagic fever viruses Junin, Machupo, and Guanarito. *J Virol.* 2008;82(12):6045-51.
228. Buchmeier MJ, Southern PJ, Parekh BS, Wooddell MK, Oldstone MB. Site-specific antibodies define a cleavage site conserved among arenavirus GP-C glycoproteins. *J Virol.* 1987;61(4):982-5.
229. Agnihothram SS, York J, Nunberg JH. Role of the stable signal peptide and cytoplasmic domain of G2 in regulating intracellular transport of the Junin virus envelope glycoprotein complex. *J Virol.* 2006;80(11):5189-98.
230. Damonte EB, Mersich SE, Candurra NA. Intracellular processing and transport of Junin virus glycoproteins influences virion infectivity. *Virus Res.* 1994;34(3):317-26.
231. Hawkins JL, Robbins MD, Warren LC, Xia D, Petras SF, Valentine JJ, et al. Pharmacologic inhibition of site 1 protease activity inhibits sterol regulatory element-binding protein processing and reduces lipogenic enzyme gene expression and lipid synthesis in cultured cells and experimental animals. *J Pharmacol Exp Ther.* 2008;326(3):801-8.
232. Pasquato A, Pullikotil P, Asselin MC, Vacatello M, Paolillo L, Ghezzi F, et al. The proprotein convertase SKI-1/S1P. In vitro analysis of Lassa virus glycoprotein-derived substrates and ex vivo validation of irreversible peptide inhibitors. *J Biol Chem.* 2006;281(33):23471-81.
233. Zapata JC, Salvato MS. Arenavirus variations due to host-specific adaptation. *Viruses.* 2013;5(1):241-78.

234. York J, Nunberg JH. Intersubunit interactions modulate pH-induced activation of membrane fusion by the Junin virus envelope glycoprotein GPC. *J Virol.* 2009;83(9):4121-6.
235. Briknarova K, Thomas CJ, York J, Nunberg JH. Structure of a zinc-binding domain in the Junin virus envelope glycoprotein. *J Biol Chem.* 2011;286(2):1528-36.
236. Eichler R, Lenz O, Strecker T, Eickmann M, Klenk HD, Garten W. Lassa virus glycoprotein signal peptide displays a novel topology with an extended endoplasmic reticulum luminal region. *J Biol Chem.* 2004;279(13):12293-9.
237. York J, Nunberg JH. A novel zinc-binding domain is essential for formation of the functional Junin virus envelope glycoprotein complex. *J Virol.* 2007;81(24):13385-91.
238. Eichler R, Lenz O, Garten W, Strecker T. The role of single N-glycans in proteolytic processing and cell surface transport of the Lassa virus glycoprotein GP-C. *Virology.* 2006;341:3-11.
239. York J, Romanowski V, Lu M, Nunberg JH. The signal peptide of the Junin arenavirus envelope glycoprotein is myristoylated and forms an essential subunit of the mature G1-G2 complex. *J Virol.* 2004;78(19):10783-92.
240. Froeschke M, Basler M, Groettrup M, Dobberstein B. Long-lived signal peptide of lymphocytic choriomeningitis virus glycoprotein pGP-C. *J Biol Chem.* 2003;278(43):41914-20.
241. York J, Nunberg JH. Role of the stable signal peptide of Junin arenavirus envelope glycoprotein in pH-dependent membrane fusion. *J Virol.* 2006;80(15):7775-80.
242. Braakman I, van Anken E. Folding of viral envelope glycoproteins in the endoplasmic reticulum. *Traffic.* 2000;1(7):533-9.
243. Bonhomme CJ, Capul AA, Lauron EJ, Bederka LH, Knopp KA, Buchmeier MJ. Glycosylation modulates arenavirus glycoprotein expression and function. *Virology.* 2011;409(2):223-33.
244. Zhu X, Liu Y, Guo J, Cao J, Wang Z, Xiao G, et al. Effects of N-Linked Glycan on Lassa Virus Envelope Glycoprotein Cleavage, Infectivity, and Immune Response. *Virology.* 2021;36(4):774-83.
245. Sommerstein R, Flatz L, Remy MM, Malinge P, Magistrelli G, Fischer N, et al. Arenavirus Glycan Shield Promotes Neutralizing Antibody Evasion and Protracted Infection. *PLoS Pathog.* 2015;11(11):e1005276.
246. Klaus JP, Eisenhauer P, Russo J, Mason AB, Do D, King B, et al. The intracellular cargo receptor ERGIC-53 is required for the production of infectious arenavirus, coronavirus, and filovirus particles. *Cell Host Microbe.* 2013;14(5):522-34.
247. Casabona JC, Livingston Macleod JM, Loureiro ME, Gomez GA, Lopez N. The RING domain and the L79 residue of Z protein are involved in both the rescue of nucleocapsids and the incorporation of glycoproteins into infectious chimeric arenavirus-like particles. *J Virol.* 2009;83(14):7029-39.
248. Manning JT, Yun NE, Seregin AV, Koma T, Sattler RA, Ezeomah C, et al. The Glycoprotein of the Live-Attenuated Junin Virus Vaccine Strain Induces Endoplasmic Reticulum Stress and Forms Aggregates prior to Degradation in the Lysosome. *J Virol.* 2020;94(8):1-11.
249. Manning JT, Maruyama J, Wanninger T, Reyna RA, Stevenson HL, Peng BH, et al. The roles of XJ13 and XJ44-specific mutations within the Candid #1 GPC in Junin virus attenuation. *Front Immunol.* 2023;14:1172792.
250. Wu G, Yan S. Reasoning of spike glycoproteins being more vulnerable to mutations among 158 coronavirus proteins from different species. *J Mol Model.* 2005;11(1):8-16.
251. Seregin AV, Yun NE, Miller M, Aronson J, Smith JK, Walker AG, et al. The glycoprotein precursor gene of Junin virus determines the virulence of the Romero strain and the attenuation of the Candid #1 strain in a representative animal model of Argentine hemorrhagic fever. *J Virol.* 2015;89(11):5949-56.
252. Kranzusch PJ, Schenk AD, Rahmeh AA, Radoshitzky SR, Bavari S, Walz T, et al. Assembly of a functional Machupo virus polymerase complex. *Proc Natl Acad Sci U S A.* 2010;107(46):20069-74.

253. Wilda M, Lopez N, Casabona JC, Franze-Fernandez MT. Mapping of the tacaribe arenavirus Z-protein binding sites on the L protein identified both amino acids within the putative polymerase domain and a region at the N terminus of L that are critically involved in binding. *J Virol.* 2008;82(22):11454-60.
254. Brunotte L, Lelke M, Hass M, Kleinstauber K, Becker-Ziaja B, Gunther S. Domain structure of Lassa virus L protein. *J Virol.* 2011;85(1):324-33.
255. Vieth S, Torda AE, Asper M, Schmitz H, Günther S. Sequence analysis of L RNA of Lassa virus. *Virology.* 2004;318(1):153-68.
256. Poch O, Blumberg BM, Bougueleret L, Tordo N. Sequence comparison of five polymerases (L proteins) of unsegmented negative-strand RNA viruses: theoretical assignment of functional domains. *J Gen Virol.* 1990;71 (Pt 5):1153-62.
257. Poch O, Sauvaget I, Delarue M, Tordo N. Identification of four conserved motifs among the RNA-dependent polymerase encoding elements. *EMBO J.* 1989;8(12):3867-74.
258. Gerlach P, Malet H, Cusack S, Reguera J. Structural Insights into Bunyavirus Replication and Its Regulation by the vRNA Promoter. *Cell.* 2015;161(6):1267-79.
259. Muller R, Poch O, Delarue M, Bishop DH, Bouloy M. Rift Valley fever virus L segment: correction of the sequence and possible functional role of newly identified regions conserved in RNA-dependent polymerases. *J Gen Virol.* 1994;75 (Pt 6):1345-52.
260. Vieth S, Torda AE, Asper M, Schmitz H, Gunther S. Sequence analysis of L RNA of Lassa virus. *Virology.* 2004;318(1):153-68.
261. Morin B, Coutard B, Lelke M, Ferron F, Kerber R, Jamal S, et al. The N-terminal domain of the arenavirus L protein is an RNA endonuclease essential in mRNA transcription. *PLoS Pathog.* 2010;6(9):e1001038.
262. Fuller-Pace FV, Southern PJ. Detection of virus-specific RNA-dependent RNA polymerase activity in extracts from cells infected with lymphocytic choriomeningitis virus: in vitro synthesis of full-length viral RNA species. *J Virol.* 1989;63(5):1938-44.
263. Hass M, Lelke M, Busch C, Becker-Ziaja B, Gunther S. Mutational evidence for a structural model of the Lassa virus RNA polymerase domain and identification of two residues, Gly1394 and Asp1395, that are critical for transcription but not replication of the genome. *J Virol.* 2008;82(20):10207-17.
264. Lelke M, Brunotte L, Busch C, Gunther S. An N-terminal region of Lassa virus L protein plays a critical role in transcription but not replication of the virus genome. *J Virol.* 2010;84(4):1934-44.
265. Kang H, Cong J, Wang C, Ji W, Xin Y, Qian Y, et al. Structural basis for recognition and regulation of arenavirus polymerase L by Z protein. *Nat Commun.* 2021;12(1):4134.
266. Ma J, Zhang S, Zhang X. Structure of Machupo virus polymerase in complex with matrix protein Z. *Nat Commun.* 2021;12(1):6163.
267. Xu X, Peng R, Peng Q, Wang M, Xu Y, Liu S, et al. Cryo-EM structures of Lassa and Machupo virus polymerases complexed with cognate regulatory Z proteins identify targets for antivirals. *Nat Microbiol.* 2021;6(7):921-31.
268. Sanchez AB, de la Torre JC. Genetic and biochemical evidence for an oligomeric structure of the functional L polymerase of the prototypic arenavirus lymphocytic choriomeningitis virus. *J Virol.* 2005;79(11):7262-8.
269. Young PR, Howard CR. Ribonucleoprotein complexes associated with virions of Pichinde virus and Pichinde virus-infected cells. *Med Microbiol Immunol.* 1986;175(2-3):79-83.
270. Vezza AC, Clewley JP, Gard GP, Abraham NZ, Compans RW, Bishop DH. Virion RNA species of the arenaviruses Pichinde, Tacaribe, and Tamiami. *J Virol.* 1978;26(2):485-97.

271. Holm T, Kopicki JD, Busch C, Olschewski S, Rosenthal M, Uetrecht C, et al. Biochemical and structural studies reveal differences and commonalities among cap-snatching endonucleases from segmented negative-strand RNA viruses. *J Biol Chem*. 2018;293(51):19686-98.
272. Raju R, Raju L, Hacker D, Garcin D, Compans R, Kolakofsky D. Nontemplated bases at the 5' ends of Tacaribe virus mRNAs. *Virology*. 1990;174(1):53-9.
273. Polyak SJ, Zheng S, Harnish DG. 5' termini of Pichinde arenavirus S RNAs and mRNAs contain nontemplated nucleotides. *J Virol*. 1995;69(5):3211-5.
274. Lehmann M, Pahlmann M, Jerome H, Busch C, Lelke M, Gunther S. Role of the C terminus of Lassa virus L protein in viral mRNA synthesis. *J Virol*. 2014;88(15):8713-7.
275. Garcin D, Kolakofsky D. A novel mechanism for the initiation of Tacaribe arenavirus genome replication. *J Virol*. 1990;64(12):6196-203.
276. Auperin DD, Galinski M, Bishop DH. The sequences of the N protein gene and intergenic region of the S RNA of pichinde arenavirus. *Virology*. 1984;134(1):208-19.
277. Tortorici MA, Albarino CG, Posik DM, Ghiringhelli PD, Lozano ME, Rivera Pomar R, et al. Arenavirus nucleocapsid protein displays a transcriptional antitermination activity in vivo. *Virus Res*. 2001;73(1):41-55.
278. Cornu TI, Feldmann H, de la Torre JC. Cells expressing the RING finger Z protein are resistant to arenavirus infection. *J Virol*. 2004;78(6):2979-83.
279. Jacamo R, Lopez N, Wilda M, Franze-Fernandez MT. Tacaribe virus Z protein interacts with the L polymerase protein to inhibit viral RNA synthesis. *J Virol*. 2003;77(19):10383-93.
280. D'Antuono A, Loureiro ME, Foscaldi S, Marino-Buslje C, Lopez N. Differential contributions of tacaribe arenavirus nucleoprotein N-terminal and C-terminal residues to nucleocapsid functional activity. *J Virol*. 2014;88(11):6492-505.
281. Iwasaki M, Ngo N, Cubitt B, de la Torre JC. Efficient Interaction between Arenavirus Nucleoprotein (NP) and RNA-Dependent RNA Polymerase (L) Is Mediated by the Virus Nucleocapsid (NP-RNA) Template. *J Virol*. 2015;89(10):5734-8.
282. Kerber R, Rieger T, Busch C, Flatz L, Pinschewer DD, Kummerer BM, et al. Cross-species analysis of the replication complex of Old World arenaviruses reveals two nucleoprotein sites involved in L protein function. *J Virol*. 2011;85(23):12518-28.
283. Pontremoli C, Forni D, Cagliani R, Pozzoli U, Riva S, Bravo IG, et al. Evolutionary analysis of Old World arenaviruses reveals a major adaptive contribution of the viral polymerase. *Mol Ecol*. 2017;26(19):5173-88.
284. Forni D, Pontremoli C, Pozzoli U, Clerici M, Cagliani R, Sironi M. Ancient Evolution of Mammarenaviruses: Adaptation via Changes in the L Protein and No Evidence for Host-Virus Codivergence. *Genome Biol Evol*. 2018;10(3):863-74.
285. Grande-Perez A, Martin V, Moreno H, de la Torre JC. Arenavirus Quasispecies and Their Biological Implications. *Curr Top Microbiol Immunol*. 2016;392:231-76.
286. Golden JW, Hammerbeck CD, Mucker EM, Brocato RL. Animal Models for the Study of Rodent-Borne Hemorrhagic Fever Viruses: Arenaviruses and Hantaviruses. *Biomed Res Int*. 2015;2015:793257.
287. Matloubian M, Kolhekar SR, Somasundaram T, Ahmed R. Molecular determinants of macrophage tropism and viral persistence: importance of single amino acid changes in the polymerase and glycoprotein of lymphocytic choriomeningitis virus. *J Virol*. 1993;67(12):7340-9.
288. Bergthaler A, Flatz L, Hegazy AN, Johnson S, Horvath E, Lohning M, et al. Viral replicative capacity is the primary determinant of lymphocytic choriomeningitis virus persistence and immunosuppression. *Proc Natl Acad Sci U S A*. 2010;107(50):21641-6.
289. Ng CT, Sullivan BM, Oldstone MB. The role of dendritic cells in viral persistence. *Curr Opin Virol*. 2011;1(3):160-6.

290. Sullivan BM, Teijaro JR, de la Torre JC, Oldstone MB. Early virus-host interactions dictate the course of a persistent infection. *PLoS Pathog.* 2015;11(1):e1004588.
291. McLay L, Lan S, Ansari A, Liang Y, Ly H. Identification of virulence determinants within the L genomic segment of the pichinde arenavirus. *J Virol.* 2013;87(12):6635-43.
292. Zhang LK, Xin QL, Zhu SL, Wan WW, Wang W, Xiao G. Activation of the RLR/MAVS Signaling Pathway by the L Protein of Mopeia Virus. *J Virol.* 2016;90(22):10259-70.
293. Khamina K, Lercher A, Caldera M, Schliehe C, Vilagos B, Sahin M, et al. Characterization of host proteins interacting with the lymphocytic choriomeningitis virus L protein. *PLoS Pathog.* 2017;13(12):e1006758.
294. Fang J, Pietzsch C, Witwit H, Tsapralis G, Crynen G, Cho KF, et al. Proximity interactome analysis of Lassa polymerase reveals eRF3a/GSPT1 as a druggable target for host-directed antivirals. *Proc Natl Acad Sci U S A.* 2022;119(30):e2201208119.
295. Pontremoli C, Forni D, Cagliani R, Pozzoli U, Riva S, Bravo IG, et al. Evolutionary analysis of Old World arenaviruses reveals a major adaptive contribution of the viral polymerase. *Molecular Ecology.* 2017;26(19):5173-88.
296. Pontremoli C, Forni D, Cagliani R, Sironi M. Analysis of Reptarenavirus genomes indicates different selective forces acting on the S and L segments and recent expansion of common genotypes. *Infection, Genetics and Evolution.* 2018;64(May):212-8.
297. Kouba T, Vogel D, Thorkelsson SR, Quemin ERJ, Williams HM, Milewski M, et al. Conformational changes in Lassa virus L protein associated with promoter binding and RNA synthesis activity. *Nat Commun.* 2021;12(1):7018.
298. Peng R, Xu X, Jing J, Wang M, Peng Q, Liu S, et al. Structural insight into arenavirus replication machinery. *Nature.* 2020;579(7800):615-9.
299. Parvathy ST, Udayasuriyan V, Bhadana V. Codon usage bias. *Mol Biol Rep.* 2022;49(1):539-65.
300. Zhou Z, Dang Y, Zhou M, Li L, Yu CH, Fu J, et al. Codon usage is an important determinant of gene expression levels largely through its effects on transcription. *Proc Natl Acad Sci U S A.* 2016;113(41):E6117-E25.
301. Quax TE, Claassens NJ, Soll D, van der Oost J. Codon Bias as a Means to Fine-Tune Gene Expression. *Mol Cell.* 2015;59(2):149-61.
302. Palidwor GA, Perkins TJ, Xia X. A general model of codon bias due to GC mutational bias. *PLoS One.* 2010;5(10):e13431.
303. Plotkin JB, Kudla G. Synonymous but not the same: the causes and consequences of codon bias. *Nat Rev Genet.* 2011;12(1):32-42.
304. Sueoka N, Kawanishi Y. DNA G+C content of the third codon position and codon usage biases of human genes. *Gene.* 2000;261(1):53-62.
305. Rosenberg MS, Subramanian S, Kumar S. Patterns of transitional mutation biases within and among mammalian genomes. *Mol Biol Evol.* 2003;20(6):988-93.
306. Duchene S, Ho SY, Holmes EC. Declining transition/transversion ratios through time reveal limitations to the accuracy of nucleotide substitution models. *BMC Evol Biol.* 2015;15:36.
307. Zhang Z, Gerstein M. Patterns of nucleotide substitution, insertion and deletion in the human genome inferred from pseudogenes. *Nucleic Acids Res.* 2003;31(18):5338-48.
308. Jiang C, Zhao Z. Mutational spectrum in the recent human genome inferred by single nucleotide polymorphisms. *Genomics.* 2006;88(5):527-34.
309. Pauly MD, Procaro MC, Lauring AS. A novel twelve class fluctuation test reveals higher than expected mutation rates for influenza A viruses. *Elife.* 2017;6.

310. Zhou M, Guo J, Cha J, Chae M, Chen S, Barral JM, et al. Non-optimal codon usage affects expression, structure and function of clock protein FRQ. *Nature*. 2013;495(7439):111-5.
311. Chaney JL, Clark PL. Roles for Synonymous Codon Usage in Protein Biogenesis. *Annu Rev Biophys*. 2015;44:143-66.
312. Liu Y, Yang Q, Zhao F. Synonymous but Not Silent: The Codon Usage Code for Gene Expression and Protein Folding. *Annu Rev Biochem*. 2021;90:375-401.
313. Tao P, Dai L, Luo M, Tang F, Tien P, Pan Z. Analysis of synonymous codon usage in classical swine fever virus. *Virus Genes*. 2009;38(1):104-12.
314. Belalov IS, Lukashev AN. Causes and implications of codon usage bias in RNA viruses. *PLoS One*. 2013;8(2):e56642.
315. Lauring AS. Within-Host Viral Diversity: A Window into Viral Evolution. *Annu Rev Virol*. 2020;7(1):63-81.
316. Wong EH, Smith DK, Rabadan R, Peiris M, Poon LL. Codon usage bias and the evolution of influenza A viruses. *Codon Usage Biases of Influenza Virus*. *BMC Evol Biol*. 2010;10:253.
317. Singh NK, Tyagi A. A detailed analysis of codon usage patterns and influencing factors in Zika virus. *Arch Virol*. 2017;162(7):1963-73.
318. Ata G, Wang H, Bai H, Yao X, Tao S. Edging on Mutational Bias, Induced Natural Selection From Host and Natural Reservoirs Predominates Codon Usage Evolution in Hantaan Virus. *Front Microbiol*. 2021;12:699788.
319. Noor F, Ashfaq UA, Bakar A, Qasim M, Masoud MS, Alshammari A, et al. Identification and characterization of codon usage pattern and influencing factors in HFRS-causing hantaviruses. *Front Immunol*. 2023;14:1131647.
320. Mogro EG, Bottero D, Lozano MJ. Analysis of SARS-CoV-2 synonymous codon usage evolution throughout the COVID-19 pandemic. *Virology*. 2022;568:56-71.
321. Malhotra H, Kumar A. Codon usage signatures in Sabia and Chapare for host adaptation. *Bioinformatics*. 2021;17(10):891-8.
322. Sakabe S, Cubitt B, Martinez-Sobrido L, de la Torre JC. Molecular Engineering of a Mammarenavirus with Unbreachable Attenuation. *J Virol*. 2023;97(1):e0138522.
323. Cheng BY, Ortiz-Riano E, Nogales A, de la Torre JC, Martinez-Sobrido L. Development of live-attenuated arenavirus vaccines based on codon deoptimization. *J Virol*. 2015;89(7):3523-33.
324. Kustin T, Stern A. Biased Mutation and Selection in RNA Viruses. *Mol Biol Evol*. 2021;38(2):575-88.
325. Butt AM, Nasrullah I, Tong Y. Genome-wide analysis of codon usage and influencing factors in chikungunya viruses. *PLoS One*. 2014;9(3):e90905.
326. Kumar U, Khandia R, Singhal S, Puranik N, Tripathi M, Pateriya AK, et al. Insight into Codon Utilization Pattern of Tumor Suppressor Gene EPB41L3 from Different Mammalian Species Indicates Dominant Role of Selection Force. *Cancers (Basel)*. 2021;13(11).
327. Sharp PM, Li WH. The codon Adaptation Index--a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Res*. 1987;15(3):1281-95.
328. Mueller S, Papamichail D, Coleman JR, Skiena S, Wimmer E. Reduction of the rate of poliovirus protein synthesis through large-scale codon deoptimization causes attenuation of viral virulence by lowering specific infectivity. *J Virol*. 2006;80(19):9687-96.
329. Rahman SU, Yao X, Li X, Chen D, Tao S. Analysis of codon usage bias of Crimean-Congo hemorrhagic fever virus and its adaptation to hosts. *Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases*. 2018;58:1-16.

330. Khandia R, Singhal S, Kumar U, Ansari A, Tiwari R, Dhama K, et al. Analysis of Nipah Virus Codon Usage and Adaptation to Hosts. *Front Microbiol.* 2019;10:886.
331. Tort FL, Castells M, Cristina J. A comprehensive analysis of genome composition and codon usage patterns of emerging coronaviruses. *Virus Res.* 2020;283:197976.
332. Shackelton LA, Parrish CR, Holmes EC. Evolutionary basis of codon usage and nucleotide composition bias in vertebrate DNA viruses. *J Mol Evol.* 2006;62(5):551-63.
333. Hu JS, Wang QQ, Zhang J, Chen HT, Xu ZW, Zhu L, et al. The characteristic of codon usage pattern and its evolution of hepatitis C virus. *Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases.* 2011;11(8):2098-102.
334. Tsai CJ, Sauna ZE, Kimchi-Sarfaty C, Ambudkar SV, Gottesman MM, Nussinov R. Synonymous mutations and ribosome stalling can lead to altered folding pathways and distinct minima. *J Mol Biol.* 2008;383(2):281-91.
335. Sueoka N. Directional mutation pressure and neutral molecular evolution. *Proc Natl Acad Sci U S A.* 1988;85(8):2653-7.
336. Laguens M, Chambo JG, Laguens RP. In vivo replication of pathogenic and attenuated strains of Junin virus in different cell populations of lymphatic tissue. *Infect Immun.* 1983;41(3):1279-83.
337. Laguens RM, Chambo JG, Laguens RP. Splenic dendritic cells and Junin virus. *Med Microbiol Immunol.* 1986;175(2-3):187-9.
338. Gomez RM, Pozner RG, Lazzari MA, D'Atri LP, Negrotto S, Chudzinski-Tavassi AM, et al. Endothelial cell function alteration after Junin virus infection. *Thrombosis and Haemostasis.* 2003;90(2):326-33.
339. Lander HM, Grant AM, Albrecht T, Hill T, Peters CJ. Endothelial cell permeability and adherens junction disruption induced by Junin virus infection. *American Journal of Tropical Medicine and Hygiene.* 2014;90(6):993-1002.
340. Kunz S. The role of the vascular endothelium in arenavirus haemorrhagic fevers. *Thromb Haemost.* 2009;102(6):1024-9.
341. Meyer BJ, Schmaljohn CS. Persistent hantavirus infections: characteristics and mechanisms. *Trends Microbiol.* 2000;8(2):61-7.
342. Stephan BI, Lozano ME, Goni SE. Watching every step of the way: junin virus attenuation markers in the vaccine lineage. *Curr Genomics.* 2013;14(7):415-24.
343. Hallam SJ, Manning JT, Maruyama J, Seregin A, Huang C, Walker DH, et al. A single mutation (V64G) within the RING Domain of Z attenuates Junin virus. *PLoS Negl Trop Dis.* 2020;14(9):e0008555.
344. Agnese G. Entre controversias científico-médicas y movilizaciones populares. *Población epidémica y vacunas contra la fiebre hemorrágica argentina 1958-1990.* *Asclepio.* 2013;65:p011.
345. Goni SE, Stephan BI, Iserle JA, Contigiani MS, Lozano ME, Tenorio A. Viral diversity of Junin virus field strains. *Virus Res.* 2011;160(1-2):150-8.
346. Al-Masri C, Trozzi F, Lin SH, Tran O, Sahni N, Patek M, et al. Investigating the conformational landscape of AlphaFold2-predicted protein kinase structures. *Bioinform Adv.* 2023;3(1):vbad129.
347. Del Alamo D, Sala D, McHaourab HS, Meiler J. Sampling alternative conformational states of transporters and receptors with AlphaFold2. *Elife.* 2022;11.
348. Stein RA, McHaourab HS. SPEACH_AF: Sampling protein ensembles and conformational heterogeneity with AlphaFold2. *PLoS Comput Biol.* 2022;18(8):e1010483.
349. Wayment-Steele HK, Ovchinnikov S, Colwell L, Kern DJB. Prediction of multiple conformational states by combining sequence clustering with AlphaFold2. 2022:2022.10. 17.512570.
350. Zhang Y, Li L, Liu X, Dong S, Wang W, Huo T, et al. Crystal structure of Junin virus nucleoprotein. *J Gen Virol.* 2013;94(Pt 10):2175-83.

351. Ng WM, Sahin M, Krumm SA, Seow J, Zeltina A, Harlos K, et al. Contrasting Modes of New World Arenavirus Neutralization by Immunization-Elicited Monoclonal Antibodies. *mBio*. 2022;13(2):e0265021.
352. Loureiro ME, Wilda M, Levingston Macleod JM, D'Antuono A, Foscaldi S, Buslje CM, et al. Molecular Determinants of Arenavirus Z Protein Homo-Oligomerization and L Polymerase Binding. *Journal of Virology*. 2011;85(23):12304-14.
353. Cornu TI, de la Torre JC. RING Finger Z Protein of Lymphocytic Choriomeningitis Virus (LCMV) Inhibits Transcription and RNA Replication of an LCMV S-Segment Minigenome. *Journal of Virology*. 2001;75(19):9415-26.
354. Levingston Macleod JM, D'Antuono A, Loureiro ME, Casabona JC, Gomez GA, Lopez N. Identification of Two Functional Domains within the Arenavirus Nucleoprotein. *Journal of Virology*. 2011;85(5):2012-23.
355. York J, Nunberg JH. Epistatic Interactions within the Junin Virus Envelope Glycoprotein Complex Provide an Evolutionary Barrier to Reversion in the Live-Attenuated Candid#1 Vaccine. *J Virol*. 2018;92(1).
356. Wilda M, Lopez N, Casabona JC, Franze-Fernandez MT. Mapping of the Tacaribe Arenavirus Z-Protein Binding Sites on the L Protein Identified both Amino Acids within the Putative Polymerase Domain and a Region at the N Terminus of L That Are Critically Involved in Binding. *Journal of Virology*. 2008;82(22):11454-60.
357. Honorato RV, Koukos PI, Jimenez-Garcia B, Tsaregorodtsev A, Verlato M, Giachetti A, et al. Structural Biology in the Clouds: The WeNMR-EOSC Ecosystem. *Front Mol Biosci*. 2021;8:729513.
358. van Zundert GCP, Rodrigues J, Trellet M, Schmitz C, Kastrius PL, Karaca E, et al. The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. *J Mol Biol*. 2016;428(4):720-5.
359. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, et al. Scalable molecular dynamics with NAMD. *J Comput Chem*. 2005;26(16):1781-802.
360. Emonet SF, Seregin AV, Yun NE, Poussard AL, Walker AG, de la Torre JC, et al. Rescue from cloned cDNAs and in vivo characterization of recombinant pathogenic Romero and live-attenuated Candid #1 strains of Junin virus, the causative agent of Argentine hemorrhagic fever disease. *J Virol*. 2011;85(4):1473-83.
361. Sayers EW, Bolton EE, Brister JR, Canese K, Chan J, Comeau DC, et al. Database resources of the national center for biotechnology information. *Nucleic Acids Res*. 2022;50(D1):D20-D6.
362. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an open-source platform for biological-image analysis. *Nat Methods*. 2012;9(7):676-82.
363. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. *J Mol Graph*. 1996;14(1):33-8, 27-8.
364. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem*. 2004;25(13):1605-12.
365. Okonechnikov K, Golosova O, Fursov M, team U. Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics*. 2012;28(8):1166-7.
366. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol Biol Evol*. 2018;35(6):1547-9.
367. Lemoine F, Correia D, Lefort V, Doppelt-Azeroual O, Mareuil F, Cohen-Boulakia S, et al. NGPhylogeny.fr: new generation phylogenetic services for non-specialists. *Nucleic Acids Res*. 2019;47(W1):W260-W5.
368. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021;596(7873):583-9.
369. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res*. 2018;46(W1):W296-W303.

370. Studer G, Tauriello G, Bienert S, Biasini M, Johner N, Schwede T. ProMod3-A versatile homology modelling toolbox. *PLoS Comput Biol.* 2021;17(1):e1008667.
371. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: A program to check the stereochemical quality of protein structures. *J Appl Cryst.* 1993;26:283-91.
372. Luthy R, Bowie JU, Eisenberg D. Assessment of protein models with three-dimensional profiles. *Nature.* 1992;356(6364):83-5.
373. Robert X, Gouet P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* 2014;42(Web Server issue):W320-4.
374. Laimer J, Hofer H, Fritz M, Wegenkittl S, Lackner P. MAESTRO--multi agent stability prediction upon point mutations. *BMC Bioinformatics.* 2015;16:116.
375. Puigbo P, Bravo IG, Garcia-Vallve S. CAIcal: a combined set of tools to assess codon usage adaptation. *Biol Direct.* 2008;3:38.
376. Team RC. R: A Language and Environment for Statistical Computing. 2023.
377. Le S, Josse J, Husson F. FactoMineR: A Package for Multivariate Analysis. *Journal of Statistical Software.* 2008;25:1-18.
378. Wei T, Simko V. R package 'corrplot': Visualization of a Correlation Matrix. 2021.