

Identificación de ARNs no codificantes largos que forman parte de los complejos ribonucleoprotéicos del ARNg del Virus de la Inmunodeficiencia Humana de tipo 1

Tesis

Entregada A La

Universidad De Chile

En Cumplimiento Parcial De Los Requisitos

Para Optar Al Grado De

Doctora en Ciencias con Mención en Microbiología

Facultad De Ciencias

Por

Paulina Andrea Aguilera Cortés

Diciembre, 2021

Director de Tesis: Dr. Ricardo Soto Rifo

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

Se informa a la Escuela de Postgrado de la Facultad de Ciencias que la Tesis de Doctorado presentada por la candidata:

Paulina Andrea Aguilera Cortés

Ha sido aprobada por la comisión de Evaluación de la tesis como requisito para optar al grado de Doctora en Ciencias con Mención en Microbiología en el examen de Defensa Privada de Tesis rendido el día

Director de Tesis:

Dr. Ricardo Soto Rifo

.....

Comisión de Evaluación de la Tesis

Dr. Eugenio Spencer

.....

Dr. Omar Orellana

.....

Dra. Ana María Sandino

.....

Dr. Marcelo López

.....

ÍNDICE

1	INTRODUCCIÓN	x
1.1	CAPÍTULO 1: Virus de la Inmunodeficiencia Humana de tipo 1 (VIH-1).....	1
1.1.1	Características generales de VIH-1	1
1.1.2	Ciclo replicativo de VIH-1.	4
1.2	CAPÍTULO 2: Complejos Ribonucleoproteicos (RNPs).....	9
1.2.1	Composición de los complejos Ribonucleoproteicos.	10
1.2.2	Complejos Ribonucleoproteicos de VIH-1.	11
1.3	CAPÍTULO 3: ARNs no codificantes.	14
1.3.1	ARNnc largos: Clasificación y características generales	15
2	HIPÓTESIS Y OBJETIVOS.....	22
2.1	Hipótesis	22
2.2	Objetivo general	22
2.3	Objetivos específicos	22
3	MATERIALES Y MÉTODOS	23
3.1	Plasmidios.....	23
3.2	Cultivo Celular.....	23
3.3	Activación de células J-Lat.	24
3.4	Preparación de partículas virales.	24
3.5	Infección de células.....	24
3.6	Extracción de ARN total.	25
3.7	RT-PCR en tiempo real.	25
3.8	Extracción de proteínas totales.	26
3.9	Western blot.....	27
3.10	Preparación de la sonda biotinilada.....	28
3.11	Purificación de ARN Antisentido (RAP).	30
3.12	Sobreexpresión de HAR1A.	31
4	RESULTADOS.....	33
4.1	OE1. Aislar los mRNPs del ARNg de VIH-1.	33
4.1.1	Preparación de sonda antisentido biotinilada.	33
4.1.2	Captura del ARNg de VIH-1 a partir de células J-Lat.....	35
4.1.3	Captura del ARNg de VIH-1 a partir de células Jurkat infectadas.....	40

4.2	OE2. Identificar y seleccionar ARNnc largos asociados al ARNg de VIH-1.	42
4.2.1	Identificación de ARNs capturados a partir de células J-Lat activadas.	42
4.2.2	Identificación de ARNs capturados a partir de células Jurkat infectadas y sin infectar.	44
4.2.3	Re-análisis de datos de secuenciación.	46
4.2.4	Control de calidad y limpieza de las secuencias.....	48
4.2.5	Alineamiento contra ARNr y ARNt.....	48
4.2.6	Alineamiento contra el genoma humano (hg38).	50
4.2.7	Alineamiento o mapeo contra el genoma de VIH-1 (pNL4.3).....	51
4.2.8	Visualización del alineamiento contra el genoma de VIH-1.	52
4.2.9	Cuantificación y enriquecimiento de secuencias alineadas contra el genoma humano.....	52
4.2.10	Visualización de secuencias alineadas contra el genoma humano.	56
4.2.11	Diseño de partidores y validación mediante RT-qPCR.....	56
4.3	OE3. Determinar el efecto de un ARNnc largo seleccionado sobre la replicación viral.....	58
4.3.1	Evaluación de los niveles de expresión relativa de los ARNnc largos seleccionados en diferentes líneas celulares.....	58
4.3.2	Efecto de HAR1A sobre los niveles de Gag y del ARNg.	63
5	DISCUSIÓN	67
6	CONCLUSIONES.....	74
7	REFERENCIAS.....	75

LISTA DE TABLAS

Tabla 1. ARNnc largos que afectan el metabolismo de VIH-1.	21
Tabla 2. Plasmidios utilizados en este trabajo.	23
Tabla 3. Lista de partidores utilizados en este trabajo	26
Tabla 4. Reacción para la transcripción in vitro	28
Tabla 5. ARNnc largos intergénicos más enriquecidos a partir de la captura de ARNg en células J-Lat	44
Tabla 6. ARNnc largos intergénicos más enriquecidos a partir de la captura de ARNg en células Jurkat	46
Tabla 7. Porcentaje de alineamiento contra una base de datos de ARNr y ARNt.	50
Tabla 8. Porcentaje de alineamiento contra el genoma humano (hg38)	51
Tabla 9. Porcentaje de alineamiento contra el genoma de VIH-1 (pNL4.3ΔEnv)	52
Tabla 10. ARNnc largos intergénicos seleccionados de acuerdo al análisis de expresión diferencial.	56

LISTA DE FIGURAS

Figura 1. Ciclo replicativo de VIH-1	4
Figura 2. Clasificación de los ARNnc largos según su locus genómico	16
Figura 3. Regulación postranscripcional de los ARNnc largos.....	19
Figura 4. Preparación de la sonda de ARN antisentido de VIH-1	34
Figura 5. Cinética de activación de células J-Lat.....	35
Figura 6. Purificación de ARN antisentido (RAP).....	37
Figura 7. Ensayo piloto de captura del ARNg de VIH-1	38
Figura 8. Enriquecimiento de ARNg y ARNnc	39
Figura 9. Cinética de infección de células Jurkat.....	41
Figura 10. Identificación de ARNs asociados al ARNg en células J-Lat.....	43
Figura 11. Identificación de ARNs asociados al ARNg en células Jurkat infectadas con pseudotipos de VIH-1	45
Figura 12. Mapa de flujo para el análisis de datos de secuenciación de ARN	47
Figura 13. Control de calidad de una muestra representativa.....	49
Figura 14. Visualización del alineamiento contra el genoma de VIH-1.....	53
Figura 15. Análisis de expresión diferencial en un único modelamiento	54
Figura 16. Análisis de datos para expresión diferencial de muestras mock e infectadas por separado	55
Figura 17. Visualización de dos de los ARNnc largos intergénicos seleccionados de acuerdo al análisis de expresión diferencial.....	57
Figura 18. Infección de 6 líneas celulares diferentes con VIH-1.	59
Figura 19. Expresión relativa de ARNnc largos en distintas líneas celulares	60
Figura 20. Expresión relativa de Linc00938 en células infectadas.....	61
Figura 21. Expresión relativa de Inc-ISL2-1 en células infectadas	62
Figura 22. Expresión relativa de HAR1A en células infectadas	63
Figura 23. Efecto de la sobreexpresión de HAR1A sobre los niveles de la proteína viral Gag...	65
Figura 24. Efecto de la sobreexpresión de HAR1A sobre los niveles de ARNg de VIH-1	66

LISTA DE ABREVIATURAS

Abreviaturas:

5'UTR	Región 5' no traducida
ADN	Ácido desoxirribonucleico
ADNc	Ácido desoxirribonucleico complementario
ARN	Ácido ribonucleico
ARNg	Ácido ribonucleico genómico
ARNm	Ácido ribonucleico mensajero
ARNnc	Ácido ribonucleico no codificante
ARNr	Ácido ribonucleico ribosomal
ARNt	Ácido ribonucleico de transferencia
CA	Cápside
CBP	Proteína de unión a Cap
CCR5	Receptor de quimioquinas C-C tipo 5
CD4	Cúmulo de diferenciación 4
CXCR4	Receptor de quimioquinas C-X-C tipo 4
DIS	Señal de dimerización
DMSO	Dimetil sulfóxido
EDTA	Ácido etilendiaminotetraacético
EGTA	Ácido etilenglicol-bis(2-aminoetiléter)-N,N,N'N'-tetraacético
EJC	Complejo de unión a exones
Env	Envoltura
ESE	Activador de <i>splicing</i> exónicos
ESS	Silenciador de <i>splicing</i> exónicos
Gag	Antígeno grupo específico
GAPDH	Gliceraldehído-3-fosfato deshidrogenasa
gp120	Glicoproteína 120

gp41	Glicoproteína 41
HCl	Ácido clorhídrico
hnRNP	Ribonucleoproteína nuclear heterogénea
HRP	Peroxidasa de rábano
IRES	Sitio interno de entrada de ribosomas
ISS	Silenciador de <i>splicing</i> intrónico
LiCl	Cloruro de Litio
LTR	Repetido largo terminal
MA	Matriz
mRNPs	Complejos Ribonucleoprotéicos
NaCl	Cloruro de sodio
NC	Nucleocápside
Nef	Factor regulador negativo
ORF	Marco de lectura abierto
PABP	Proteína de unión a poliA
PBS	Buffer fosfato salino
PEI	Polietilenimina
PMA	Forbol 12-miristato 13-acetato
Pol	Polimerasa
Psi	Señal de empaquetamiento
PVDF	Difluoruro de polivinilideno
qPCR	Reacción en cadena de la polimerasa cuantitativa
RAP	Purificación de ARN antisentido
RBP	Proteína de unión a ARN
Rev	Regulador del virión
rNTP	Ribonucleosido trifosfato
RT	Transcripción Reversa
SIDA	Síndrome de inmunodeficiencia adquirida

TAR	Elemento de respuesta de trans-activación
Tat	Trans-activador de la transcripción
TBE	Buffer Tris-Borato-EDTA
TBS-T	Tris buffer salino - Tween 20
TCEP	Tris-(2-carboxietil) fosfina, hidrocloreuro
TPM	Transcritos por millón
UV	Ultravioleta
Vif	Factor de infectividad viral
VIH-1	Virus de la inmunodeficiencia humana tipo 1
Vpr	Proteína viral R
Vpu	Proteína viral U
VSV	Virus de la estomatitis vesicular

Unidades:

µg	microgramo
µl	microlitro
Å	Angstrom
Da	Dalton
hpa	Horas post-activación
hpi	Horas post-infección
hpt	Horas post-transfección
Kb	Kilobase
kDa	Kilodalton
ml	Mililitro
mM	Milimolar
ng	Nanogramo
nm	Nanómetro
nt	Nucleótido
pb	Par de bases
rpm	Revoluciones por minuto

RESUMEN

Durante las etapas tardías del ciclo replicativo, el ARN genómico (ARNg) del Virus de la Inmunodeficiencia Humana de tipo 1 (VIH-1) se ensambla en complejos ribonucleoprotéicos (mRNPs) específicos cuya composición determina su destino celular. Una vez en el citoplasma, el ARNg de VIH-1 cumple dos funciones relevantes: 1) es el mensajero a partir del cual se sintetizan las proteínas virales Gag y Gag-Pol y 2) es el genoma empaquetado en las nuevas partículas virales sintetizadas, características que lo han hecho de particular interés. Diversas proteínas virales y celulares que forman parte de los mRNPs del ARNg de VIH-1 han sido identificadas, pero poco se sabe si ARN regulatorios como los ARN no codificantes largos (ARNnc largos) son componentes de estos complejos. A pesar de que los ARNnc largos han sido identificados como importantes reguladores transcripcionales y postranscripcionales, sólo se han reportado diez que participan en la replicación de VIH-1 y, hasta la fecha, no se había identificado ninguno que forme parte de los mRNPs del ARNg. El objetivo del presente trabajo fue implementar una estrategia de captura del ARNg de VIH-1, identificar ARNnc largos asociados a este para, finalmente, determinar si alguno de ellos afecta la replicación del virus. Se desarrolló una estrategia de captura utilizando una sonda antisentido que permitió la purificación de los mRNPs del ARNg, con un enriquecimiento de aproximadamente 6000 veces a partir de Linfocitos T CD4+ infectados. Mediante ARN-seq se identificaron cuatro ARNnc largos enriquecidos en los mRNPs capturados: Linc00938, Lnc-ISL2-1, HAR1A y Lnc-TMEM105-2. HAR1A mostró bajos niveles de expresión en diversas líneas celulares y se evaluó el efecto de su sobreexpresión sobre la replicación del virus. Sin embargo, no se observó ningún cambio en los niveles de Gag y ARNg intracelulares. Este es el primer reporte de la presencia de ARNnc largos que forman parte de los mRNPs del ARNg de VIH-1. El efecto de éstos sobre el ciclo replicativo del virus es una interrogante abierta a futuras investigaciones.

ABSTRACT

During the late steps of viral replication, the Human Immunodeficiency Virus type 1 (HIV-1) genomic RNA (gRNA) is assembled into specific ribonucleoprotein complexes (mRNPs) whose composition determines its cellular fate. Once in the cytoplasm, the HIV-1 gRNA fulfills two relevant functions: 1) as the messenger from which the viral proteins Gag and Gag-Pol are synthesized and 2) as the genome packaged into newly synthesized viral particles. Several viral and host proteins have been shown to be part of the mRNPs of the HIV-1 gRNA, but little is known if regulatory RNAs such as long non-coding RNAs (lncRNAs) are components of these complexes. Despite they appear as important transcriptional and post-transcriptional regulators, only ten have been related to HIV-1 replication. To date, no lncRNAs have been reported to be part of the mRNPs of HIV-1 gRNA. The aim of this work was to perform an HIV-1 gRNA capture strategy, identify lncRNAs that are associated with it and assess if any of them affect viral replication. An interactome capture strategy was developed using an antisense probe which allowed a 6000-fold enrichment of the HIV-1 gRNA from infected CD4+ T lymphocytes. By using RNA-seq, we identified four lncRNAs enriched in the capture samples: Linc00938, Lnc-ISL2-1, HAR1A and Lnc-TMEM105-2. HAR1A showed low levels of relative expression in various cell lines and the effect of its overexpression on HIV-1 replication was assessed. However, no change in intracellular Gag and mRNA levels was detected. This is the first report that shows the presence of lncRNAs in the mRNPs of HIV-1 mRNA. The effect of these lncRNAs in the replicative cycle of the virus is an interesting question that must be addressed in future investigations.

1 INTRODUCCIÓN

1.1 CAPÍTULO 1: Virus de la Inmunodeficiencia Humana de tipo 1 (VIH-1).

En 1981 se reportaron las primeras observaciones de manifestaciones clínicas del denominado Síndrome de Inmunodeficiencia Adquirida (SIDA) (Gottlieb et al., 1981) caracterizado por la adquisición de infecciones oportunistas, el desarrollo de ciertos tipos de cáncer y leucopenia. En 1983, a partir de un cultivo derivado de un nodo linfático de un paciente con linfadenopatía generalizada, se aisló un nuevo retrovirus humano (Barre-Sinoussi et al., 1983) que posteriormente fue confirmado como el agente causal del SIDA y se denominó Virus de la Inmunodeficiencia Humana de tipo 1 (VIH-1) (Barré-Sinoussi et al., 2013). Desde el inicio de la pandemia, aproximadamente 79 millones de personas se han infectado con VIH-1 y 36 millones han muerto por enfermedades relacionadas con el SIDA. En 2020, 37 millones de personas vivían con VIH-1 y, ese mismo año, 1,5 millones de personas nuevas se infectaron con el virus (UNAIDS, 2021). El VIH/SIDA no tiene cura, pero sí tratamiento, por lo que actualmente se considera una enfermedad crónica que requiere tratamiento de por vida.

1.1.1 Características generales de VIH-1

VIH-1 pertenece a la familia *Retroviridae*, cuyos miembros presentan partículas virales con envoltura que contienen un genoma de ARN de hebra simple, lineal, no segmentado y de polaridad positiva. En la fase temprana de su ciclo replicativo el genoma de los retrovirus sufre transcripción reversa, originando un intermediario de ADN de doble hebra que se integra en el genoma de la célula hospedera, estado que se denomina provirus (Coffin et al., 1997). En particular, VIH-1 pertenece al género *Lentivirus* (Chiu et al., 1985) cuyos miembros se caracterizan por generar enfermedades “lentas”, es decir, de largos periodos de incubación y lenta aparición de signos y síntomas relacionados con la infección. Además, los lentivirus se

diferencian de los otros retrovirus pues son capaces de infectar tanto células que se dividen, como células que no se dividen (Lewis & Emerman, 1994).

1.1.1.1 Organización genómica y proteínas virales.

VIH-1 tiene un genoma de aproximadamente 9 Kb que codifica para 15 proteínas a partir de 9 marcos de lectura abiertos. El genoma de todos los retrovirus contiene los genes *gag*, *pol* y *env*, que codifican proteínas estructurales y las enzimas virales (Coffin et al., 1997). VIH-1 es un retrovirus complejo por lo que, además de estos 3 genes, en su genoma se encuentran genes que codifican proteínas reguladoras (genes *tat* y *rev*) y accesorias (genes *vif*, *vpr*, *vpu* y *nef*) (Cullen, 1991; Gallo et al., 1988; Strebel et al., 1988). Gag (p55) es una poliproteína de 55 kDa y es la proteína más abundante de las partículas virales. Una partícula viral contiene alrededor de 5000 moléculas de Gag (Briggs et al., 2004). Además de su rol estructural, cumple funciones fundamentales en la selección y empaquetamiento del genoma, así como en el ensamblaje de nuevas partículas virales. A partir de la proteólisis de Gag, se generan las proteínas: Matriz o MA (p17), Cápside o CA (p24), Nucleocápside o NC (p7) y p6^{gag} (Bell & Lever, 2013; Freed, 2015). Gag-Pol (p160) es una poliproteína producida gracias a un corrimiento en el marco de lectura en la región p6 de Gag (Jacks et al., 1987). A partir de la proteólisis de Gag-Pol se generan las mismas proteínas que en Gag: Matriz, Cápside y Nucleocápside. Además, se produce la proteína p6^{pol}, y las tres enzimas esenciales para la replicación del virus: Proteasa, Transcriptasa Reversa e Integrasa (Freed, 2015). Cabe mencionar que, tanto la proteólisis de Gag como de Gag-Pol es llevada a cabo por la Proteasa viral. Por otro lado, Env (gp160) es una glicoproteína que al ser procesada por proteasas celulares, origina a las proteínas de superficie gp120 y de transmembrana gp41 (Robey et al., 1985; Veronese et al., 1985), que permiten el reconocimiento y la entrada del virus a la célula hospedera. Tat y Rev son proteínas reguladoras, que cumplen roles en la activación de la transcripción (Rice & Mathews, 1988) y

en la exportación nuclear de transcritos virales (Malim et al., 1989), respectivamente. Por último, Vif, Vpu, Vpr y Nef son proteínas accesorias, que pueden ser dispensables para la infección *in vitro*, pero son relevantes para la replicación viral *in vivo*, pues contrarrestan factores de restricción (Sheehy et al., 2002), inducen el arresto celular (Jowett et al., 1995), permiten la infección en células que no se dividen (Popov et al., 1998) e inducen la evasión del sistema inmune (Collins et al., 1998; Schwartz et al., 1996), entre otras funciones (Malim & Emerman, 2008).

1.1.1.2 Partícula viral.

Las partículas virales de VIH-1 son relativamente heterogéneas en tamaño, con un diámetro que puede variar entre los 100-200 nm aproximadamente (Briggs et al., 2003; Wilk et al., 2001). Tal como se mencionó anteriormente, los viriones están envueltos por una bicapa lipídica de origen celular, desde la cual protruye la glicoproteína gp120/gp41 en forma de trímero (Wu et al., 2010). Dentro de la partícula madura, justo debajo de la envoltura, se encuentra la proteína de la Matriz. En el centro de la partícula está la cápside viral que tiene forma cónica con una geometría de tipo fullereno, formada por aproximadamente 250 hexámeros y 12 pentámeros de la proteína de la Cápside (Briggs et al., 2003; S. Li et al., 2000). Dentro de la cápside, hay dos copias de genoma viral recubierto por proteínas de la Nucleocápside. A su vez, las partículas virales contienen las enzimas Proteasa, Transcriptasa Reversa e Integrasa, las proteínas accesorias Vpr (Cohen et al., 1990), Vif (H. Liu et al., 1995) y Nef (Welker et al., 1996), así como también proteínas y factores de origen celular (Franke et al., 1994; Y. Huang et al., 1994).

1.1.2 Ciclo replicativo de VIH-1.

Al igual que otros parásitos intracelulares, VIH-1 se aprovecha de la maquinaria celular para realizar distintas etapas de su ciclo replicativo, el que se puede dividir en dos fases principales: temprana y tardía. La fase temprana abarca desde la unión del virus con los receptores de la célula hospedera, hasta la integración del ADN viral. La fase tardía, abarca desde la transcripción del genoma viral, hasta la liberación y maduración de nuevas partículas virales (Figura 1).

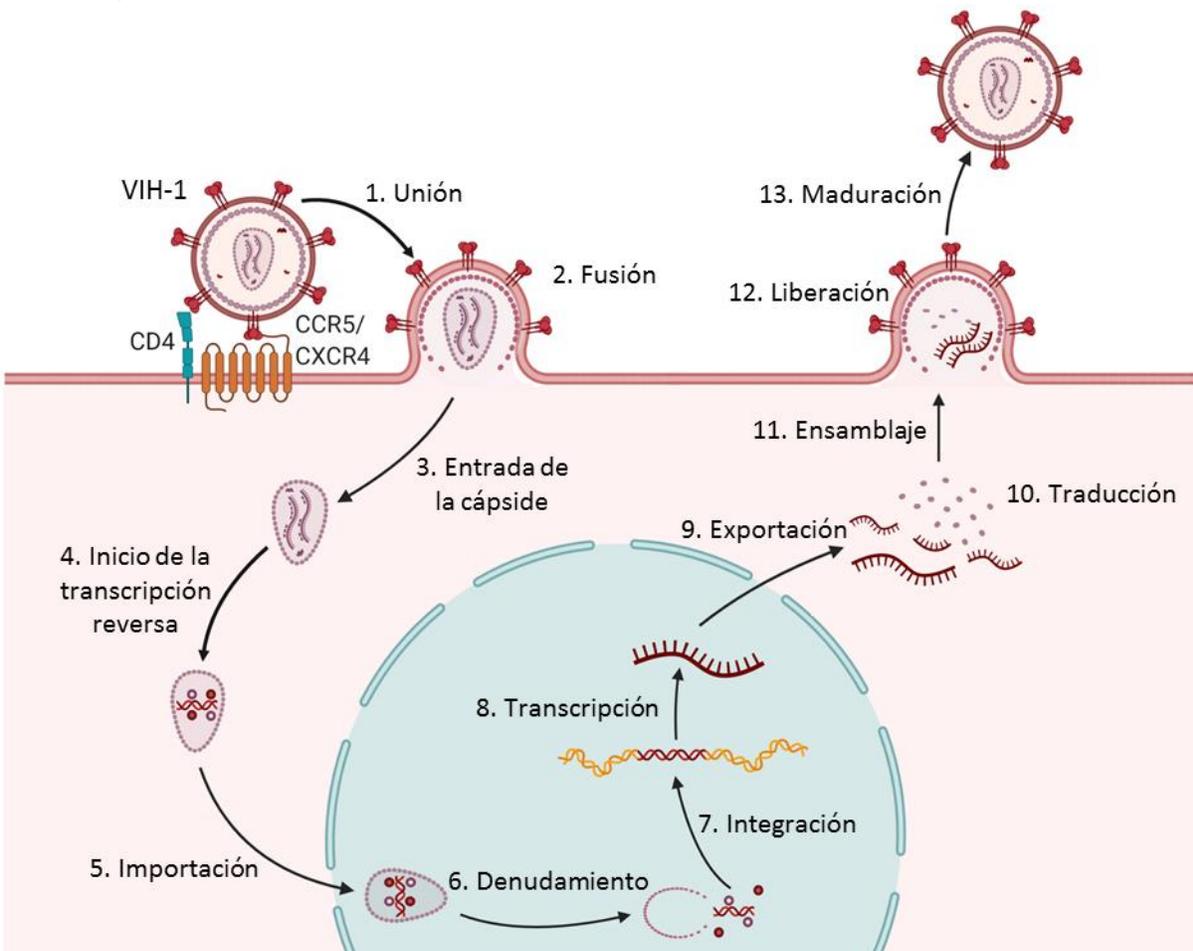


Figura 1. Ciclo replicativo de VIH-1. El ciclo replicativo de VIH-1 consta de las siguientes etapas: 1. Unión, 2. Fusión, 3. Entrada de la cápside, 4. Inicio de la transcripción reversa, 5. Importación, 6. Denudamiento, 7. Integración, 8. Transcripción, 9. Exportación, 10. Traducción, 11. Ensamblaje, 12. Liberación y 13. Maduración (Basado en Ramdas et al., 2020).

1.1.2.1 Fase temprana.

Entrada, transcripción reversa, desnudamiento e integración. El ciclo replicativo comienza con la unión entre la proteína de superficie del virus gp120 con el receptor celular CD4 (Maddon et al., 1986; McDougal et al., 1986). Esta primera interacción permite la posterior interacción con los correceptores CCR5 o CXCR4 (Choe et al., 1996; Feng et al., 1996). La unión con los correceptores induce la exposición de gp41 y su inserción en la membrana celular, en consecuencia, la membrana viral se fusiona con la membrana celular (Wilén et al., 2012). La presencia de CD4 y correceptores, por lo tanto, determina qué células son susceptibles a la infección por VIH-1, tales como los linfocitos T CD4+, monocitos, macrófagos, microglías y células dendríticas.

Luego de la fusión, la cápside viral ingresa al citoplasma y se inicia la transcripción reversa. El genoma viral es retrotranscrito por la Transcriptasa Reversa (RT). Esta proteína es un heterodímero (p66/p51) con dos actividades enzimáticas: polimerasa de ADN dependiente de ARN y RNasa H que permiten la síntesis de ADN a partir del ARN_v y la degradación del ARN en el híbrido ARN-ADN (Lightfoote et al., 1986; Veronese et al., 1986). Al igual que otras polimerasas de ADN, la RT necesita un partidor y un molde. En el caso de VIH-1, la RT utiliza el tRNA^{Lys3} como partidor, el cual se une a un sitio específico dentro de la región 5' no traducida (5'UTR) del ARN viral (Barat et al., 1989). La 5'UTR del ARN de VIH-1 es una región altamente estructurada que contiene diversos dominios o elementos estructurales que cumplen funciones en distintas etapas del ciclo replicativo del virus.

El mecanismo mediante el que la RT sintetiza ADN a partir de ARN es relativamente complejo, y consiste en una serie de pasos donde ocurren transferencias o “saltos” de la enzima, degradación del ARN utilizado como molde y elongación de las hebras de ADN (Hu & Hughes, 2012). Cabe mencionar que en los extremos del genoma viral hay regiones que son repetidos directos (R) y también regiones únicas denominadas U5 en el extremo 5' y U3 en el

extremo 3'. Como resultado final de la retrotranscripción, se genera un ADN de doble hebra con una secuencia idéntica en los extremos, U3-R-U5, denominada repetición terminal larga o LTR. Posteriormente, la cápside viral es importada al núcleo a través del NPC, se desensambla y el ADN viral es integrado en el genoma de la célula hospedera gracias a la proteína viral Integrasa, estableciendo el estado de provirus. Hasta hace poco se pensaba que la cápside viral se desensamblaba en el citoplasma, que luego ocurría la transcripción reversa y, posteriormente, el ADN viral ingresaba al núcleo y se integraba. Sin embargo, estudios recientes sugieren que la transcripción reversa se inicia en la cápside intacta la cual ingresa al núcleo (Dharan et al., 2020; Guedán et al., 2021; C. Li et al., 2021; Zila et al., 2021) y que el desensamblaje ocurre muy cerca del sitio de integración en el cromosoma celular (Burdick et al., 2020).

1.1.2.2 Fase tardía.

Transcripción, *splicing* y exportación. Una vez integrado, el provirus es transcrito por la ARN polimerasa II celular que se une a la región LTR, en la que hay sitios de unión de factores de transcripción (Rosen et al., 1985). Una vez iniciada la transcripción, la ARN polimerasa se detiene durante la elongación, produciendo sólo transcritos cortos (Kao et al., 1987). La proteína viral Tat interactúa con el factor positivo de la elongación de la transcripción (pTEFb), un complejo que normalmente está inactivo, pero que es activado por Tat. El complejo Tat-pTEFb además interactúa con una serie de otros factores de elongación. Tat se une a una estructura de ARN denominada elemento de Respuesta a la Transactivación (TAR) ubicada en la 5'UTR inmediatamente río abajo del sitio de inicio de la transcripción, reclutando a pTEFb hacia ese lugar. Ahí, la quinasa CDK9, que forma parte del complejo pTEFb, se activa. Por un lado, CDK9 fosforila factores inhibidores de la elongación y, por otro lado, hiperfosforila al

dominio C-terminal de la ARN polimerasa II, permitiendo la transcripción completa del genoma viral (Isel & Karn, 1999; Ivanov et al., 2000).

A partir de la transcripción del provirus se origina un pre-ARNm de 9 Kb, similar al ARN genómico viral (ARNg), que en su extremo 5' tiene una estructura Cap y en su extremo 3', una cola poliA. Este pre-ARNm tiene dos regiones intrónicas y gracias al mecanismo del *splicing* alternativo, se generan más de 100 transcritos diferentes (Ocwieja et al., 2012), cuya conformación depende del sitio y del número de eventos de *splicing*. Específicamente, el pre-ARNm tiene 4 sitios donadores y 8 sitios aceptores de *splicing* óptimos y subóptimos, dependiendo de la similitud con la secuencia consenso y, además, posee elementos en *cis* cerca de los sitios donadores y aceptores de *splicing* que regulan su actividad. Estos elementos pueden ser activadores de *splicing* exónicos (ESEs), así como también silenciadores de *splicing* exónicos (ESSs) e intrónicos (ISSs) (Karn & Stoltzfus, 2012; Martin Stoltzfus, 2009; Tazi et al., 2010). Estos transcritos se pueden agrupar en 3 clases diferentes: los que sufren *splicing* completo (2 kb) o parcial (4 kb) y transcritos completos o ARNg que no sufren *splicing* (9 kb) (Purcell & Martin, 1993). La presencia de intrones en los transcritos de 4 y 9 kb impide que estos sean exportados del núcleo a través de la vía canónica de exportación de ARNm celulares mediada por NXF1-NXT1 (TAP-p15) y deben superar los mecanismos de retención y degradación de transcritos que poseen intrones. La proteína viral Rev, sintetizada a partir de los transcritos de 2 kb, permite resolver este inconveniente. Ésta se une a una región específica denominada elemento de respuesta a Rev (RRE) presente en los transcritos de VIH-1 que poseen intrones y recluta al factor de exportación CRM1, cuya interacción es estabilizada por RanGTP. Así, gracias al complejo ARN-Rev-CRM1-RanGTP, los transcritos de 4 y 9 kb son exportados a través del complejo de poro nuclear (NPC) hacia el citoplasma donde reclutan la maquinaria celular para ser traducidos (Askjaer et al., 1998; Malim et al., 1989; Pollard & Malim, 1998; Zapp & Green, 1989).

Traducción. A partir de los transcritos de 2 kb se sintetizan las proteínas reguladoras Tat, Rev y la proteína accesoria Nef. A partir de los transcritos de 4kb se sintetizan las proteínas accesorias Vif, Vpu y Vpr, y la proteína de superficie Env. Por último, a partir del transcrito de 9 kb o ARNg se sintetizan las poliproteínas estructurales Gag y Gag-Pol (Kuzembayeva et al., 2014). VIH-1 utiliza distintas estrategias de inicio de la traducción. Se ha observado que en etapas tempranas de la infección, el ARNg de VIH-1 utiliza el mecanismo canónico de inicio de la traducción, dependiente de Cap y de eIF4E. Además de este mecanismo, existe evidencia de que VIH-1 utiliza un mecanismo de inicio de la traducción no canónico dependiente de Cap y del complejo de unión a Cap (CBC) (Sharma et al., 2012). Sin embargo, en etapas tardías de la infección, VIH-1 genera un arresto celular en la fase G2/M, suprimiendo la traducción canónica dependiente de Cap. En esta etapa, la traducción de los transcritos virales ocurre mediante mecanismos no canónicos (Monette et al., 2013; Ricci et al., 2008). Tal es el caso del inicio de la traducción mediante un mecanismo dependiente de sitios internos de entrada de ribosomas (IRES) (Barrera et al., 2020). El ARNg de VIH-1 presenta dos IRES: uno en la región 5'-UTR y otro en la región codificante de Gag (Brasey et al., 2003; Buck et al., 2001) y permiten el reclutamiento de los ribosomas, independiente de Cap.

Ensamblaje, empaquetamiento, liberación y maduración. El ensamblaje de nuevas partículas virales es un proceso orquestado por la proteína viral Gag, cuyos dominios cumplen distintas funciones, incluyendo el tráfico hacia el sitio de ensamblaje, el empaquetamiento del ARNg, la incorporación de Env y la multimerización de Gag y Gag-Pol (Freed, 2015). El sitio primario de ensamblaje corresponde a la cara interna de la membrana plasmática de la célula hospedera y es dirigido por el dominio MA de Gag (Ono, 2010). Mediante el grupo miristato ubicado en la región N-terminal, MA interactúa y se inserta en la membrana plasmática (Bryant & Ratner, 1990; Ono et al., 2000). El empaquetamiento del ARNg es un proceso que depende de estructuras presentes en la 5'UTR del ARNg y el dominio de Nucleocápside de Gag (Kutluay

& Bieniasz, 2010; Mailler et al., 2016). Por un lado, el ARNg dimeriza mediante la interacción de la estructura de dimerización (DIS) (Paillart et al., 1996; Skripkin et al., 1994). Por otro lado, el dominio NC de Gag tiene dos dominios de tipo “dedos de zinc” que permiten la interacción con la señal de empaquetamiento (Psi) del ARNg (Cimarelli et al., 2000). Así, Gag interacciona con el dímero del ARNg y empaqueta 2 copias de éste. Luego de la llegada a la membrana, el dominio CA de Gag promueve la multimerización de Gag y Gag-Pol (Mateu, 2009). Finalmente, las partículas son liberadas por yemación, mediante la fisión de la membrana gracias al complejo ESCRT que es reclutado por el dominio p6 de Gag (Sundquist & Krausslich, 2012). Las partículas liberadas son procesadas por la Proteasa viral, que procesa a Gag y Gag-Pol dando origen a partículas maduras e infectivas (Bell & Lever, 2013; Freed, 2015).

Como se puede apreciar, además de ser la proteína más abundante de las partículas virales, Gag es un regulador fundamental en las distintas etapas de la fase tardía del ciclo replicativo viral. Esta proteína se sintetiza exclusivamente a partir del ARN de 9 Kb o ARNg, el cual cumple otra función fundamental, pues es empaquetado como genoma durante la formación de nuevas partículas virales; características que lo hacen de particular interés (Kuzembayeva et al., 2014; Mailler et al., 2016).

1.2 CAPÍTULO 2: Complejos Ribonucleoproteicos (RNPs).

A lo largo de su existencia, un ARNm pasa por diversas etapas: *splicing*, exportación, localización, traducción y degradación. Durante cada una de estas etapas, el ARNm no está solo, sino que se encuentra asociado con diversos factores celulares que determinan su destino (Moore, 2005; Müller-Mcnicoll & Neugebauer, 2013; Singh et al., 2015). En otras palabras, todo lo que le ocurre a un ARNm depende de los factores que interactúen con él. Los complejos formados por un ARNm y proteínas asociadas se denominan Complejos Ribonucleoproteicos del ARNm (mRNPs). Así, las proteínas de unión a ARN (RBPs) han sido foco de gran interés,

pues median la interacción del ARNm con la maquinaria celular, brindando una capa de regulación adicional de la expresión génica.

1.2.1 Composición de los complejos Ribonucleoproteicos.

Se han identificado diversas proteínas que forman parte de los mRNPs en células eucariontes. Las primeras investigaciones detectaron proteínas que actualmente se sabe que son ubicuas en los ARNm celulares, como por ejemplo, la familia de ribonucleoproteínas heterogéneas nucleares (hnRNPs), proteínas ricas en arginina y serina (SR), proteínas de unión a poliA (PABP), proteínas de unión a Cap (CBP20/80, eIF4E) y el complejo de unión a exones (EJC) (Moore & Proudfoot, 2009; Singh et al., 2015). Con el desarrollo de técnicas como el entrecruzamiento con luz UV, captura de ARNm y espectrometría de masas, estudios recientes han identificado cientos de proteínas de unión a ARNs (Baltz et al., 2012; Bao et al., 2018; Castello et al., 2012; Urdaneta et al., 2019), llegando a sugerir que hay alrededor de 1000 proteínas capaces de unirse a los ARNm en células humanas (Hentze et al., 2018; Khong & Parker, 2020).

La composición de los mRNPs es dinámica. Se sabe que las primeras proteínas en unirse a un ARN lo hacen en el núcleo, de manera co-transcripcional, pero no lo acompañan necesariamente durante todas las etapas posteriores. Ciertos eventos celulares, como la exportación a través del NPC, la llegada al citoplasma y la primera ronda de traducción impactan de manera importante la composición de los mRNPs (Moore, 2005), pues la gran mayoría de las proteínas que estaban asociadas al ARNm se desprenden, provocando un reordenamiento de estos complejos. Aun así, hay proteínas que se unen de manera más estable. De hecho, muchos componentes de los mRNPs citoplasmáticos son adquiridos en el núcleo y permiten el acoplamiento de etapas sucesivas, como el *splicing* y la exportación

nuclear; y de otras no sucesivas, como el *splicing* y la traducción (Giorgi & Moore, 2007; Moore & Proudfoot, 2009).

1.2.2 Complejos Ribonucleoproteicos de VIH-1.

Como se vio en el capítulo anterior, el ARNg de VIH-1 es exportado por vías no canónicas hacia el citoplasma, tiene una región 5'UTR altamente estructurada y es traducido por mecanismos dependientes e independientes de Cap. Además, cumple dos funciones relevantes: i) es utilizado como mensajero para sintetizar las proteínas virales Gag y Gag-Pol y ii) es empaquetado como genoma. Al igual que los ARNm celulares, las interacciones entre el ARNg de VIH-1 con proteínas virales y celulares determinan su destino celular. Todo lo anterior, ha generado un particular interés por estudiar e identificar las proteínas que forman parte de los mRNPs del ARNg de VIH-1.

Por ejemplo, se han descrito las proteínas celulares que interactúan con Tat y promueven la transcripción (Karn, 2011). Como se vio anteriormente, Tat activa a pTEFb y lo recluta hacia el dominio TAR en la 5'UTR. Además, se han identificado una serie de factores de transcripción/coactivadores y factores de elongación que interactúan con Tat-pTEFb como AFF1, AFF4, ENL, AF9, ELL2, EAF1 y PAF1 (He et al., 2010; Sobhian et al., 2010). Este gran complejo, promueve la elongación de la transcripción del genoma viral. También se han estudiado los factores celulares que regulan el *splicing* alternativo del virus. Como se mencionó, a partir del ARNg de 9 Kb, se generan más de 100 transcritos diferentes. Esto ocurre gracias a la presencia de sitios donadores y aceptores de *splicing*, y a la presencia de elementos silenciadores (ESSs e ISSs) o activadores (ESEs). Se han descrito dos familias de proteínas que se unen a estos elementos. Por un lado, las proteínas de la familia SR generalmente se unen a ESEs y promueven el *splicing*. Entre ellas destacan las proteínas SF/ASF, SC35 y SRp40 (Mayeda et al., 1999; Ropers et al., 2004). Por otro lado, las proteínas de la familia

hnRNP generalmente se unen a ESSs e ISSs y reprimen el *splicing*. Entre ellas destacan las proteínas de la familia hnRNP A/B (hnRNP A1, A2 y B1) (Caputi et al., 1999), hnRNP H (Jacquenot et al., 2001) y hnRNP K (Marchand et al., 2011). Las proteínas hnRNPs suelen encontrarse en el núcleo, sin embargo, también son capaces de salir hacia el citoplasma y se ha observado que, además de su rol en la regulación del *splicing* alternativo, algunas de ellas forman parte de mRNPs relacionados con la exportación nuclear y traducción, como se verá más adelante.

Otro componente ampliamente estudiado de los mRNPs del ARNg de VIH-1 ha sido la proteína viral Rev y los factores celulares que interactúan con ella. Como ya se mencionó, Rev forma un complejo ARN-Rev-CRM1-RanGTP, permitiendo la exportación de los ARNs con intrones, pero también se ha observado un rol de Rev en traducción. Así, se han identificado diversos factores celulares que interactúan con Rev y que regulan distintas etapas del metabolismo del ARNg, como es el caso de las helicasas de ARN de tipo DEAD/H-box, tales como DDX1 y DDX3, entre otras (Fang et al., 2004; Fröhlich et al., 2016; Yasuda-Inoue et al., 2013; Yedavalli & Jeang, 2010); la proteína de interacción con Rev/de unión al dominio de activación Rex hRIP/RAB (Sánchez-Velazco et al., 2004); la proteína de matriz nuclear MatrIn3 (Kula et al., 2011); la helicasa MOV10 (F. Huang et al., 2015) y la proteína de degradación de ARN UPF1 (Ajamian et al., 2015), quienes participan en la estabilidad y exportación de los transcritos con intrones.

Las proteínas que participan en el inicio de la traducción también han sido un foco de estudio importante. VIH-1 utiliza distintas estrategias de inicio de la traducción, dependientes e independientes de Cap (Barrera et al., 2020; Monette et al., 2013; Ricci et al., 2008). Por un lado, se han identificado proteínas que forman parte de los mRNPs no canónicos y promueven el inicio de la traducción dependiente de Cap, entre las que se encuentran las helicasas DDX3 (Soto-Rifo et al., 2012, 2013) y DHX9/RHA (Bolinger et al., 2007, 2010), Staufen 1 (Dugré-Brisson et al., 2005), el autoantígeno La (Svitkin et al., 1994) y UPF1 (Ajamian et al., 2008). Por

otro lado, se han identificado proteínas o factores que actúan en *trans* sobre IRES (ITAFs), que son necesarios para este tipo de inicio de la traducción independiente de Cap (Barrera et al., 2020). Algunos ITAFs de VIH-1 descritos a la fecha son: DDX3 (Liu et al., 2011), eS25 (Carvajal et al., 2016), hRIP (Liu et al., 2011), HuR (Rivas-Aravena et al., 2009), GADD34 (Ishaq et al., 2020), hnRNP A1 (Monette et al., 2009), hnRNP U (Tesis Nataly Contreras 2018), hnRNP K (Olguín Araneda, 2020) y Staufen 1 (Ramos et al., 2021).

Cabe destacar que varias de las proteínas mencionadas regulan más de una etapa del metabolismo del ARNg, como DDX3 (exportación y traducción), UPF1 (exportación y degradación), hRIP (exportación y traducción) y hnRNP A1 (*splicing* alternativo y traducción), evidenciando aún más la relevancia que puede tener un factor que forma parte de los mRNPs del ARNg. Además de las mencionadas, se han identificado decenas de proteínas que interactúan con el ARNg y regulan otras etapas del ciclo replicativo de VIH-1, como la estabilidad y degradación del ARN, el ensamblaje, el empaquetamiento y la liberación de nuevas partículas. Incluso, se han identificado un grupo de proteínas que también interactúan con el ARNg de VIH-1 y son capaces de metilar, demetilar y reconocer la metilación de éste, agregando aún más complejidad a la regulación de la expresión génica viral (Kennedy et al., 2016; Lichinchi et al., 2016; Riquelme-Barrios et al., 2018; Tirumuru et al., 2016).

La gran mayoría de los estudios en que se identificaron las proteínas revisadas, se basan principalmente en técnicas de ARN interferentes, sobreexpresión y/o silenciamiento de genes, así como también, inmunoprecipitaciones de proteínas virales. A finales de 2017, se lograron identificar decenas de nuevas proteínas que forman parte de los RNPs del ARNg de VIH-1 mediante una estrategia de captura del ARNg. Específicamente, se identificaron 189 proteínas virales y celulares que interactúan con el ARNg, de las cuales 94 no habían sido previamente descritas como proteínas que interactúan con el ARNg, ni tampoco con las

proteínas virales Tat, Rev o Gag (Knoener et al., 2017). Recientemente, este mismo grupo de trabajo, logró identificar similitudes y diferencias entre el interactoma del ARNg y el interactoma de los ARN con *splicing* parcial y completo (Knoener et al., 2021)

Como se puede apreciar, se han identificado decenas de proteínas celulares que forman parte de diversos mRNPs del ARNg de VIH-1, regulando las distintas etapas por las que éste transita. El conocimiento que se tiene acerca de las proteínas que interactúan con el ARNg es bastante amplio y ha sido profundizado por diversos grupos de investigación. Sin embargo, existe un componente importante de los mRNPs que había sido ignorado hasta hace unos años: los ARN no codificantes.

1.3 CAPÍTULO 3: ARNs no codificantes.

Por mucho tiempo el concepto de “ADN basura”, es decir, secuencias de ADN que no cumplen una función aparente, era ampliamente aceptado. Actualmente, gracias a la transcriptómica, se sabe que sólo el 2% del genoma codifica proteínas, pero que aproximadamente el 80% es transcrito (Djebali et al., 2012; Harrow et al., 2012). Así fue como se identificaron los llamados ARNs no codificantes (ARNnc). Estos se pueden clasificar en dos grandes grupos de acuerdo con su función i) los ARNnc de *housekeeping*, donde se encuentran los ARN de transferencia, ARN ribosomales, ARN pequeños nucleares y ARN pequeños nucleolares; y ii) los ARNnc reguladores (Eddy, 2001). Estos últimos se han clasificado arbitrariamente de acuerdo con su tamaño en dos subgrupos: los ARNnc cortos (< 200 nt) y los ARNnc largos (> 200 nt). De los ARNnc reguladores, los más estudiados han sido los microARNs, que mediante apareamiento de bases con su ARNm blanco reprimen su traducción o inducen su degradación (Bartel, 2009). Por otro lado, los ARNnc largos han emergido como un grupo diverso y clave en la regulación de la expresión génica.

1.3.1 ARNnc largos: Clasificación y características generales

Los ARNnc largos se definen como aquellos ARNs con un tamaño mayor a 200 nt y que no tienen potencial codificante, lo que tradicionalmente se ha definido como aquellas secuencias que no presentan marcos de lectura abiertos (ORFs) mayores a 300 nt (100 aa) (Kapranov et al., 2007). Esta definición ha sido controversial, pues dentro de secuencias descritas previamente como no codificantes, se han detectado pequeños ORFs que codifican pequeñas proteínas o micropéptidos que cumplen funciones relevantes (Hartford & Lal, 2020). Sin embargo, el descubrimiento de micropéptidos funcionales no es excluyente con la función reguladora propia de los ARNnc largos.

De los 61.533 genes anotados en el genoma humano, 19.982 corresponden a genes que codifican proteínas, 7.567 a ARNnc cortos y 18.811 a ARNnc largos (GENCODE v39 13-12-2021). Estos últimos se pueden clasificar en 5 grandes grupos de acuerdo con su posición relativa en el genoma (Figura 2): i) Intergénicos. Ubicados entre dos genes codificantes, ii) Intrónicos. Ubicados en un intrón de un gen codificante, iii) Antisentido. Ubicados en la hebra complementaria de un gen codificante, iv) Bidireccionales o Promotores. Transcritos desde un promotor, en la dirección opuesta a su gen codificante y v) Enhancers. Ubicados en regiones enhancer (Fernandes et al., 2019).

Los ARNnc largos son transcritos por la ARN Polimerasa II, por lo tanto, generalmente tienen estructura Cap en 5' y cola poliA en 3'. Sin embargo, cada una de las subclases tienen características que comparten y que los diferencian. Los ARNnc bidireccionales tienen un tamaño entre 200-600 nt y presentan estructura 5'-Cap y poliA. En cambio, los ARNnc enhancers tienen un tamaño menor a 2000 nt, tienen estructura 5'-Cap, pero son procesados por el complejo Integrator por lo que carecen de cola poliA. Ambos tipos de ARNnc son retenidos en el núcleo, tienen una corta vida media y no está claro si son funcionales o no (H.

Wu et al., 2017). Aproximadamente la mitad de los ARNnc largos son intergénicos (Ransohoff et al., 2018), siendo esta la subclase más abundante y estudiada. Esto ha llevado a que algunos grupos de trabajo se refieran indistintamente a ARNnc largos y ARNnc largos intergénicos. Los ARNnc largos intergénicos se localizan en el núcleo y en el citoplasma, pero se encuentran más enriquecidos en el núcleo en comparación con los ARNm. Además, pueden sufrir o no *splicing*, son menos abundantes que los ARNm y su expresión es altamente tejido-específica (Cabili et al., 2011; Rinn & Chang, 2020). En promedio tienen un tamaño de 1 kb, pero se han descrito algunos que pueden superar las 15 kb (Brockdorff et al., 1992; J. T. Lee et al., 1999; Sunwoo et al., 2009). Los ARNnc antisentido tienen características muy similares a los ARNnc intergénicos, siendo la principal diferencia la región en el genoma donde están codificados (Wu et al., 2017).

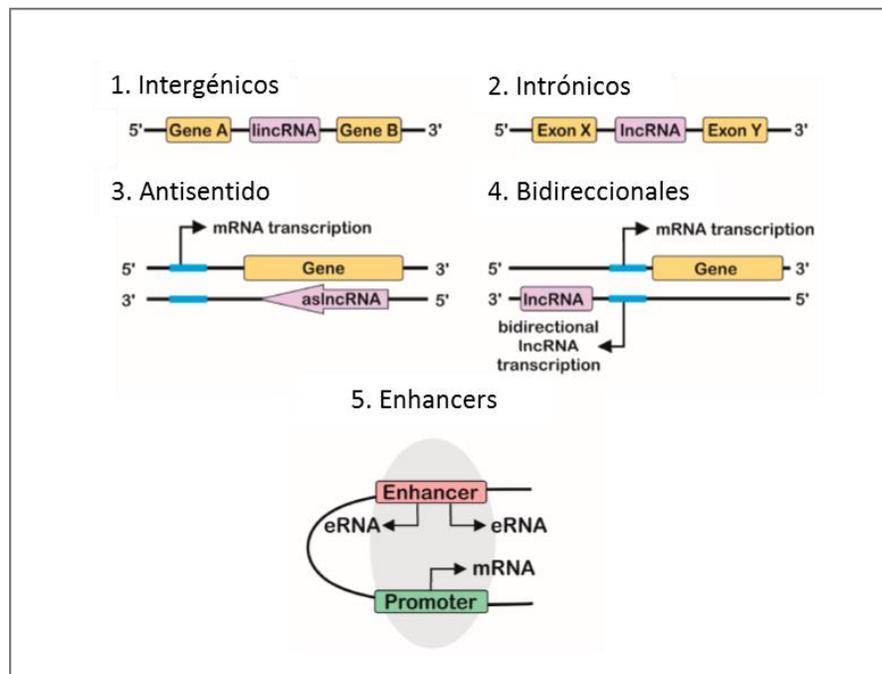


Figura 2. Clasificación de los ARNnc largos según su locus genómico. Los ARNnc largos se pueden clasificar en 5 grupos de acuerdo a su posición en el genoma: 1. Intergénicos, 2. Intrónicos, 3. Antisentido, 4. Bidireccionales y 5. Enhancers (Fernandes et al., 2019).

Debido a su gran tamaño, los ARNnc largos pueden adquirir estructuras tridimensionales y son capaces de interactuar con ARN, ADN y proteínas, mediante apareamiento de bases o interacciones estructurales. Pueden actuar secuestrando ARN o proteínas, impidiendo su unión con sus moléculas blanco o guiándolas hacia ellas y también pueden actuar como andamios promoviendo el ensamblaje de mRNPs (Guttman & Rinn, 2012; Kazimierczyk et al., 2020; Rinn & Chang, 2012). Todo esto los convierte en reguladores particularmente versátiles de la expresión génica.

1.3.1.1 ARNnc largos como reguladores transcripcionales y postranscripcionales.

A pesar de que casi un tercio de los genes anotados en el genoma humano corresponden a ARNnc largos, su funcionalidad está recién comenzando a ser develada. Se ha identificado la función de alrededor de 3700 (Ma et al., 2019) de los cuales la mayoría son biomarcadores de ciertas enfermedades, principalmente cáncer (X. Zhang et al., 2019). Sin embargo, los mecanismos mediante los cuales ejercen su efecto han sido escasamente profundizados y sólo algunas decenas han sido caracterizadas. Los primeros ARNnc largos descritos fueron identificados como reguladores transcripcionales, pues modulan el estado de metilación de la cromatina ya sea activándola o silenciándola. Un paradigmático ejemplo es Xist, un ARNnc largo de 17 kb que participa en la inactivación del cromosoma X (Brown C et al., 1991) . Durante el desarrollo embrionario, Xist se une en *cis* a lo largo del cromosoma X desde el cual es transcrito. Como consecuencia de esta unión ocurre el silenciamiento de un gran número de genes ubicados en dicho cromosoma. Se han identificado 6 dominios funcionales de Xist, los que interactúan con diversas proteínas entre las que se encuentran factores de modificación de la cromatina como el complejo de represivo PRC1 y PRC2 (Loda & Heard, 2019). Otros ejemplos de ARNnc largos que inducen el silenciamiento de genes mediante el reclutamiento de proteínas que modulan el estado de la cromatina son: H19 (Bartolomei et al.,

1991), ANRIL (Yap et al., 2010) y HOTAIR (Rinn et al., 2007), entre otros. A diferencia de Xist, HOTAIR actúa en *trans* sobre sus genes blanco silenciando, por ejemplo, genes HOX mediante el reclutamiento de PRC2 (Rinn et al., 2007). Por otro lado, existen ARNnc largos que activan genes, como es el caso de HOTTIP, que es transcrito en el extremo 5' del locus *HOXA* y actúa en *cis* sobre genes de este mismo locus, promoviendo su activación mediante el reclutamiento del complejo de metilación WDR-MLL (K. C. Wang et al., 2011).

La participación de ARNnc largos en etapas postranscripcionales era desconocida hasta hace menos de 15 años. Actualmente se sabe que se pueden localizar tanto en el núcleo como en el citoplasma e incluso en las mitocondrias (Aillaud & Schulte, 2020) y que, además de ser importantes reguladores transcripcionales, están involucrados directamente en todos los procesos postranscripcionales de la expresión génica (Figura 3) (Noh et al., 2018; Statello et al., 2021; Yoon et al., 2013). Por ejemplo, MALAT1 regula el *splicing* alternativo (Tripathi et al., 2010); GAS5 y LincRNA-p21 inhiben la traducción (G. Hu et al., 2014; Yoon et al., 2012) y Uchl1-AS la activa (Carrieri et al., 2012). TINCR y BACE-AS promueven la estabilización de ARNm (Faghihi et al., 2008; Kretz et al., 2013) y ½-sbsRNA induce su degradación (Gong & Maquat, 2011). Otra función ampliamente descrita de los ARNnc largos es que actúan como competidores endógenos o “esponjas” de microARNs, impidiendo su interacción con sus moléculas blanco. Cabe mencionar que un mismo ARNnc largo puede participar en la regulación de más de un proceso celular.

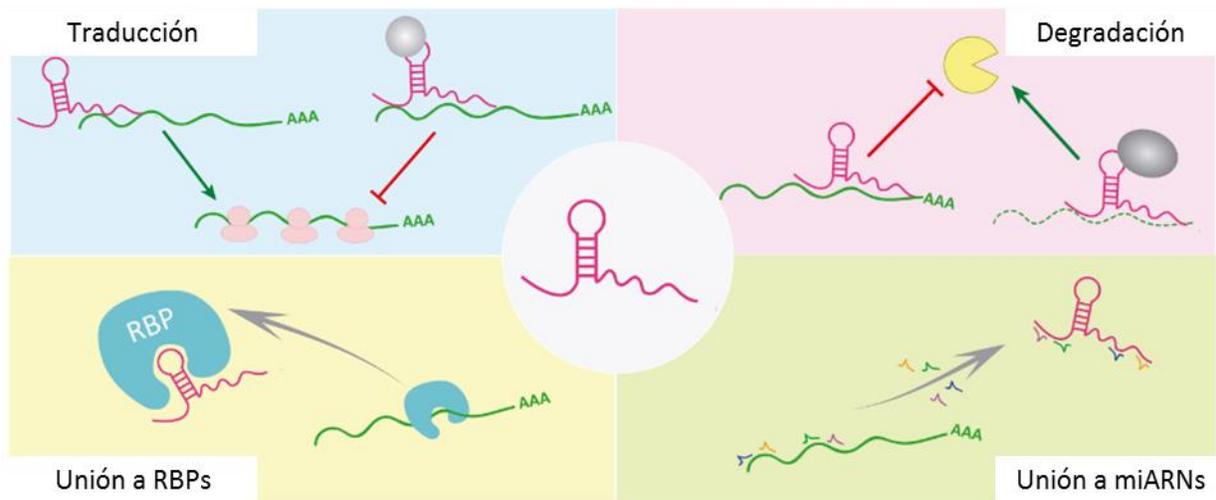


Figura 3. Regulación postranscripcional de los ARNnc largos. Los ARNnc largos pueden regular la expresión génica en distintas etapas postranscripcionales: Activando o inhibiendo la traducción, promoviendo o inhibiendo la degradación de ARNs, secuestrando proteínas o miARNs e impidiendo su interacción con sus moléculas blanco.

1.3.1.2 ARNnc largos y VIH-1.

En los últimos años, se han estudiado principalmente las interacciones entre VIH-1 y microARNs celulares y se han identificado decenas que interactúan con los transcritos virales regulando su traducción o estabilidad (Balasubramaniam et al., 2018; Barichievy et al., 2015). Los ARNnc largos que participan en el metabolismo de VIH-1 son un foco de atención emergente (Ray & Morris, 2020). Mediante el uso de microarreglos y herramientas de secuenciación masiva, se han encontrado decenas de ARNnc largos que tienen una expresión diferencial en diversos modelos de infección con VIH-1 (Postler et al., 2017; Trypsteen et al., 2016, 2019). Hasta la fecha, se han caracterizado funcionalmente diez ARNnc largos que afectan de alguna manera la replicación del virus, de los cuales ocho han sido caracterizados en los últimos dos años (Tabla 1). La mayoría ejerce una regulación a nivel transcripcional,

como es el caso de NRON (Imam et al., 2015; Li et al., 2016), uc002yug.2 (Huan et al., 2018), MALAT1 (Qu et al., 2019), HEAL (Chao et al., 2019) y el recientemente descrito AK130181 (Li et al., 2020). Otros ejercen una regulación indirecta afectando el metabolismo de las células infectadas, como es el caso de SAF, un factor anti-apoptótico (Boliar et al., 2019), Lincp21 un factor pro-apoptótico (Barichiev et al., 2018) y GAS5 que inhibe la replicación viral mediante el secuestro de miR-873 (Chen et al., 2018). El único ARNnc largo descrito hasta la fecha que regula el metabolismo de VIH-1 a nivel postranscripcional es NEAT-1, el cual ha sido previamente asociado con la formación de subdominios nucleares denominados *paraspeckles* (Clemson et al., 2009). Al eliminar el gen NEAT1 disminuye la formación de *paraspeckles* y aumenta la exportación de ARN y la síntesis de proteínas virales. Por ello, se ha propuesto que NEAT1 secuestra los ARN con intrones de VIH-1 en dichas estructuras (Zhang et al., 2013). Aun así, no se ha demostrado una interacción directa entre NEAT-1 y el ARNg de VIH-1.

Pese a que los ARNnc largos son capaces de regular la expresión génica en etapas transcripcionales y postranscripcionales y a que son moléculas versátiles capaces de interactuar con ácidos nucleicos y proteínas, a la fecha no existen estudios que muestren la presencia de ARNnc largos asociados a los mRNPs del ARNg de VIH-1 ni su rol sobre la replicación viral. En esta tesis, se propuso evaluar estas posibilidades.

Tabla 1. ARNnc largos que afectan el metabolismo de VIH-1.

Nombre	Tipo	Efecto sobre el metabolismo de VIH-1	Referencia
NEAT-1	Intergénico	Promueve la formación de paraspeckles y retiene los ARNs con intrones de VIH-1 en el núcleo, inhibiendo la replicación viral.	(Zhang et al., 2013)
NRON	Antisentido	Inhibe la transcripción viral, pues estabiliza la forma inactiva del factor de transcripción NFAT e induce la degradación de la proteína viral Tat.	(Imam et al., 2015)
uc002yug.2	Intergénico	Activa la transcripción viral, pues reduce los niveles de RUNX1b/1c y aumenta los niveles de Tat.	(Huan et al., 2018)
GAS5	Antisentido	Inhibe la replicación viral, pues GAS5 inhibe la actividad de miR-873, un microARN que promueve de la replicación de VIH-1.	(Chen et al., 2018)
Lincp21	Intergénico	El virus secuestra las proteínas HuR en el núcleo y hRNP-K en el citoplasma, lo que impide que lincp21 ejerza su función proapoptótica.	(Barichievy et al., 2018)
MALAT1	Intergénico	Activa la transcripción viral, pues impide que el complejo de silenciamiento génico PRC2 se una a la región LTR de VIH-1.	(Qu et al., 2019)
SAF	Antisentido	Macrófagos infectados con VIH-1 tienen mayor expresión de SAF, que inhibe la actividad de caspasas apoptóticas.	(Boliar et al., 2019)
HEAL	Intergénico	Activa la transcripción viral, pues al interactuar con FUS, promueve la acetilación de H3K27 y el enriquecimiento de p-TEFb en el promotor de VIH-1.	(Chao et al., 2019)
CCR5AS	Antisentido	Estabiliza al ARNm CCR5, promoviendo su expresión en la superficie celular.	(Kulkarni et al., 2019)
AK130181	Intrónico	Inhibe la transcripción viral, pues inhibe a NF-kB	(Li et al., 2020)

2 HIPÓTESIS Y OBJETIVOS

2.1 Hipótesis

“Los mRNPs del ARNg de VIH-1 contienen ARNnc largos que podrían afectar la replicación del virus”.

2.2 Objetivo general

Identificar ARNnc largos que forman parte de los mRNPs del ARNg de VIH-1 y determinar si alguno afecta la replicación del virus.

2.3 Objetivos específicos

- 1) Aislar los mRNPs del ARNg de VIH-1.
- 2) Identificar los ARNnc largos asociados al ARNg de VIH-1.
- 3) Determinar el efecto de un ARNnc largo asociado al ARNg de VIH-1 sobre la replicación viral.

3 MATERIALES Y MÉTODOS

3.1 Plasmidios

Todos los plasmidios utilizados en este trabajo están descritos en la Tabla 2.

Tabla 2. Plasmidios utilizados en este trabajo.

Plasmidio	Características	Referencia
pNL4.3ΔEnv	Vector proviral que codifica todos los genes del VIH-1, excepto <i>env</i> donde hay una inserción de 4pb que genera un corrimiento del marco de lectura que genera un codón de término prematuro, impidiendo síntesis de la proteína de la envoltura.	Toro-Ascuy et al., 2018
pCMV-VSVg	Vector que codifica la proteína de la envoltura del virus de la estomatitis vesicular bajo el control del promotor de citomegalovirus.	Addgene
pcDNA-HAR1A	Vector que codifica al ARNnc largo HAR1A bajo el control del promotor de citomegalovirus	Este trabajo
pcDNA-GlobinRen	Vector que codifica a la luciferasa de <i>Renilla</i> , pero carece de intrones por lo que tiene bajos niveles de traducción.	Ricci et al., 2009
pBSKGagPol	Vector que codifica los genes <i>gag</i> y <i>pol</i> , flanqueados por sitios promotores T3 y T7	Soto-Rifo et al., 2013

3.2 Cultivo Celular.

Las líneas celulares adherentes: HEK293T, HeLa, Huh7, U2OS y C20 fueron cultivadas en medio DMEM (Hyclone) y las células en suspensión J-Lat y Jurkat en medio RPMI (Hyclone). Los medios fueron suplementados con suero fetal bovino al 10% v/v final, solución de antibióticos y antimicóticos 1X final (Sigma-Aldrich) y L-Glutamina 2mM final (Thermo Fisher). Las células fueron mantenidas a 37°C con una concentración de 5% CO₂ en la atmósfera. Dependiendo de la escala de los experimentos, las células fueron sembradas en placas de 6

pocillos con 2 ml de medio por pocillo, en placas de 100 mm con 10 ml o en placas de 150mm con 20 ml.

3.3 Activación de células J-Lat.

Se trabajó con células J-Lat clon 10.6 (Jordan et al., 2003). Las células se separaron en 2 cultivos a una concentración de 1 millón cél/ml. A uno de ellos se les agregó forbol 12-miristato 13-acetato (PMA) (Merck) a 10 ng/ml final en DMSO (Merck). Al cultivo control se le agregó un volumen de DMSO equivalente. Las células se cosecharon a distintos tiempos post-activación, se centrifugaron a 100 x g y se lavaron 2 veces con PBS 1X, para posteriormente realizar la extracción de ARN y proteínas según corresponda.

3.4 Preparación de partículas virales.

Las partículas virales se prepararon mediante la transfección de células HEK293T. Para ello se sembraron 1,5 millones de células en una placa de 100 mm. Luego de 24 h, se realizó la transfección utilizando polietilenimina (PEI) de ~25000 Da, siguiendo el protocolo previamente estandarizado por nuestro laboratorio (Fröhlich et al., 2016; García-de-Gracia et al., 2021; Toro-Ascuy et al., 2018). Brevemente, se mezclaron los plasmidios pCMV-VSVg y pNL4.3ΔEnv en partes iguales con PEI en NaCl al 0,9%. Se utilizó una razón de 1/15 de µg de ADN/µl de PEI. La mezcla de ADN con PEI se incubó a temperatura ambiente por 20 min y luego se agregó a las células. Las partículas virales se recuperaron 72 h post-transfección a partir del sobrenadante del cultivo el cual fue filtrado (filtro de celulosa 0,22 µm; TCL) y utilizado inmediatamente o guardado a 4°C hasta un máximo de 24 h.

3.5 Infección de células.

Las células adherentes se sembraron para ser infectadas 24 h después a un 70-80% de confluencia. Las células Jurkat fueron sembradas a 1 millón de cél/ml para ser infectadas 2-4 h después. Se retiró el 50% del medio y se le agregó un volumen equivalente de los virus

filtrados. Se mantuvieron a 37°C y 5% de CO₂ hasta ser cosechadas a distintas horas post-infección. Las células adherentes fueron lavadas 2 veces con PBS 1X directo en la placa y las células en Jurkat se centrifugaron a 100 x g por 5 min y también se lavaron 2 veces con PBS 1X, para posteriormente realizar la extracción de ARN y proteínas según corresponda.

3.6 Extracción de ARN total.

Las células cosechadas previamente y lavadas con PBS 1X se centrifugaron a 5000 x rpm por 5 min a 4°C. El *pellet* celular se resuspendió en 200 µl de PBS y se agregó 1 mL de TRIzol® (Thermo Fisher). Se mezcló con vórtex por 30 s y se incubó a temperatura ambiente por 5 min. Se agregaron 200 µl de cloroformo, se mezcló con vórtex por 30 s y se centrifugó a 10000 x rpm por 5 min. Se recuperaron 500 µl de la fase acuosa y se le agregó 500 µl de isopropanol a temperatura ambiente. Se mezcló por vórtex por 30 s, se incubó a temperatura ambiente por 5 min y se centrifugó a 13000 x rpm por 20 min a 4°C. Luego, se descartó el sobrenadante y el *pellet* se lavó cuidadosamente con etanol 70% centrifugando a la misma velocidad. Se descartó el sobrenadante, el *pellet* se secó a temperatura ambiente y se resuspendió en 17 µl de agua ultrapura. El ARN obtenido fue tratado con DNasa utilizando el kit TURBO DNA-free™ (Thermo Fisher) por 30 min a 37°C siguiendo las indicaciones del fabricante. La concentración del ARN obtenido se cuantificó midiendo la absorbancia a 260 nm con el lector de placas SYNERGY HTX multi-mode (BioTek).

3.7 RT-PCR en tiempo real.

La transcripción reversa se realizó utilizando el kit High-Capacity RNA-to-cDNA™ (Thermo Fisher) con partidores al azar, siguiendo las indicaciones del fabricante. Para cada reacción se utilizaron 300 ng de ARN total. Para el PCR cuantitativo (qPCR) cada reacción se hizo en 25 µl de volumen total. Así, para cada reacción se agregaron 5 µl de ADNc (previamente diluido 1/10), 12,5 µl de Brilliant II SYBR® Green QPCR Master Mix 2X (Agilent),

0,4 µl de mezcla partidores sentido y antisentido (10 µM cada uno) y 7,1 µl de agua ultrapura. La reacción fue amplificada con el equipo AriaMx Real-Time PCR System (Agilent Technologies) y el programa consistió en una incubación inicial a 95°C por 10 min y 40 ciclos de: 95°C por 15 s, 60°C por 15 s y 72°C por 15 s. Los niveles relativos del ARN de interés se calcularon respecto a los niveles de GAPDH, utilizando el método o ecuación de Pfaffl, que incorpora la eficiencia de los partidores. Los partidores se presentan en la Tabla 3.

Tabla 3. Lista de partidores utilizados en este trabajo

Partidor	Secuencia (5'→3')
ARNg de VIH-1	Fw: GCAGTGGCGCCCGAACAGG Rv: TTTTGGCGTACTCACCAGTC
GAPDH	Fw: AGCCACATCGCTCAGACAC Rv: GCCCAATACGACCAAATCC
7SL	Fw: ATCGGGTGTCCGCACTAAG Rv: CACCCCTCCTTAGGCAACCT
tRNALys3	Fw: GTCGGTAGAGCATCAGACTT Rv: CGCCCGAACAGGGACTT
Linc00938	Fw: CCAACGACTTCTACCACCTCT Rv: CGGAGCAGTCACCTCAGC
lnc-ISL2-1	Fw: GTGCCTTCTGTTAGCCAT Rv: CAAGCCTTCCATCATCTG
HAR1A	Fw: AAGTTTAGCATCCAATGACCCA Rv: CTCAAGGCTCGCTCTGTG
lnc-TMEM105-2	Fw: TAGGAGCAGAACTGTTGT Rv: ATAATGAAGGGGCTGTGA

3.8 Extracción de proteínas totales.

Las células adherentes se lavaron con PBS y se lisaron con 80 µl de buffer de lisis frío (NaCl 100 mM; NP-40 0,5%; EDTA 1mM y Tris-HCl 10mM pH 7,5) agregado directamente en la placa de cultivo. Las células con buffer de lisis se soltaron con un rastrillo y se traspasaron a un tubo. Las células en suspensión se lavaron con PBS mediante centrifugaciones a 100 x g y se les agregó buffer de lisis en el tubo. Una vez en buffer de lisis, se mezclaron intensamente por

vórtex por 30 s y se congelaron a -20°C por al menos 24 h antes de continuar con la extracción. El lisado se descongeló en hielo, se centrifugó a 5000 x rpm por 5 min en hielo, se recuperó el sobrenadante y se cuantificaron las proteínas totales con el método de Bradford utilizando el reactivo Protein Assay Dye Reagent Concentrate (Bio-Rad) siguiendo las indicaciones del fabricante y con el lector de placas SYNERGY HTX multi-mode (BioTek).

3.9 Western blot.

Una vez cuantificadas las proteínas, a cada muestra se les agregó buffer de carga Laemeli y se calentaron a 95°C por 5 min. Se cargaron 30 µg de proteínas por pocillo en un gel de acrilamida de dos fases: una fase concentradora compuesta por acrilamida al 5% y una fase separadora compuesta por acrilamida al 10%. La electroforesis se realizó en buffer de corrida (Tris 25mM; glicina 190 mM y SDS 0,1% finales) a un voltaje constante de 80 V durante 20 min y 120 V durante 2 h o hasta que el frente de migración cayera del gel. Luego se realizó la transferencia desde el gel a una membrana de difluoruro de polivinilideno (PVDF), previamente activada con metanol, a un amperaje constante de 0,4 A durante 1h 30 min a 4°C en buffer de transferencia (Tris 25 mM pH 8,0; glicina 190 mM y metanol 20% finales). Posteriormente, la membrana se incubó en solución de bloqueo (5% de Blotting grade blocker (BioRad) en TBS-T) con agitación durante 1 hora a temperatura ambiente. A continuación, la membrana se incubó con el anticuerpo primario correspondiente diluido en solución de bloqueo al 0,5% en TBS-T durante toda noche con agitación a 4°C. Los anticuerpos primarios utilizados fueron anti-p24 diluido 1:1000 (NIH AIDS Reagents Program, Cat N° #3537), anti-GAPDH diluido 1/5000 (Santa Cruz Biotechnologies, Cat N° #sc-51905). Posteriormente, la membrana se lavó 3 veces con buffer TBS-T y se agregó el anticuerpo secundario Anti-Ratón HRP diluido 1:5000 solución de bloqueo al 0,5% en TBS-T (Jackson ImmunoResearch) y se incubaron con agitación durante 3-4 h a temperatura ambiente. Nuevamente, la membrana se lavó 3 veces con TBS-T y incubó

utilizando el sustrato Immobilon Forte Western HRP (Merck) o Pierce™ ECL Western blotting (Thermo Scientific). Para la detección de la reacción quimioluminiscente se utilizó el equipo HD9 Western Blot Imaging System (Uvitec Cambridge).

3.10 Preparación de la sonda biotinilada.

Transcripción in vitro: Se preparó la sonda antisentido contra el ARNg de VIH-1 utilizando el vector pBSKGag-Pol (Figura 4A). Para comenzar, se digirió el vector con la enzima *EcoRI* durante 4 h. El vector digerido se purificó agregando 2,5 volúmenes de etanol 100% frío e incubando durante toda la noche a -20°C. Se centrifugó a 13000 rpm durante 30 min a 4°C, se lavó con etanol al 70% y se resuspendió en agua ultrapura. Se visualizó el vector digerido mediante electroforesis en gel de agarosa y se cuantificó midiendo la absorbancia a 260 nm. Luego, se realizó la reacción de transcripción *in vitro* utilizando la ARN Polimerasa T3 (Promega) siguiendo las instrucciones del fabricante. La reacción se hizo en un volumen total de 25 µl de acuerdo con lo mostrado en la Tabla 4.

Tabla 4. Reacción para la transcripción in vitro

Volumen (µl)	Reactivo
0,75	DTT (100 mM)
4	Mezcla de rNTPs (*)
5	Buffer de Transcripción (5X)
0,5	RNasa in (40 U/µl)
1	ADN molde(1 µg/µl)
0,46	ARN pol T3 (18.4 U/µl)
13,29	Agua ultrapura

La mezcla de rNTPs (*) con 50% de UTP-16-biotina (Epicentre) consistió en: rATP, rCTP y rGTP 10mM final, UTP 5mM final y UTP-16-biotina 5mM final. La reacción se incubó a 37°C

durante 5 h. Luego, se hizo un tratamiento con DNasa utilizando el kit TURBO DNA-free™ (Thermo Fisher) por 30 min a 37°C siguiendo las indicaciones del fabricante.

Purificación: El ARN transcrito se purificó utilizando las perlas magnéticas Agencourt® RNAClean® XP (Beckman Coulter) siguiendo las instrucciones del fabricante. La reacción de 25 µl se llevó a un volumen final de 50 µl con agua ultrapura. Se agregaron 1,8 volúmenes (90 µl) de perlas magnéticas. Se homogeneizaron mediante pipeteo y se incubaron por 5 min. Los tubos se insertaron en una gradilla magnética, se incubaron durante 5 min y se descartó el sobrenadante. Se lavaron 2 veces con 180 µl de etanol 70% sin perturbar las perlas. Se descartó el etanol, se centrifugó a 1000 x g durante 1 min, se insertaron los tubos en la gradilla magnética, se incubaron 30 s y se retiró todo el etanol excedente. Se incubaron los tubos destapados durante 2-5 min cuidando que las perlas nunca se secaran por completo. Finalmente, se resuspendieron en 12 µl de agua ultrapura, se incubaron durante 3 min, se insertaron en la gradilla magnética, se incubaron 3 min y se recuperó el eluido en un tubo nuevo. Todo fue realizado a temperatura ambiente. El ARN recuperado se cuantificó midiendo la absorbancia a 260 nm con el lector de placas SYNERGY HTX multi-mode (BioTek).

Fragmentación y purificación desde gel: La fragmentación se llevó a cabo con el kit Ambion™ fragmentation reagents (Thermo Fisher) siguiendo las instrucciones del fabricante. Brevemente, se realizó la reacción de fragmentación con 5 µg de ARN en 10 µl de volumen total. Se incubó a 70°C por 15 min. Inmediatamente se agregó el reactivo para detener la reacción y se mantuvo en hielo. Para recuperar la sonda fragmentada se hizo una electroforesis en gel de Acrilamida-Urea, el cual se precalentó a 17 watts constantes durante 1 hora, luego se cargaron las muestras y se corrieron a la misma corriente durante 1 hora en buffer TBE. El gel se tiñó con SYBR™Gold (Thermo Fisher) en buffer TBE durante 10 min protegido de la luz, luego se visualizó en un transiluminador con luz UV y se cortó la sonda del tamaño deseado. El gel se trituró en buffer de elución de gel (NaCl 30 mM y EDTA 10 mM) , se incubó en agitación

durante toda la noche y se recuperó el sobrenadante. Se cuantificó midiendo la absorbancia a 260 nm, se alicuotó y se mantuvo a -80°C hasta ser utilizada.

3.11 Purificación de ARN Antisentido (RAP).

Se trabajó con células J-Lat activadas o células Jurkat infectadas y su respectivo control. Para los ensayos piloto se usaron 5 millones de células y para los ensayos definitivos se usaron 15-30 millones de células. Todos los buffers utilizados fueron frescos, preparados el día anterior. El ensayo descrito corresponde a 30 millones de células, pero se puede escalar según corresponda.

Las células se centrifugaron a 100 x *g* por 5 min y se lavaron con PBS 1X. Luego se resuspendieron en 10 ml de PBS 1X y se les agregó 27 µl de formaldehído 37% (0,1% final) y se incubaron por 15 min a temperatura ambiente. Se agregó 1,25 ml solución de *quenching* (Glicina 2 M; Tris-HCl 25 mM pH 7,0), se centrifugaron a 100 x *g* por 5 min, se eliminó el sobrenadante y las células se resuspendieron con buffer hipotónico frío (Tris-HCl 20 mM pH 7,5; NaCl 15 mM; EDTA 0,5 mM; Tritón X-100 0,1% e inhibidor de proteasas) y se incubaron en hielo por 10 min. Posteriormente, se sonicaron a 40% de amplitud con el equipo Microson™ Ultrasonic cell disruptor XL por un total de 30 segundos en frío con un régimen de 2 s encendido y 10 s apagado. El lisado celular se alicuotó en tubos para microcentrífuga y se centrifugaron a 15000 x *g* por 10 min para remover el material insoluble. Se recuperó el sobrenadante en un tubo cónico y se le agregó un volumen equivalente de buffer de hibridación 2X (Tris-HCl 20 mM pH 7,5; EDTA 7 mM; EGTA 3 mM; LiCl 150 mM; NP-40 1%; N-lauroilsarcosina 0,2%; deoxicolato de sodio 0,1%; tiocianato de guanidinio 3 M y TCEP 2,5 mM finales). Se agregaron 180 µl de perlas magnéticas acopladas a estreptavidina (Promega) previamente lavadas 2 veces con buffer de hibridación 1X. Se incubaron a 55°C por 30 min a 1400 rpm. Se insertaron los tubos en una gradilla magnética y se transfirió sobrenadante a un tubo nuevo. Se guardó una alícuota de la muestra como input. Luego, se agregaron 3 µg de la

sonda biotinilada (1 µg de sonda por cada 10 millones de células) y se incubó la reacción de hibridación durante 2 h a 55°C a 1400 rpm. Para la captura se agregaron 180 µl de perlas acopladas a estreptavidina Dynabeads™ MyOne™ Streptavidin C1 (Thermo Fisher) previamente lavadas 2 veces con buffer de hibridación 1X y se incubaron durante 30min a 55°C a 1400 rpm. Los tubos se insertaron en una gradilla magnética y se descartó el sobrenadante. Las perlas con los complejos asociados, se lavaron 2 veces con buffer de lavado (20 mM Tris-HCl pH 7,5; 10 mM EDTA; 1% NP-40; 0,2% N-lauroilsarcosina; 0,1% de deoxicolato de sodio y 3M de tiocianato de guanidinio) y los complejos se eluyeron por 5 min a 95°C con buffer de elución (20 mM de Tris-HCl pH 7,5; 10 mM de EDTA; 2% de N-lauroilsarcosina y 2,5 mM de TCEP). Se recuperó lo eluido y se trató con Proteinasa K. Finalmente, se hizo una extracción de ARN con TRIzol® y se analizó el enriquecimiento de ARNg mediante RT-qPCR.

3.12 Sobreexpresión de HAR1A.

En primer lugar, la secuencia de HAR1A (ENSG00000225978) fue sintetizada y clonada en el sitio *NheI/XbaI* del vector pcDNA3.1 (pcDNA-HAR1A) mediante el servicio otorgado por GeneScript. Se sobreexpresó HAR1A en 3 líneas celulares: HeLa, Huh7 y C20. Cada línea celular se sembró por duplicado en placas de 6 pocillos. Se sembraron 250.000 células HeLa, 200.000 Huh7 y 200.000 C20. Luego de 24 h se transfectaron las distintas líneas celulares con 1 µg y 2 µg de pcDNA-HAR1A y se utilizó como control de transfección el vector pcDNA-GlobinRen. Así, en todas las condiciones se transfectaron siempre 2 µg de ADN total: el control se transfectó con 2 µg de pcDNA-GlobinRen; la condición 1) se transfectó con 1 µg de pcDNA-GlobinRen y 1 µg de pcDNA-HAR1A y la condición 2) se transfectó con 2 µg de pcDNA-HAR1A. Se hizo la mezcla de transfección para cada pocillo compuesta por 2 µg de ADN, 10 µl de PEI y 188 µl de NaCl al 0,9%, se incubó a temperatura ambiente por 20 min y se agregó a las células.

A las 24 hpt, se lavaron las células con PBS 1X, se agregó 1 ml de medio y 1 ml de partículas virales preparadas como se describe en la sección 3.3. A las 20 hpi se lavaron 2 veces con PBS 1X y cosecharon las células. Posteriormente, se realizó una extracción de ARN total y de proteínas siguiendo lo indicado en las secciones 3.5 y 3.7.

4 RESULTADOS

4.1 OE1. Aislar los mRNPs del ARNg de VIH-1.

4.1.1 Preparación de sonda antisentido biotinilada.

Con el objetivo de capturar el ARNg de VIH-1 con los ARNs que formarían parte de los mRNPs, se desarrolló un protocolo de preparación de sonda de ARN antisentido de VIH-1 mediante transcripción *in vitro* con el 50% de los rUTPs biotinilados. Para ello, se usó como molde el vector pBSKGagPol, donde los genes Gag-Pol se encuentran flaqueados por secuencias promotoras T7 (para transcripción sentido) y T3 (para transcripción antisentido). El vector se cortó con la enzima *EcoRI*, se purificó mediante precipitación con etanol y se realizó la transcripción *in vitro* con la polimerasa T3 (Figura 4). El producto de la transcripción se trató con TURBO™ DNasa, se purificó con perlas Agencourt RNA Clean XP y se fragmentó con un kit de fragmentación de RNA. Por un lado, se evaluaron distintos tiempos de fragmentación y se encontró que con 15 min de incubación se obtienen fragmentos de 300 nt o menor tamaño, por lo que se continuó preparando la sonda con ese tiempo de incubación (Figura 4A). Para recuperar la sonda y optimizar su especificidad, se probaron distintos métodos de precipitación y purificación con el objetivo de eliminar nucleótidos libres y fragmentos muy pequeños (Figura 4B). Sin embargo, con todos los métodos de precipitación, además de los fragmentos del tamaño deseado, se recuperaron fragmentos de tamaños muy pequeños. Al intentar purificar la sonda con perlas Agencourt RNA Clean XP, la eficiencia de recuperación fue muy baja (menor al 10%). Por lo tanto, se decidió purificar la sonda directamente desde un gel de Acrilamida-Urea. Para ello, luego de correr las muestras fragmentadas en el gel, se recuperó la sonda de un tamaño entre 300 y 50 nt aproximadamente, cortando directamente desde el gel.

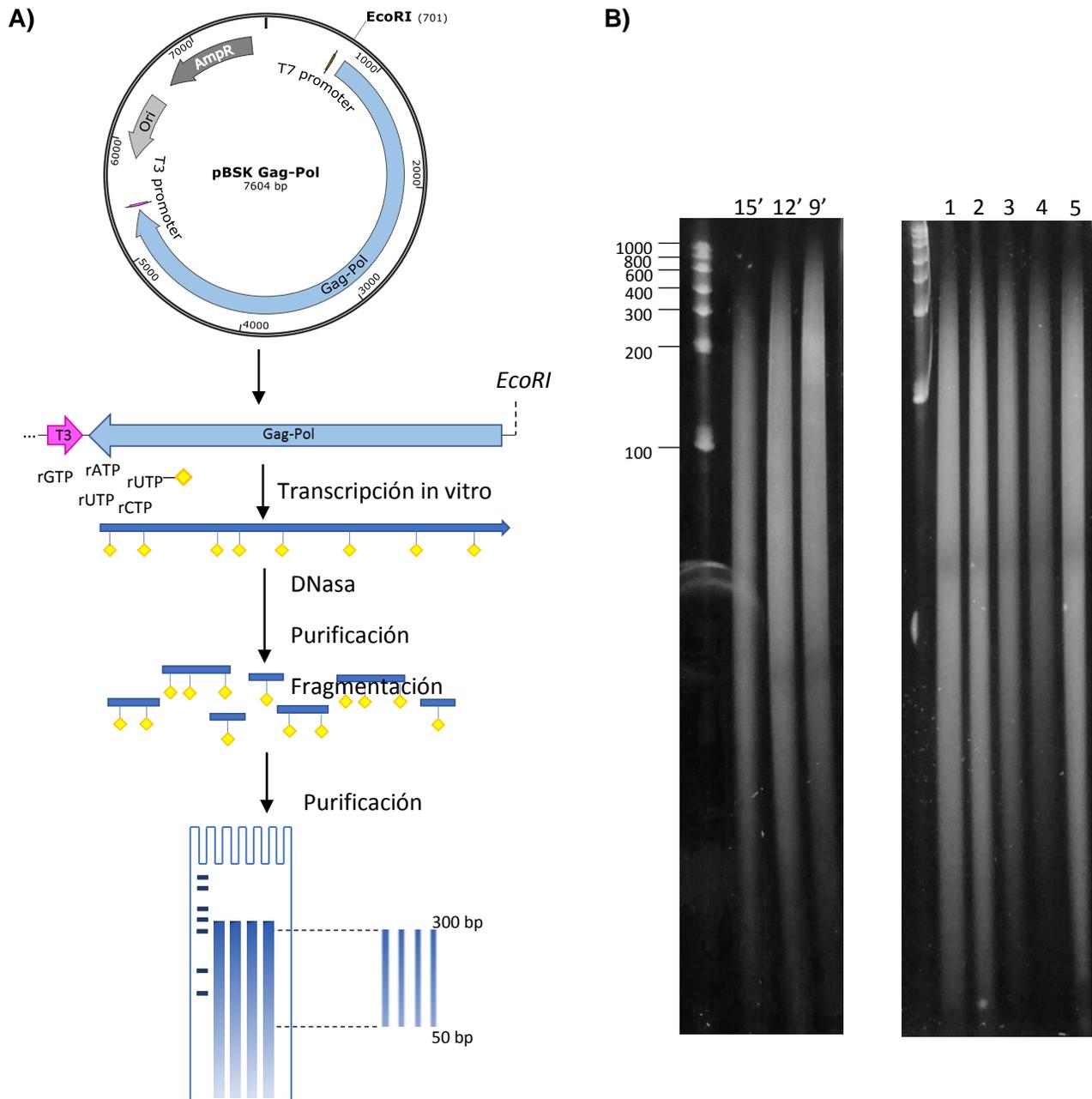


Figura 4. Preparación de la sonda de ARN antisentido de VIH-1. A) La sonda se preparó mediante transcripción in vitro usando como molde el vector pBSKGagPol digerido con *EcoRI*. El 50% de los rUTPs incorporados corresponden a 16-UTP-biotina. El producto de la transcripción fue tratado con DNasa, purificado con perlas Agencourt y fragmentado. Los fragmentos fueron cargados en un gel de Acrilamida-Urea desde el cual se cortó y recuperó la sonda con un tamaño entre 300 y 50 bp aproximadamente. B) El tamaño de la sonda se visualizó mediante Urea-PAGE. Izquierda: se evaluaron 3 tiempos de fragmentación: 15, 12 y 9 min. Derecha: se evaluaron distintos métodos de recuperación de la sonda: 1. Acetato de amonio + glicógeno; 2. Acetato de amonio; 3. LiCl; 4. Perlas Agencourt. 5. Sin purificar.

4.1.2 Captura del ARNg de VIH-1 a partir de células J-Lat.

En una primera etapa, se trabajó con células J-Lat clon 10.6, una línea celular derivada de linfocitos T (Jurkat) utilizadas como modelo de latencia de VIH-1 pues tienen integrado el genoma completo de VIH-1, pero con un corrimiento en el marco de lectura del gen *env* y con el gen *gfp* en el sitio del gen *nef*. La transcripción viral en estas células puede ser activada con compuestos como el forbol 12-miristato 13-acetato (PMA). Con el objetivo de determinar un tiempo óptimo de activación de dichas células, se realizó una cinética de activación. Se cosecharon las células entre 0 y 24 h post-activación (hpa) en intervalos de 4 h. De las células colectadas se hizo una extracción de ARN para analizar la expresión de ARNg mediante RT-qPCR y una extracción de proteínas totales para analizar mediante Western blot la expresión de la proteína viral Gag (Figura 5).

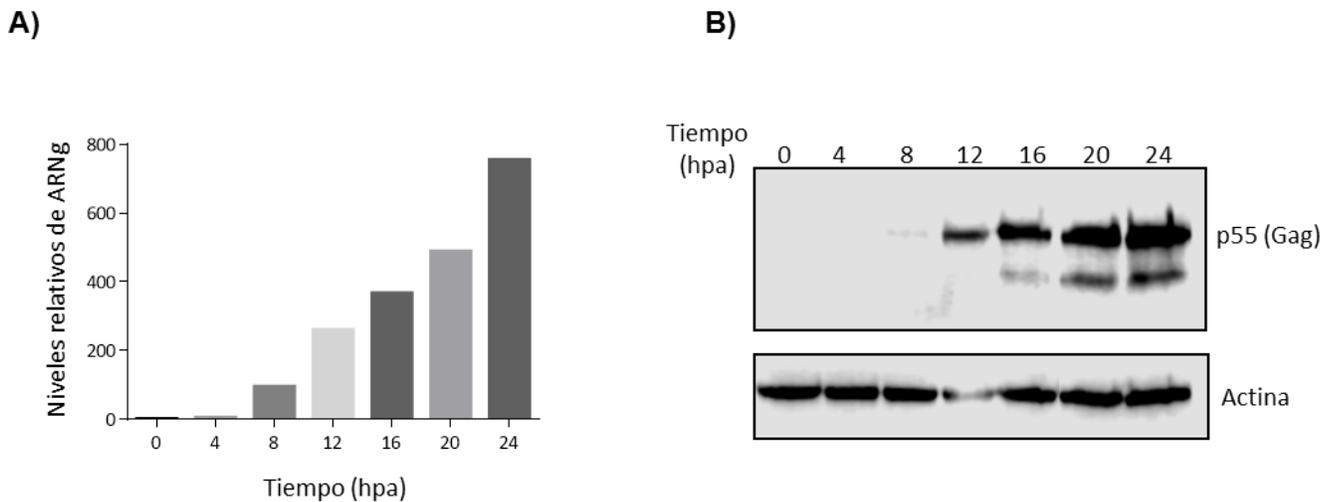


Figura 5. Cinética de activación de células J-Lat. Se hizo una cinética de activación de células J-Lat con PMA a 10 ng/ml. Se cosecharon las células entre 0 y 24 h post-activación en intervalos de 4 h. **A)** Niveles relativos de ARNg, medidos por RT-qPCR. Los valores de expresión relativa del ARNg fueron normalizados por los niveles de GAPDH **B)** Nivel de expresión de la proteína viral Gag analizados mediante Western blot usando anticuerpo monoclonal anti-p24. Se usó anti-actina como control de carga.

A partir de estos resultados se decidió que la captura de ARNg se realizaría 14 hpa, pues en ese tiempo se observan altos niveles de ARNg, pero no tan altos niveles de Gag. En otras palabras, se espera que la población de ARNg capturado a las 14 hpa esté en fases previas o iniciales de la traducción que justamente son las fases de regulación de interés en esta tesis. Una vez definido el tiempo post-activación en el que se realizaría la captura de los mRNPs de VIH-1, se hizo un ensayo piloto de captura con 5 millones de células J-Lat a las 14 hpa y 24 hpa y un control sin activar, utilizando la sonda antisentido previamente preparada. Se ajustó y estandarizó el protocolo para poder capturar los mRNPs a partir de lo publicado por McHugh et al., 2015, como se detalla en la sección 3.11 y se comenta brevemente a continuación (Figura 6). Las células activadas y sin activar se trataron con formaldehído 0,1%, se lisaron con buffer hipotónico, se sonicaron y centrifugaron. Se recuperó el sobrenadante, se le agregó buffer de hibridación y se incubó con perlas acopladas a estreptavidina. Posteriormente se retiraron las perlas, se agregó la sonda antisentido biotilada y se incubó durante 2 h. Los mRNPs del ARNg se capturaron con perlas magnéticas acopladas a estreptavidina. Las perlas con los mRNPs asociados, se lavaron con buffer de lavado y los mRNPs se eluyeron con buffer de elución. Se recuperó lo eluido y se trató con proteinasa K. Finalmente, se hizo una extracción de RNA y se analizó mediante RT-qPCR. En este primer ensayo se logró enriquecer 200 y 400 veces el ARNg capturado en células activadas 14 y 24 hpa respectivamente, relativo a GAPDH como ARN control (Figura 7). Este resultado muestra que esta estrategia de captura y que el protocolo de preparación y recuperación de la sonda antisentido permiten aislar específicamente el ARNg de VIH-1.

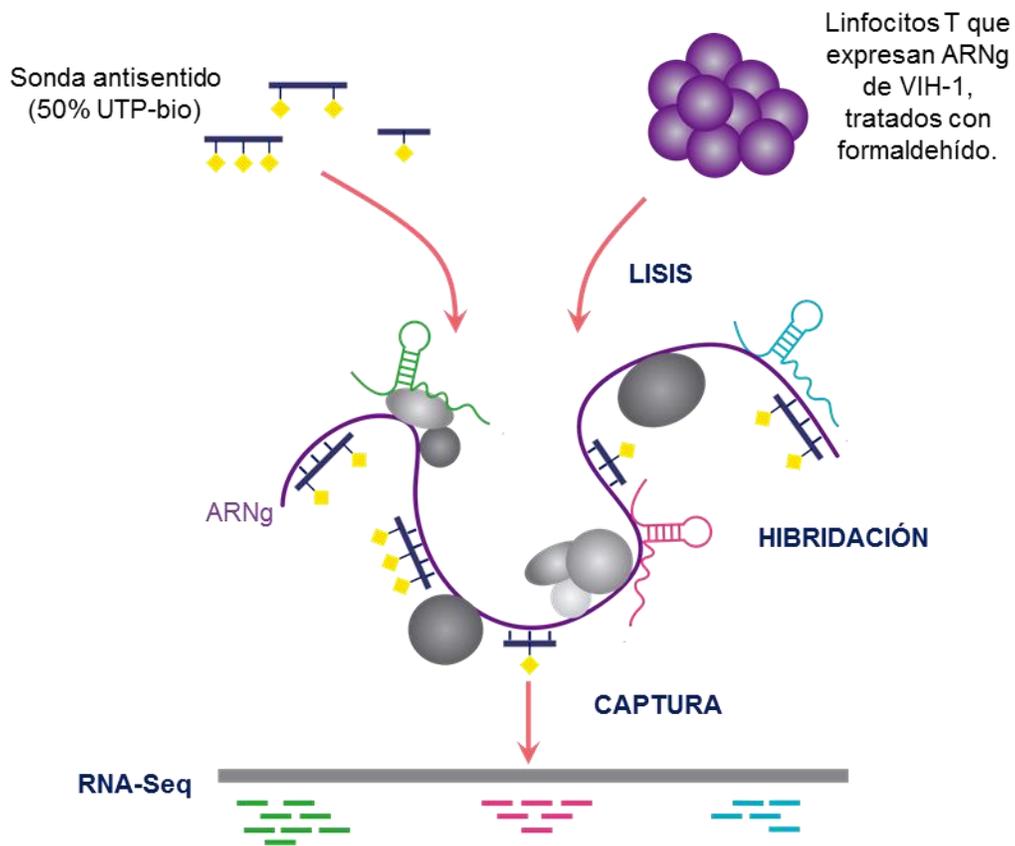


Figura 6. Purificación de ARN antisentido (RAP). Linfocitos T que expresan el ARNg de VIH-1 fueron tratados con formaldehído y lisados con buffer hipotónico y sonicación. Se agregó la sonda biotinilada al lisado celular y se incubó 2h a 55°C (1400 rpm). La sonda unida a los mRNPs del ARNg fueron capturados con perlas magnéticas acopladas a estreptavidina. Los complejos fueron eluidos a 95°C y tratados con proteinasa K. Finalmente, se hizo una extracción con fenol:cloroformo y los ARN recuperados fueron analizados mediante RNA-Seq.

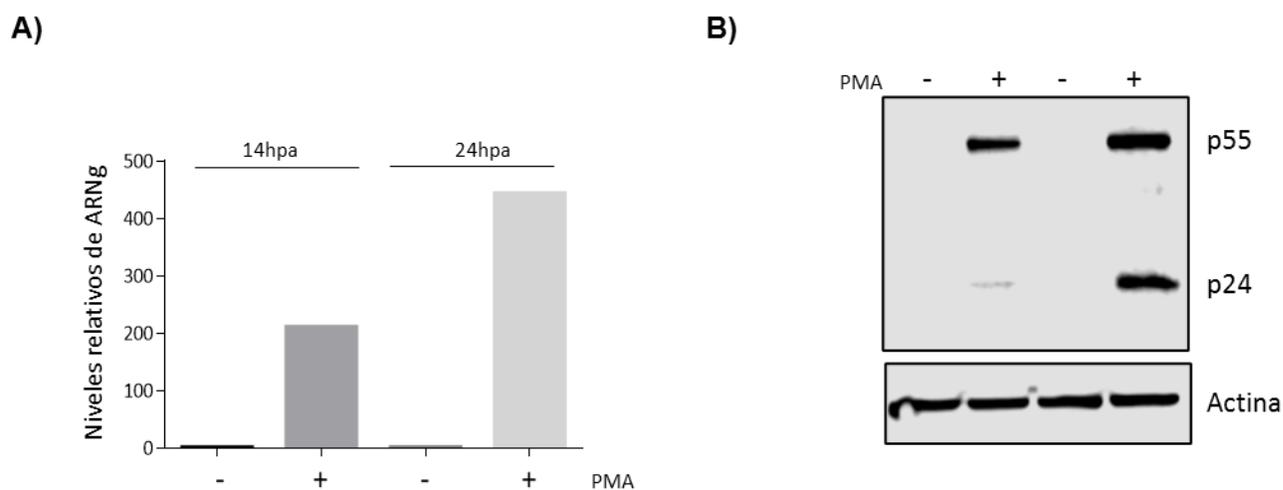


Figura 7. Ensayo piloto de captura del ARNg de VIH-1. Se capturó el ARNg a partir de 5 millones células J-Lat 14hpa, 24 hpa y control sin activar, utilizando la sonda antisentido preparada previamente. **A)** Niveles relativos de ARNg, medidos por RT-qPCR. Los valores de expresión del ARNg fueron normalizados por los niveles de GAPDH. **B)** Nivel de expresión de la proteína viral Gag analizados mediante Western blot usando anticuerpo monoclonal anti-p24. Se usó anti-actina como control.

Una vez definidos los parámetros para la purificación de los mRNPs del ARNg, se realizó la captura del ARNg a partir de 15 millones de células J-Lat a las 14 hpa y un control sin activar (Figura 8A). Al analizar los niveles de ARNg mediante RT-qPCR se logró enriquecer aproximadamente 6000 veces el ARNg de VIH-1, relativo a GAPDH como control. Lo anterior muestra que la estrategia permite capturar específicamente el ARNg de VIH-1, sin embargo, no proporciona información acerca de si es posible capturar ARNs que formen parte de los mRNPs de éste. Para ello, se quiso evaluar la presencia de dos ARNnc: tRNA^{Lys3} y 7SL. El tRNA^{Lys3} fue seleccionado como control positivo porque normalmente interactúa con el ARNg de VIH-1 dado que es utilizado como partidador para la transcripción reversa (Kleiman et al., 2010) y 7SL fue elegido como control negativo, pues es un ARNnc largo empaquetado por el virus, pero que no interactúa directamente con el ARNg (Wang et al., 2007). Al comparar los niveles de ambos ARNs, se observó que como era de esperar sólo el tRNA^{Lys3} y no el RNA 7SL está enriquecido

relativo al mock (Figura 8B), lo que indica que el protocolo de captura permite purificar el ARNg de VIH-1 y al menos un ARN asociado a él. Por lo tanto, esta estrategia permitiría capturar otros ARNs que forman parte de los mRNPs del ARNg de VIH-1.

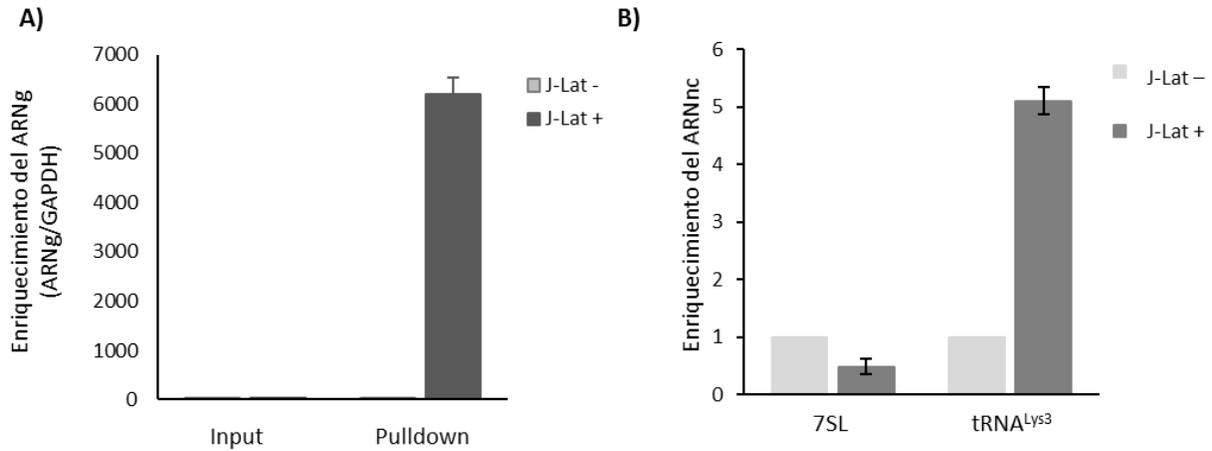


Figura 8. Enriquecimiento de ARNg y ARNnc. A) Se capturó el ARNg a partir de 15 millones células J-Lat 14 hpa y control sin activar y se evaluaron los niveles relativos de ARNg en input y pulldown por RT-qPCR. Estos valores fueron normalizados por los niveles de GAPDH. **B)** Enriquecimiento de tRNA^{Lys3} y 7SL medido por RT-qPCR en muestras provenientes de células J-Lat sin activar y activadas.

Las muestras de ARN capturado se secuenciaron mediante la plataforma Illumina HiSeq, servicio prestado por Genoma Mayor, quienes además realizaron la eliminación de los ARN ribosomales de las muestras input con el Kit RiboMinus Eukaryote System v2 (Thermo Fisher) y preparación de las genotecas con el kit TruSeq Stranded Total RNA LT (Illumina). Se encontró que una de las réplicas biológicas no cumplía con la calidad necesaria para la preparación de genotecas, por lo que se repitió el ensayo de captura. Sin embargo, nuevamente una de las réplicas no pasó la prueba de calidad, incluyendo el control, por lo que únicamente se generaron genotecas y se secuenciaron muestras correspondientes al input de ARN (ARN total inicial) y el pulldown de ARN (ARN capturado) de las células activadas. A partir de esta única

réplica biológica, se analizó la presencia de ARNnc largos asociados al ARNg de VIH-1. Dicho análisis se describe en la sección 4.2.1.

4.1.3 Captura del ARNg de VIH-1 a partir de células Jurkat infectadas.

En una segunda etapa, a diferencia del modelo de latencia anterior, se trabajó con un modelo de infección. Para ello, se infectaron células Jurkat (línea celular de linfocitos T) con virus pseudotipados. Las partículas virales se prepararon transfectando células HEK 293T con el vector pCMV-VSVg, que codifica la glicoproteína G del virus de la estomatitis vesicular, y con pNL4.3- Δ Env, que codifica todas las proteínas de VIH-1 excepto las de la envoltura. Se realizó una cinética de infección para definir cuántas horas post-infección (hpi) se realizaría la captura (Figuras 9A y 9B). Siguiendo la misma lógica que en el modelo de latencia, se encontró que a las 20 hpi hay altos niveles de ARNg y no tan altos niveles de Gag. Se realizó un ensayo piloto de captura del ARNg con 5 millones de células 20 hpi y se observó un enriquecimiento del ARNg de aproximadamente 6000 veces (Figura 9C), indicando que en este modelo de infección también es posible capturar y enriquecer el ARNg de VIH-1 con la estrategia desarrollada previamente. Así, se realizó la captura del ARNg a partir de 30 millones de células infectadas, donde nuevamente se observó un enriquecimiento en los niveles de ARNg capturado de aproximadamente 6000 veces. Se enviaron a secuenciar muestras de dos réplicas biológicas, incluyendo tanto los ARN de los inputs como de los pulldown de células infectadas y sin infectar. Todas las muestras pasaron los controles de calidad y se pudieron generar las genotecas y secuenciar.

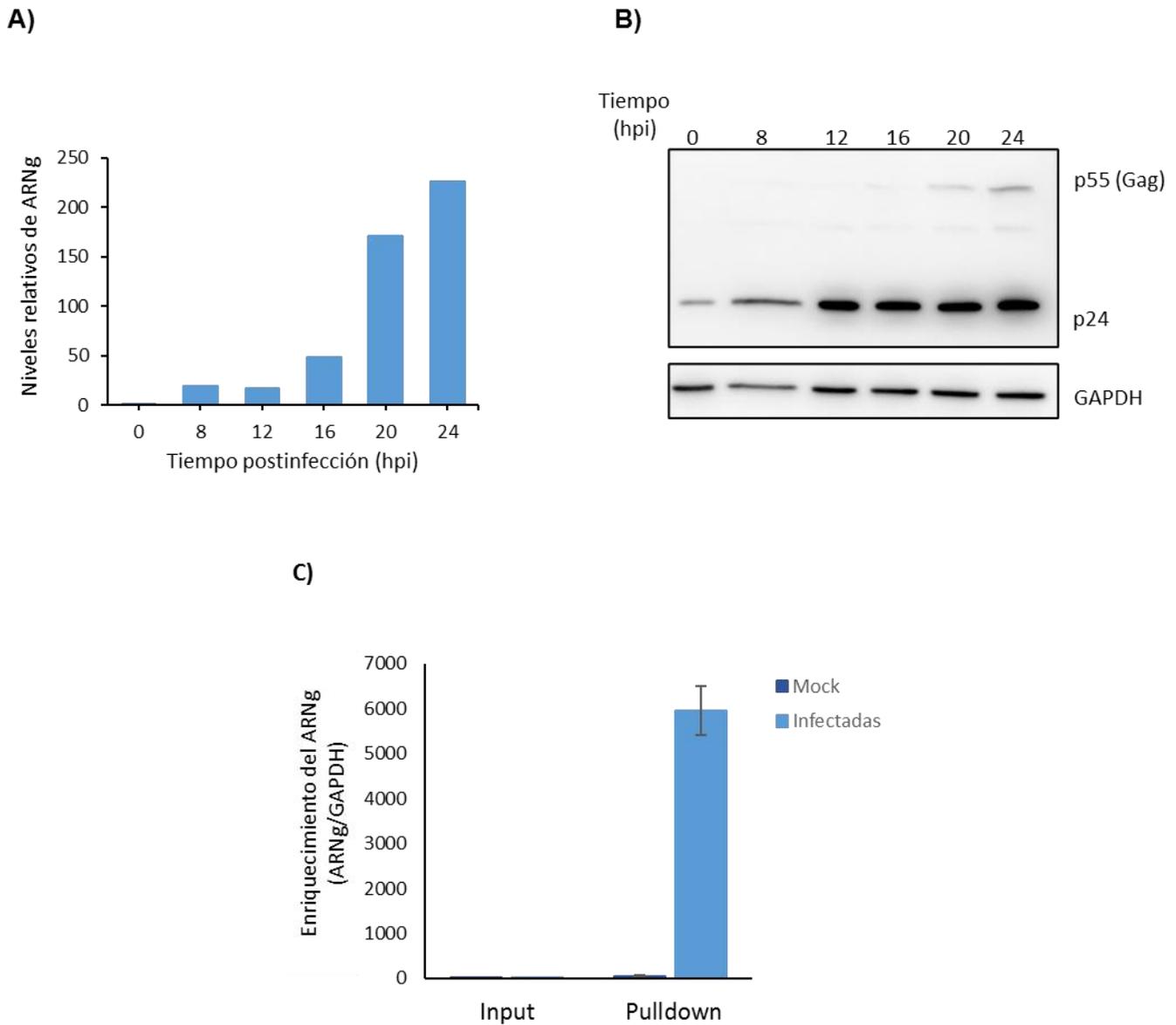


Figura 9. Cinética de infección de células Jurkat. Se hizo una cinética de infección de células Jurkat con virus preparados con los vectores pNL4.3 Δ Env y VSVg. Se cosecharon las células entre 0 y 24 h post-infección. **A)** Niveles relativos de ARNg, medidos por RT-qPCR. Los valores de expresión del ARNg fueron normalizados por los niveles de GAPDH **B)** Nivel de expresión de la proteína viral Gag analizados mediante Western blot usando anticuerpo monoclonal anti-p24. Se usó anti-GAPDH como control. **C)** Se capturó el ARNg a partir de 30 millones células Jurkat 20 hpi y mock y se evaluaron los niveles relativos de ARNg en input y pulldown por RT-qPCR. Estos valores fueron normalizados por los niveles de GAPDH.

4.2 OE2. Identificar ARNnc largos asociados al ARNg de VIH-1.

4.2.1 Identificación de ARNs capturados a partir de células J-Lat activadas.

Para identificar los ARNs presentes y posteriormente seleccionar lo más enriquecidos en la captura, fue necesario realizar un análisis bioinformático de los datos obtenidos por RNA-Seq. En el primer ensayo de captura de ARNg en células J-Lat activadas, las secuencias fueron analizadas con un flujo de trabajo bioinformático donde después del control de calidad y de la limpieza o *trimming* se calculó la abundancia de los transcritos por millón de kilobases o TPM. Tal como se mencionó, en este primer ensayo sólo fue posible secuenciar una única réplica biológica de células J-Lat activadas. Por lo tanto, una vez obtenidos los TPM de las muestras, se definió el enriquecimiento directamente como $\text{TPM pulldown}/\text{TPM input} > 2$. Los transcritos identificados se clasificaron de acuerdo con los biotipos del sistema de anotación Ensembl (Figura 10). En total se detectaron 33149 ARNs enriquecidos, donde la mayoría corresponde a ARNs codificantes (39%), seguidos por transcritos que retienen intrones (20%) y transcritos procesados sin ORFs (15%). El 13% de los transcritos enriquecidos corresponden a ARNnc largos, de los cuales el 6% son ARNs antisentido y otro 6% son ARNnc largos intergénicos (Figura 10). Debido a que estos últimos son los ARNnc largos más estudiados, de los cuáles hay más información y su validación es más confiable pues se encuentran en regiones intergénicas, se decidió continuar el estudio exclusivamente con ellos. Así, del 6% de ARNnc largos intergénicos se seleccionaron aquellos más enriquecidos (Tabla 5). De ellos sólo dos tenían funciones descritas, relacionadas con el desarrollo de cáncer. Para validar la presencia de estos ARNs seleccionados, se diseñaron partidores para cada uno de ellos y se evaluó su presencia en las muestras de ARN capturados. Sin embargo, no se detectó ninguno de los ARNs seleccionados mediante RT-qPCR. En este ensayo no fue posible eliminar el posible ruido o inespecificidad de la técnica pues no se secuenciaron las muestras de células sin activar, lo que podría en parte explicar la discrepancia entre lo observado por RNA-Seq y por

RT-qPCR. Por otro lado, se discutió y concluyó que las células J-Lat no eran el mejor modelo para realizar esta estrategia de captura, pues las células sin activar tienen un nivel de expresión basal de ARNg relativamente alto (detectado por RNA-seq, pero no así por RT-qPCR) y por lo tanto no se pueden utilizar como control.

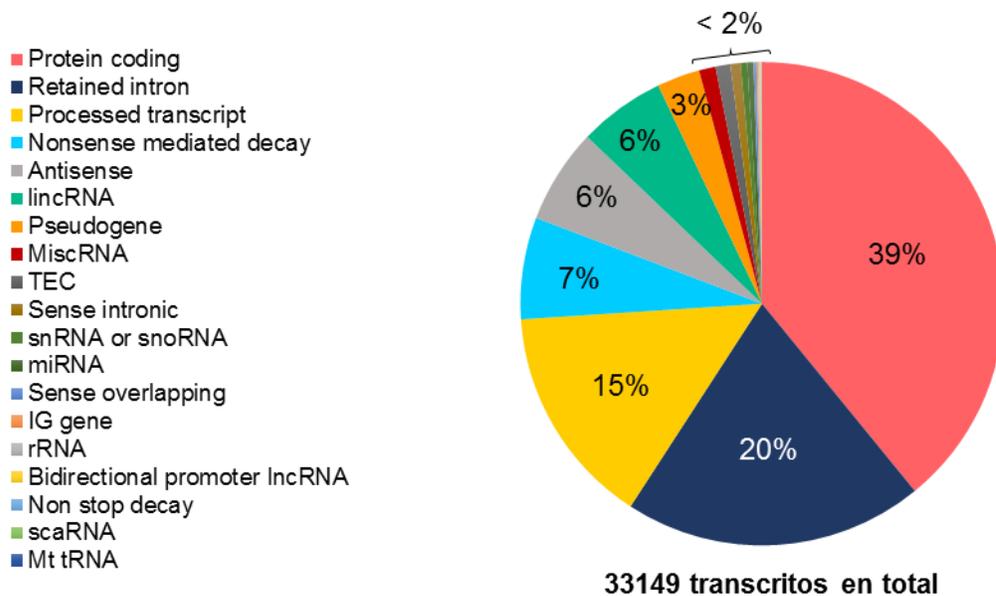


Figura 10. Identificación de ARNs asociados al ARNg en células J-Lat. La expresión relativa e identificación de ARNs asociados al gRNA de VIH-1 purificados por RAP se determinó mediante RNA-seq. La abundancia de los transcritos se determinó calculando los Transcritos por millón (TPM) y su enriquecimiento se definió como TPM pulldown/TPM input (>2). En este caso los resultados corresponden a una única réplica biológica de células J-Lat activadas. El gráfico muestra el enriquecimiento de los diferentes tipos de ARNs identificados por RNA-seq.

Tabla 5. ARNnc largos intergénicos más enriquecidos a partir de la captura de ARNg en células J-Lat

Nombre	Tamaño (bp)	Función
lnc-KCNE1B-155, 156 y 157	923	Desconocida
DANCR:12	748	Promueve crecimiento cáncer hepatocelular y colorrectal
LINC01237:8	470	Desconocida
SNHG17:12	973	Promueve cáncer gástrico
lnc-ZNF837-1:11	549	Desconocida
lnc-TET3-2:1	390	Desconocida
lnc-CYP2D6-1:1	667	Desconocida

4.2.2 Identificación de ARNs capturados a partir de células Jurkat infectadas y sin infectar.

Debido a lo anterior, en una segunda etapa se secuenciaron los ARN capturados en el modelo de infección descrito en el primer objetivo, que corresponde a células Jurkat (Linfocitos T) infectadas con pseudotipos de VIH-1. En este ensayo se lograron secuenciar tanto los input como los pulldown de dos réplicas biológicas de células infectadas y células sin infectar (Mock), por lo que para seleccionar los ARNnc largos más enriquecidos se eliminaron aquellos que aparecían en niveles similares tanto en el pulldown sin infectar como en el pulldown mock lo que se tradujo en la eliminación de aquellos RNAs donde $TPM_{pulldown\ infectadas} / TPM_{pulldown\ mock} < 2$. Luego, se seleccionaron aquellos enriquecidos respecto a su input $TPM_{pulldown\ infectadas} / TPM_{input\ infectadas} > 2$. Al aplicar dichos criterios de selección, se encontraron 2257 ARNs enriquecidos, donde al igual que lo observado en la captura en células J-Lat, la mayoría de los transcritos corresponde a ARNs codificantes (34%). Sin embargo, en este caso los segundos transcritos más abundantes corresponden a ARNnc largos intergénicos (31%), seguido por ARNs antisentido (11%) (Figura 11A). De los ARNnc largos intergénicos se seleccionaron los diez más enriquecidos. Además, se hizo un cruce de datos con otro ensayo

de captura realizado en el laboratorio en el que se capturaron por inmunoprecipitación los ARNs asociados a la proteína viral Rev y se encontraron dos ARNnc largos intergénicos enriquecidos en ambas estrategias de captura (Figura 11B). Se diseñaron partidores para los 12 ARNs seleccionados y se intentaron validar mediante RT-qPCR. Sin embargo, nuevamente no se detectó señal en ningún caso. Como se menciona más abajo, esto pudo deberse a una degradación parcial de la muestra en la etapa de sonicación.

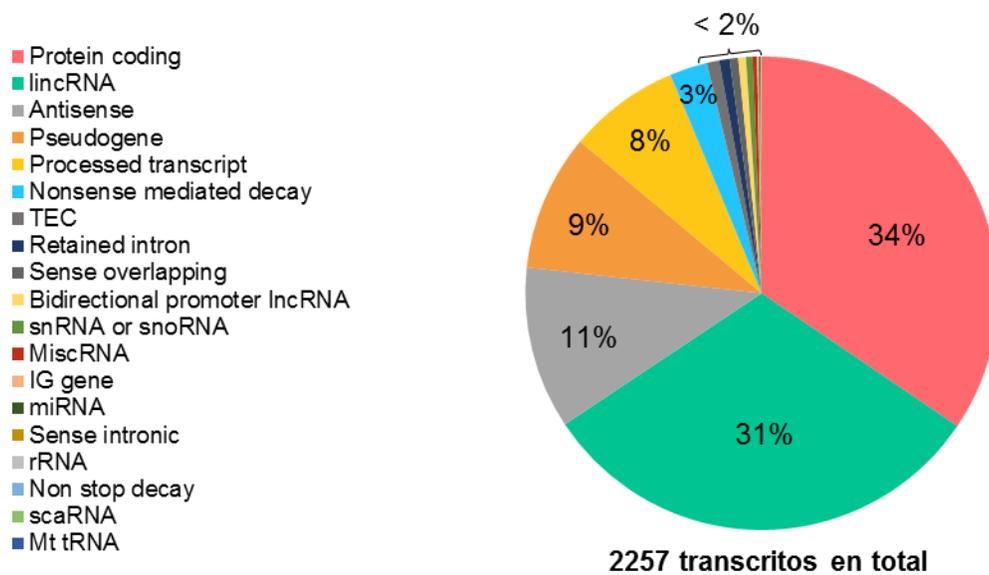


Figura 11. Identificación de ARNs asociados al ARNg en células Jurkat infectadas con pseudotipos de VIH-1. La expresión relativa e identificación de ARNs asociados al gRNA de VIH-1 purificados por RAP se determinó mediante RNA-seq. La abundancia de los transcritos se determinó calculando los Transcritos por millón (TPM) y su enriquecimiento se definió como i) TPM pulldown infectadas/TPM pulldown mock > 2 y ii) TPM pulldown/TPM input > 2. Los resultados corresponden a dos réplicas biológicas. El gráfico muestra el enriquecimiento de los diferentes tipos de ARNs identificados por RNA-seq.

Tabla 6. ARNnc largos intergénicos más enriquecidos a partir de la captura de ARNg en células Jurkat

Nombre	Tamaño (bp)	Función
Inc-ISL2-1:1	724	Desconocida
Inc-ITGB1-1:1	2051	Desconocida
Inc-LMNTD1-1:1	572	Desconocida
Inc-KIAA1755-15:1	720	Desconocida
LINC02361	1477	Desconocida
Inc-SEC24D-1:5	1512	Desconocida
Inc-MSH4-1:2	969	Desconocida
Inc-FTO-7:3	1458	Desconocida
Inc-CTSV-5:1	484	Desconocida
Inc-CNPY1-10:9	1741	Desconocida
Inc-KCNE1B-157:1	923	Desconocida
SNHG19:2	342	Desconocida

4.2.3 Re-análisis de datos de secuenciación.

Además de determinar la abundancia de transcritos calculando los TPM, existen diversas maneras de realizar un análisis de enriquecimiento. Debido a que no se pudo realizar la validación de los ARNs seleccionados, se decidió realizar un análisis bioinformático utilizando herramientas computacionales más robustas con los datos de secuenciación obtenidos a partir de la captura ARNs de células Jurkat infectadas y sin infectar. Esto se llevó a cabo con la ayuda de expertos en metabolismo y secuenciación de ARNs en el marco de una estadía corta de investigación en el laboratorio del Dr. Emiliano Ricci en la Ecole Normale Supérieure de Lyon en Francia. Allí, a partir de los datos crudos se realizó un mapa de flujo de trabajo presentado en la Figura 12. Cada uno de los pasos realizados se detalla a continuación.

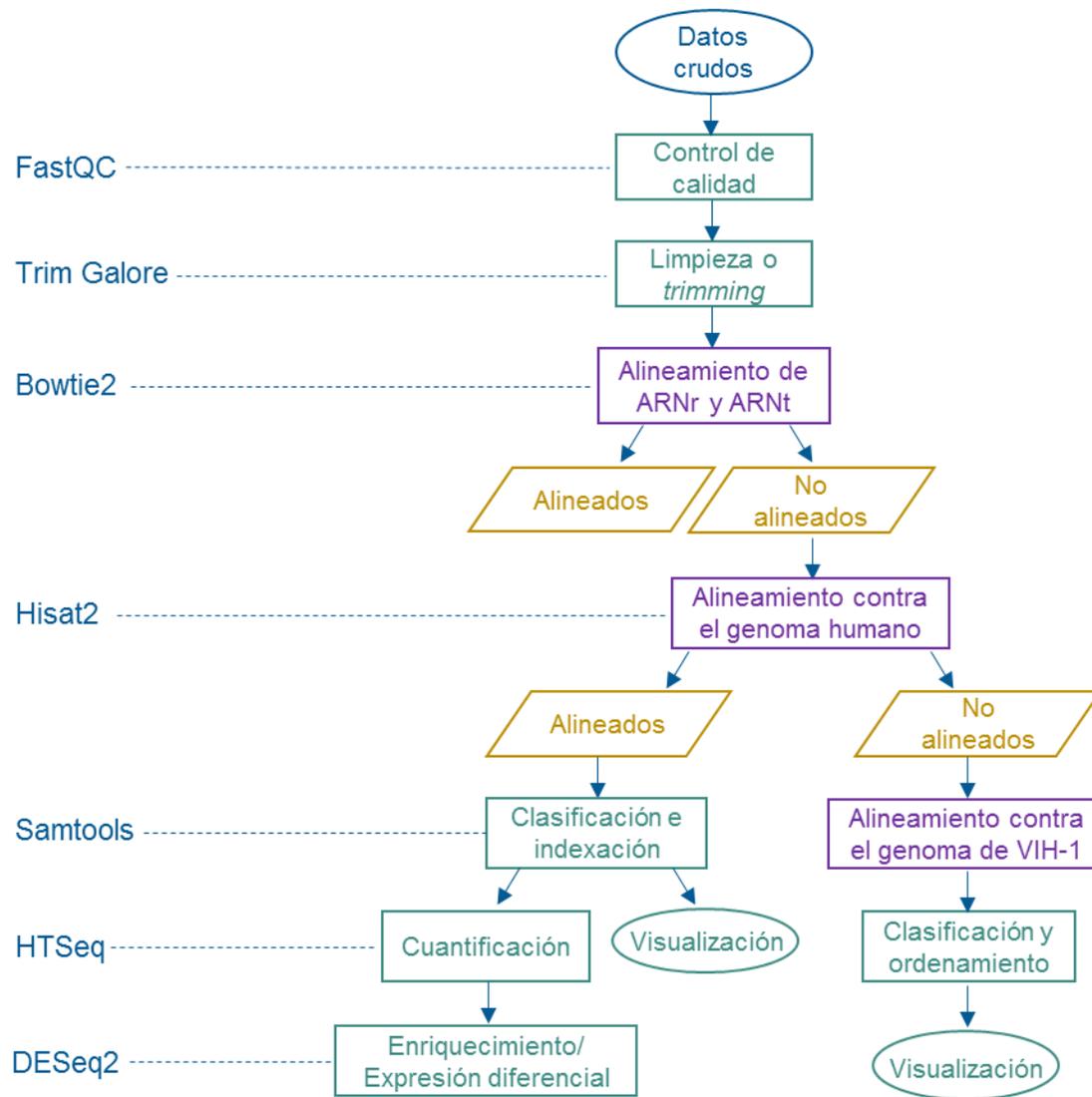


Figura 12. Mapa de flujo para el análisis de datos de secuenciación de ARN. En los recuadros se muestra secuencialmente los pasos seguidos desde el ingreso de los datos crudos hasta el resultado final de expresión diferencial en las muestras de ARN capturados de células Jurkat infectadas y sin infectar. En azul se muestran los programas utilizados para cada uno de los pasos.

4.2.4 Control de calidad y limpieza de las secuencias.

Antes de realizar cualquier análisis de los resultados de secuenciación, es fundamental realizar un control de calidad para asegurar que no existe ningún sesgo o problema que pudiese afectar el análisis de resultados y las conclusiones que se puedan obtener de éstos. Para evaluar la calidad de las secuencias se utilizó el programa FastQC v0.11.5. FastQC es una herramienta ampliamente utilizada para evaluar la calidad de los datos crudos de secuenciación masiva, pues permite tener una visión general de la calidad de distintos aspectos de las secuencias y, a su vez, muestra cada análisis de una manera visualmente sencilla e indica advertencias en aquellos análisis que presenten algún problema. Al realizar el control de calidad de los datos de secuenciación de todas las muestras, se encontró que el tamaño promedio de las secuencias o *reads* fue de 100 pb aproximadamente (Figura 13A y 13D). A su vez la calidad de la secuencia por base (Figura 13B) y por secuencia (Figura 13C) de todas las muestras mostraron valores altos con un *score* mayor a 30, lo que indica que la calidad fue muy buena y permite continuar con su análisis. Posteriormente se realizó la limpieza o *trimming* de las secuencias utilizando el programa TrimGalore v0.6.4 que identifica y elimina las secuencias de adaptadores. TrimGalore utiliza las primeras 13 pb de los adaptadores estándar de Illumina, pero también permite otras secuencias de adaptadores y es compatible con librerías *paired-end*.

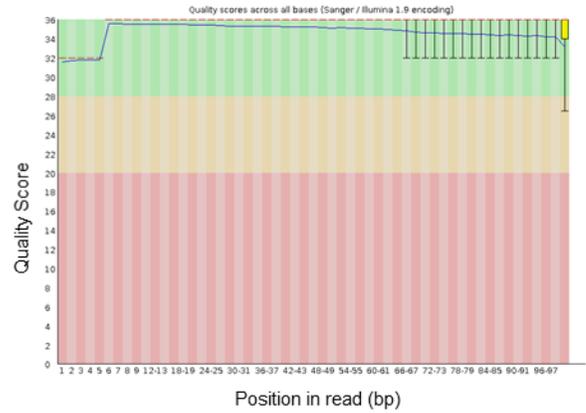
4.2.5 Alineamiento contra ARNr y ARNt.

A pesar de que previo a la preparación de genotecas las muestras fueron tratadas con RiboMinus para eliminar los ARN ribosomales (ARNr), se sabe que este paso no es completamente eficiente, por lo que se hizo un alineamiento de todas las secuencias de cada muestra contra una base de datos de los ARN ribosomales y de transferencia (ARNt) del genoma humano, utilizando el programa Bowtie2 v2.3.4.1.

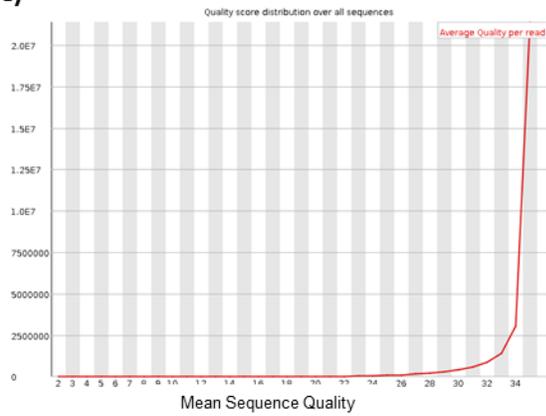
A)

Measure	Value
Filename	GM1598-1_R1_001.fastq.gz
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	28774528
Sequences flagged as poor quality	0
Sequence length	35-101
%GC	56

B)



C)



D)

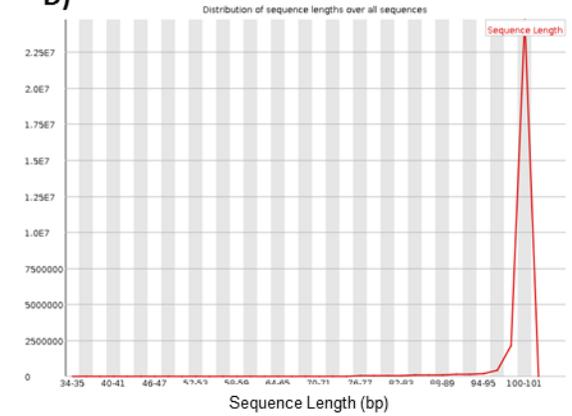


Figura 13. Control de calidad de una muestra representativa. Estos resultados corresponden al control de calidad realizado con FastQC de las secuencias de una réplica de una muestra Input mock. A) Estadísticas generales. B) Control de calidad por base. C) Control de calidad por secuencia. D) Tamaño de las secuencias o *reads*.

Bowtie2 es una herramienta desarrollada para alinear secuencias con secuencias de referencias largas de una manera eficiente en términos de uso de memoria. El porcentaje de alineamiento contra la base de datos de ARNr y ARNt en los distintos tipos de muestra se presenta en la Tabla 7. Cabe destacar que en las muestras pulldown mock se encontró entre un 39% y 55% de alineamiento y en el caso de las muestras pulldown infectadas se encontró alrededor de un 22%. Esto puede deberse a que en estas últimas hay un mayor número de secuencias correspondientes al ARNg de VIH-1. Finalmente, se descartaron todas las secuencias alineadas y se recuperaron las secuencias que no alinearon. Así, se eliminaron aquellas secuencias correspondientes a ARNr y ARNt.

Tabla 7. Porcentaje de alineamiento contra una base de datos de ARNr y ARNt.

Muestra	Tratamiento	Réplica	Nº inicial de reads	Alineamiento (%)
1	Input Mock	1	28.459.924	57,71
2	Input Infected	1	22.882.262	50,19
3	Input Mock	2	21.032.150	48,99
4	Input Infected	2	23.360.787	44,92
5	Pulldown Mock	1	46.518.849	54,77
6	Pulldown Infected	1	38.889.248	22,03
7	Pulldown Mock	2	34.755.810	38,69
8	Pulldown Infected	2	35.578.998	23,54

4.2.6 Alineamiento contra el genoma humano (hg38).

Una vez eliminadas las secuencias correspondientes a ARNr y ARNt, se realizó el mapeo contra el genoma humano (hg38) utilizando el programa Hisat2 v2.1.0 (Tabla 8). Hisat2 es una herramienta rápida y sensible que permite alinear secuencias contra genomas de referencia, como el genoma humano. Las secuencias de los inputs mostraron un alto porcentaje de alineamiento, donde se observó un alineamiento mayor al 85% en todos los casos. En las muestras pulldown mock se observó un porcentaje de alineamiento mucho mayor que en las

muestras pulldown infectadas. Al igual que en el caso anterior, esto puede deberse a que muchas de las secuencias de estas muestras corresponden al ARNg de VIH-1.

Tabla 8. Porcentaje de alineamiento contra el genoma humano (hg38)

Muestra	Tratamiento	Réplica	N° inicial de reads	Alineamiento (%)
1	Input Mock	1	12.462.631	95,25
2	Input Infected	1	11.977.030	91,3
3	Input Mock	2	11.392.070	89,07
4	Input Infected	2	13.658.758	86,73
5	Pulldown Mock	1	21.341.172	53,86
6	Pulldown Infected	1	30.426.540	25,64
7	Pulldown Mock	2	21.522.442	63,98
8	Pulldown Infected	2	27.275.884	30,36

4.2.7 Alineamiento o mapeo contra el genoma de VIH-1 (pNL4.3).

Con las secuencias que no mapearon contra el genoma humano, se hizo un mapeo contra el genoma de VIH-1 (Tabla 9). Específicamente, contra la secuencia de pNL4.3- Δ Env, pues este es el vector proviral con el que se prepararon las partículas virales en los experimentos realizados. Las muestras input infectadas mostraron un porcentaje de alineamiento cercano a un 0,6% a diferencia de las muestras input y pulldown mock que, como era de esperar, mostraron un alineamiento muy bajo. En cambio, las muestras pulldown infectadas mostraron un porcentaje de alineamiento mucho mayor, entre un 38% y un 59%. Por un lado, esto reafirma que el método de captura permite enriquecer significativamente el ARNg de VIH-1 y, por otro lado, muestra que efectivamente gran porcentaje de las secuencias de las muestras pulldown infectadas corresponden al ARNg de VIH-1 y, es por esto, que en estas muestras el porcentaje de alineamiento contra el genoma humano y los ARNr fue más bajo.

Tabla 9. Porcentaje de alineamiento contra el genoma de VIH-1 (pNL4.3ΔEnv)

Muestra	Tratamiento	Réplica	N° inicial de reads	Alineamiento(%)
1	Input Mock	1	1.279.970	0,01
2	Input Infec	1	2.216.388	0,59
3	Input Mock	2	2.626.944	0,00
4	Input Infec	2	3.670.074	0,57
5	Pulldown Mock	1	10.242.156	0,29
6	Pulldown Infec	1	22.917.266	58,92
7	Pulldown Mock	2	8.216.292	0,04
8	Pulldown Infec	2	19.265.500	38,56

4.2.8 Visualización del alineamiento contra el genoma de VIH-1.

Las secuencias que alinearon contra el genoma de VIH-1 se ordenaron e indexaron utilizando la herramienta SAMtools1.10, para posteriormente poder visualizar la cobertura del alineamiento de cada muestra con la aplicación Integrative Genomic Viewer (IGV) (Figura 14). Como era de esperar, se observó una cobertura completa en las muestras input infectadas. Sin embargo, en el caso de las muestras pulldown infectadas sólo se observó una cobertura de la primera mitad del genoma. Esa región es justamente donde hibrida la sonda antisentido utilizada para capturar el ARNg. Particularmente, corresponden a las regiones 5'UTR, *gag*, *pol*, *vif* y *vpr*.

4.2.9 Cuantificación y enriquecimiento de secuencias alineadas contra el genoma humano.

Por otro lado, las secuencias alineadas contra el genoma humano fueron cuantificadas con la herramienta HTSeq-count v0.6.1p1. Esta herramienta permite cuantificar cuántos *reads* o secuencias alinean de manera única con genes de la base de datos utilizada, cuantos no alinean o aquellos con alineamiento múltiple, entre otros. En las muestras input se encontró que entre un 20-30% de los *reads* alinean con genes. En las muestras pulldown, en cambio, este porcentaje es de un 30-40%. En total, en todas las muestras se observó que entre 4-6 millones

de *reads* alinean con genes y fueron estos los utilizados finalmente para analizar el nivel de enriquecimiento de los genes en las muestras pulldown. Posteriormente, se hizo el análisis de enriquecimiento de ARNs en las muestras. Para ello, se utilizó la herramienta DESeq2 en la plataforma RStudio. Esta herramienta importa los datos crudos obtenidos desde HTSeq y es ampliamente utilizada para análisis de expresión diferencial. Además, DESeq2 realiza una normalización interna de todos los datos, mediante un modelamiento de dispersión en el que ajusta los datos a un valor promedio y elimina aquellos que se comportan atípicamente o *outliers* (Love et al., 2014). En este caso, DESeq2 se utilizó para comparar la expresión diferencial o, en este caso, el enriquecimiento de RNAs entre las muestras pulldown e input y entre las muestras mock e infectadas.

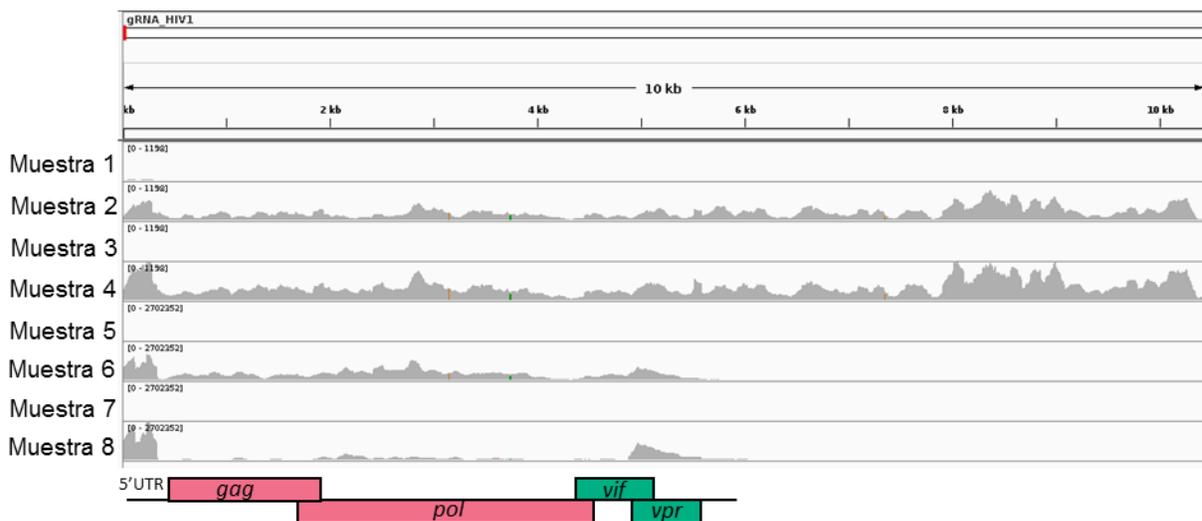


Figura 14. Visualización del alineamiento contra el genoma de VIH-1. Con el programa Hisat2 se hizo el alineamiento de las secuencias de las muestras de ARN capturados a partir de Células Jurkat infectadas y sin infectar con sus respectivos input (n=2) contra el genoma de VIH-1 (pNL4.3 ΔEnv). Los input (muestras 1, 2, 3 y 4) están en una misma escala y los pulldown (muestras 5, 6, 7 y 8) en otra escala para facilitar la visualización. Al inferior se muestra la región del ARNg que alinea en las muestras Pulldown.

Como una primera aproximación se hizo un análisis con todas las muestras en un mismo modelamiento, es decir, se incorporó en un único análisis: el tipo de muestra (input o pulldown), el tratamiento (mock o infectadas) y la réplica (n1 o n2). Se logró realizar un ajuste adecuado de los datos (Figura 15A) y un análisis de componente principal, donde se observó que el 74% de la varianza se debe al tipo de muestra (Figura 15B). Se seleccionaron los genes con un enriquecimiento mayor a 2 ($\log_{2}FC > 2$) y con un $p_{adj} < 0,05$. Así, se encontraron 83 ARNs codificantes y 6 ARNnc largos enriquecidos en las muestras pulldown infectadas. De los ARNnc largos 4 correspondían a ARNnc largos antisentido y 2 a ARNnc largos intergénicos.

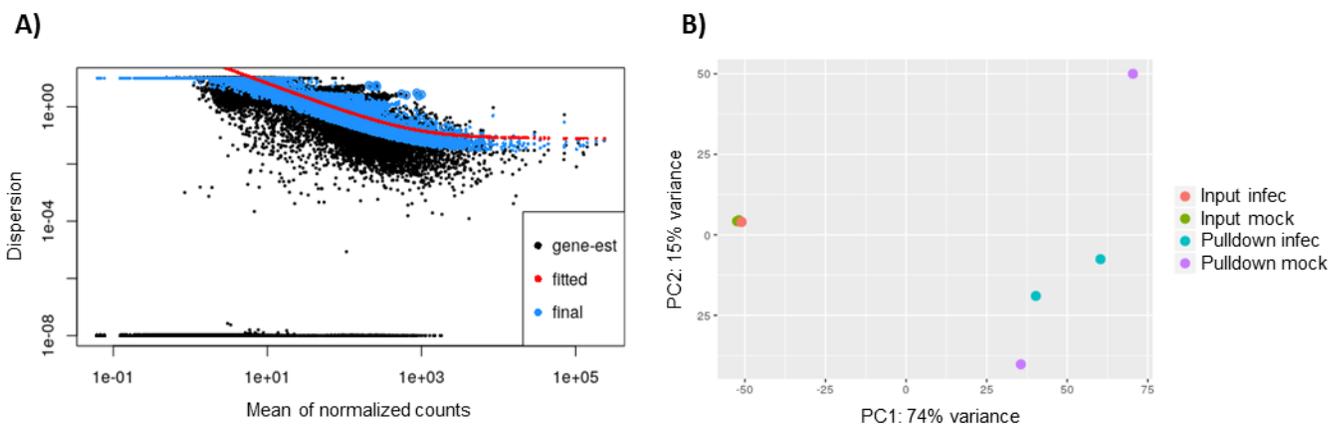


Figura 15. Análisis de expresión diferencial en un único modelamiento. Se hizo un análisis de expresión diferencial de todas las muestras en un mismo modelamiento utilizando la herramienta DESeq2. **A)** Dispersión de datos ajustados. **B)** Análisis de componente principal.

Como una segunda aproximación, se hicieron dos análisis por separado. Por un lado, se analizó el enriquecimiento en los pulldown sobre input de las muestras mock (pulldown mock/input mock) y por otro lado se realizó lo mismo en las muestras infectadas (pulldown infectadas/input infectadas). En ambos casos se observó un ajuste muy robusto de la dispersión de los datos (Figura 16A y 16C) y en el análisis de componente principal se encontró que un

84% y un 96% de la varianza se debe al tipo de muestra, en muestras mock e infectadas, respectivamente (Figura 16B y 16D). Luego, se seleccionaron aquellos ARNnc largos con un $\log_{2}FC > 2$ y $p_{adj} < 0,05$ en las muestras infectadas. De estos, se eliminaron aquellas con $\log_{2}FC < 2$ y/o $p_{adj} < 0,05$ en las muestras mock. Así quedaron 18 ARNnc largos de los cuales 9 corresponden a ARNnc largos intergénicos.

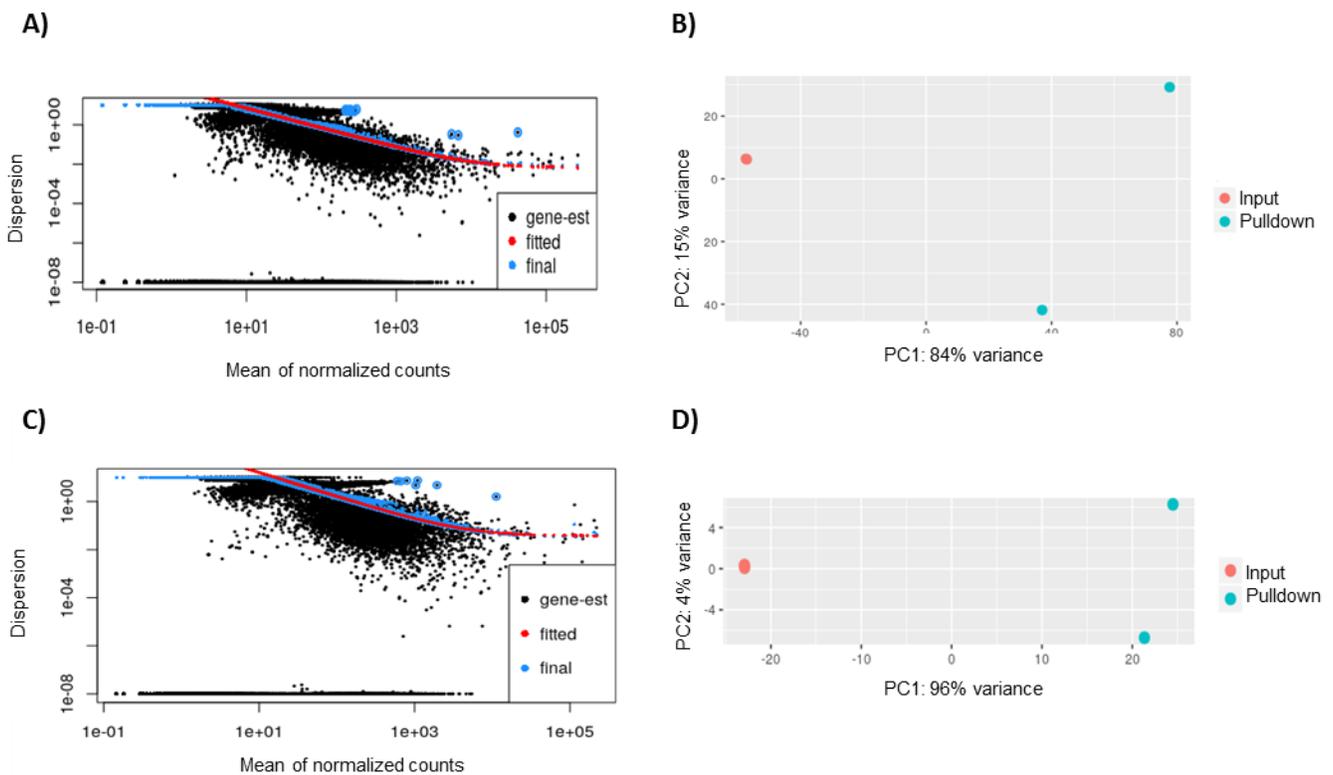


Figura 16. Análisis de datos para expresión diferencial de muestras mock e infectadas por separado. Se hizo un análisis de la expresión diferencial de pulldown/input de mock e infectadas por separado utilizando la herramienta DESeq2. **A)** Dispersión de datos ajustados de pulldown mock/input mock. **B)** Análisis de componente principal de pulldown mock/input mock. **C)** Dispersión de datos ajustados de pulldown infectadas/input infectadas. **D)** Análisis de componente principal de pulldown infectadas/input infectadas.

4.2.10 Visualización de secuencias alineadas contra el genoma humano.

Finalmente, los ARNnc largos intergénicos seleccionados con ambas aproximaciones se visualizaron en la plataforma IGV y también con Genome Browser UCSC. Con estas herramientas se descartaron algunos ARNnc largos cuyas secuencias tenían una alta cantidad de elementos repetidos o cuya cobertura fuera similar a la del mock. De esta manera, se seleccionaron 4 ARNnc largos intergénicos (Tabla 10) para los cuales se diseñaron partidores con el objetivo de validarlos mediante RT-qPCR. A partir de estos resultados se concluye que sí hay ARNnc largos que forman parte de los mRNPs del ARNg de VIH-1, de los cuales se seleccionaron 4 para ser validados.

Tabla 10. ARNnc largos intergénicos seleccionados de acuerdo al análisis de expresión diferencial.

Nombre	Función
LINC00938	Marcador de sobrevivida en cáncer de mama y de pulmón
lnc-ISL2-1	Desconocida
HAR1A	Marcador de cáncer hepatocelular
lnc-TMEM105-2	Desconocida

4.2.11 Diseño de partidores y validación mediante RT-qPCR.

Se diseñaron partidores para los 4 ARNnc largos intergénicos seleccionados donde se tuvo el cuidado de elegir las regiones donde había mayor cobertura de acuerdo con lo observado en la plataforma IGV para asegurar la amplificación independiente de la integridad de los ARNs (Figura 17). Sin embargo, no fue posible amplificar ninguno de los ARNnc largos a partir de muestras pulldown infectadas de una nueva captura.

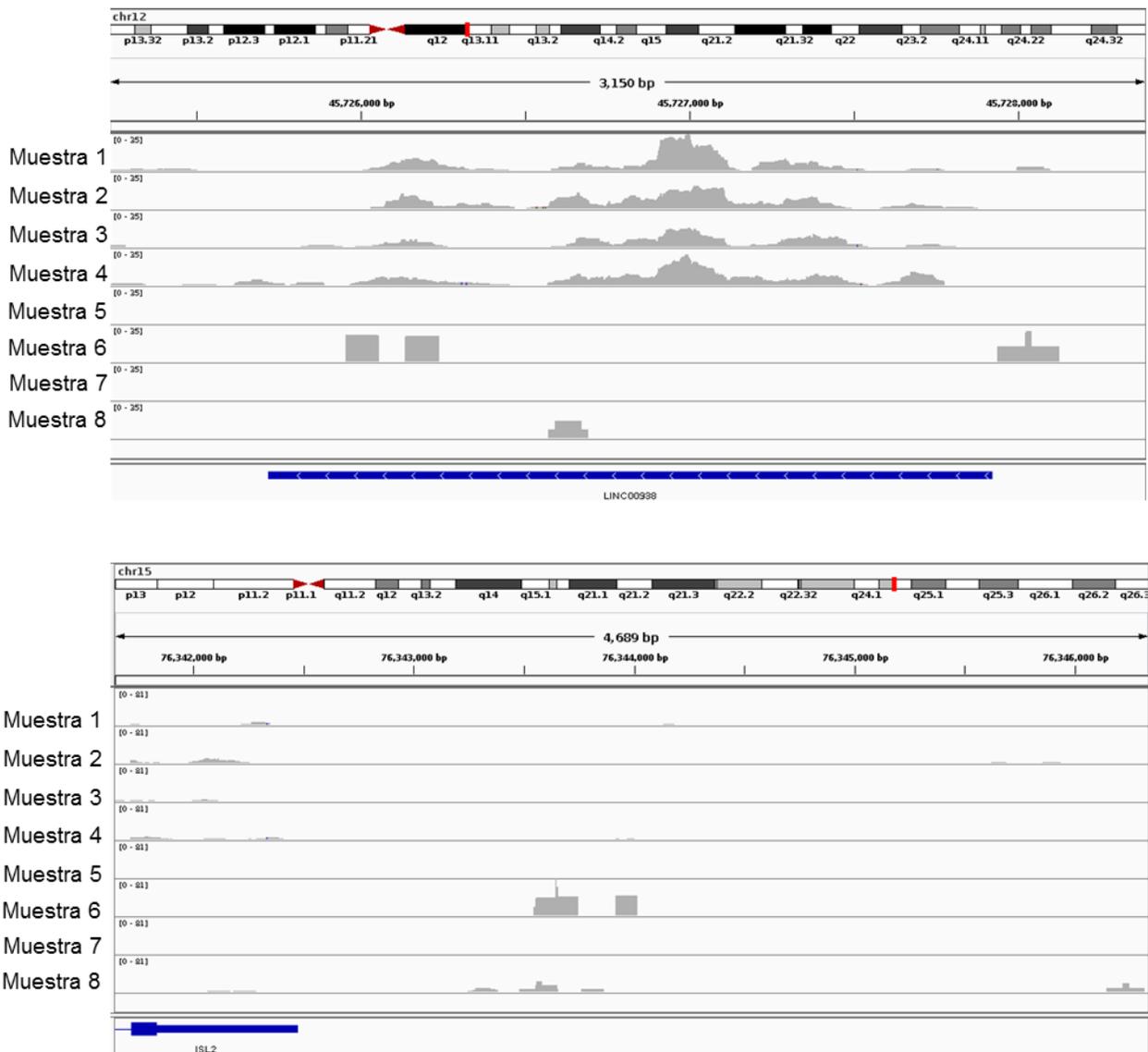


Figura 17. Visualización de dos de los ARNnc largos intergénicos seleccionados de acuerdo al análisis de expresión diferencial. A partir de los resultados de RNA-seq de los ARNs capturados de células Jurkat infectadas y mock, se hizo un re-análisis bioinformático de expresión diferencial. Se identificaron 4 ARNnc largos intergénicos que presentaron una expresión diferencial significativa. Se muestra la cobertura de dos de los ARNnc largos intergénicos seleccionados: LINC00938 y Inc-ISL2-1.

4.3 OE3. Determinar el efecto de un ARNnc largo asociado al ARNg de VIH-1 sobre la replicación viral.

4.3.1 Evaluación de los niveles de expresión relativa de los ARNnc largos seleccionados en diferentes líneas celulares.

El hecho de no haber detectado los ARNnc largos seleccionados mediante RT-qPCR, pudo deberse a diversos motivos, como por ejemplo a una baja concentración de éstos en células Jurkat. Es por esto que como estrategia para evaluar el efecto de alguno de los ARNnc largos sobre la replicación viral, se midió su nivel de expresión en distintas líneas celulares. Se eligieron 6 líneas celulares humanas diferentes: Jurkat (linfocitos T), HEK (riñón), HeLa (epitelio cervicouterino), U2OS (tejido óseo), C20 (microglías) y Huh7 (hepatocitos) y se evaluó el nivel de los ARNnc largos en condiciones con y sin infección con VIH-1. Para ello, se prepararon partículas virales mediante la transfección de células HEK 293T con el vector pCMV-VSVg, que codifica la glicoproteína G del virus de la estomatitis vesicular, y con pNL4.3ΔEnv, que codifica todas las proteínas de VIH-1 excepto las de la envoltura. Se cosecharon las células mock e infectadas 20 hpi. Para verificar si hubo una infección exitosa, se evaluaron los niveles de Gag mediante Western blot, donde se logró detectar niveles de Gag (p55) y sus derivados en todas las líneas celulares infectadas (Figura 18). A su vez, se realizó una extracción de ARN total y se midieron los niveles de los ARNnc largos seleccionados. El análisis se realizó tanto para células mock, como para células infectadas. Los niveles relativos de los ARNnc largos se normalizaron respecto a GAPDH y son presentados relativos a sus niveles en células Jurkat a las cuáles se les designó arbitrariamente un 100% de expresión de cada ARNnc largo (Figura 19). En primer lugar, se analizaron los niveles de Linc00938. Se observó una menor expresión de este ARN relativo a lo observado en Jurkat en todas las líneas celulares. Los niveles más bajos fueron detectados en células Huh7 (4 veces menos con respecto a Jurkat), tanto en células mock (Figura 19A) como infectadas (Figura 19B). En el caso de Inc-ISL2-1, se observaron niveles

bajos en todas las líneas celulares excepto en HEK. Por un lado, los niveles más bajos se observaron en células U2OS mock (Figura 19C) e infectadas (Figura 19D). Por otro lado, observamos niveles elevados con respecto a Jurkat en células HEK mock (Figura 19C) e infectadas (Figura 19D). Luego, se evaluó el nivel de expresión de HAR1A. En este caso se observaron niveles muy reducidos en todas las líneas celulares relativo a sus niveles en Jurkat, tanto en células mock (Figura 19E) como infectadas (Figura 19F). Los niveles más bajos con respecto a Jurkat fueron entre 15 veces en células U2OS y cerca de 800 veces en células HeLa infectadas. Sin embargo, la mayor diferencia se observó en células Huh7 en las cuales no se detectó expresión de HAR1A.

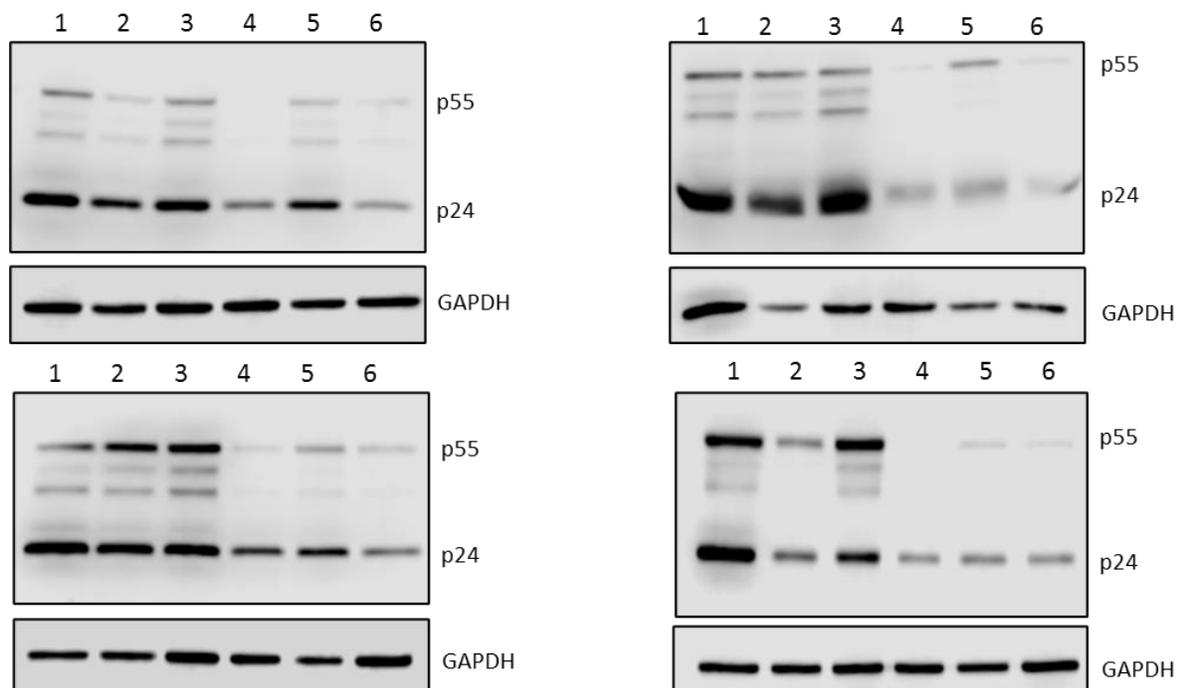


Figura 18. Infección de 6 líneas celulares diferentes con VIH-1. Se infectaron 6 líneas celulares diferentes: Jurkat, HEK, HeLa, U2OS, C20 y Huh7 con pseudotipos virales de VIH-1, preparados con los vectores pNL4.3 Δ Env y pCMV-VSVg. Se cosecharon las células 20 hpi. Los niveles de Gag y sus derivados fueron detectados mediante Western blot utilizando un anticuerpo monoclonal anti-p24. Cada imagen corresponde a un experimento independiente. (1: Jurkat, 2: HEK, 3: HeLa, 4: U2OS, 5: C20 y 6: Huh7).

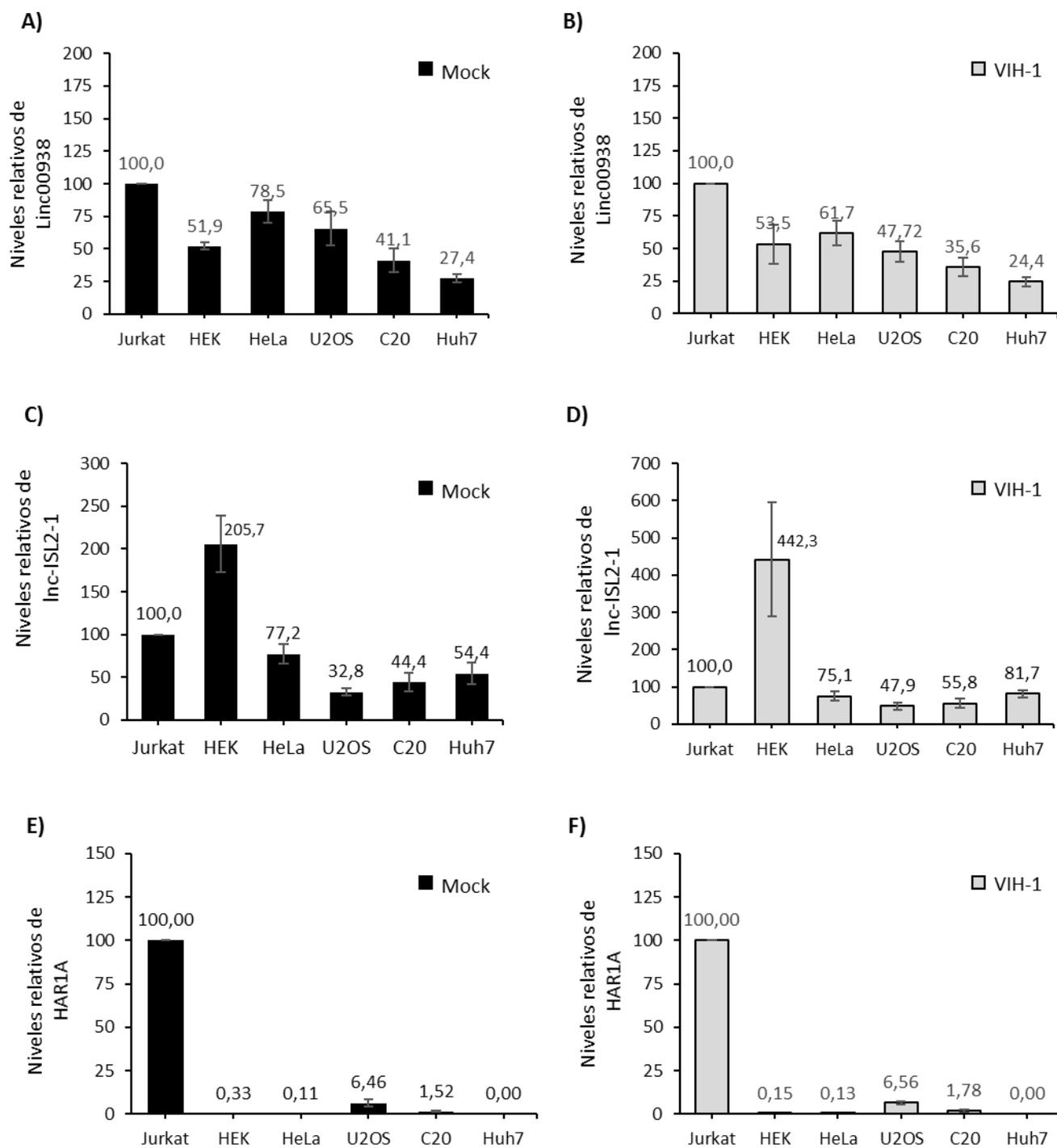


Figura 19. Expresión relativa de ARNnc largos en distintas líneas celulares. Se infectaron 6 líneas celulares diferentes: Jurkat, HEK, HeLa, U2OS, C2O y Huh7 con pseudotipos virales de VIH-1, preparados con los vectores pNL4.3 Δ Env y pCMV-VSVg. Se cosecharon las células 20 hpi. Los niveles relativos de Linc00938 (A, B); Inc-ISL2-1 (C, D) y HAR1A (E, F) se midieron por RT-qPCR en células mock e infectadas. Los valores de expresión relativa de cada ARNnc largo se muestran sobre cada columna. Éstos fueron normalizados por los niveles de GAPDH y corresponden al promedio \pm SEM de 4 experimentos independientes, relativos a su expresión en Jurkat.

Por otro lado, se quiso evaluar si había un cambio en los niveles de expresión relativa de los ARNnc largos en células infectadas. Para ello, se comparó el nivel de expresión de cada ARNnc largo normalizada por GAPDH en células infectadas relativo al mock (Figuras 20, 21 y 22). Sin embargo, no se observó una expresión diferencial significativa en ninguno de los casos. Debido a la baja expresión relativa de HAR1A en todas las líneas celulares analizadas, es que se decidió evaluar el posible efecto de este ARNnc largos sobre los niveles de la proteínas viral Gag.

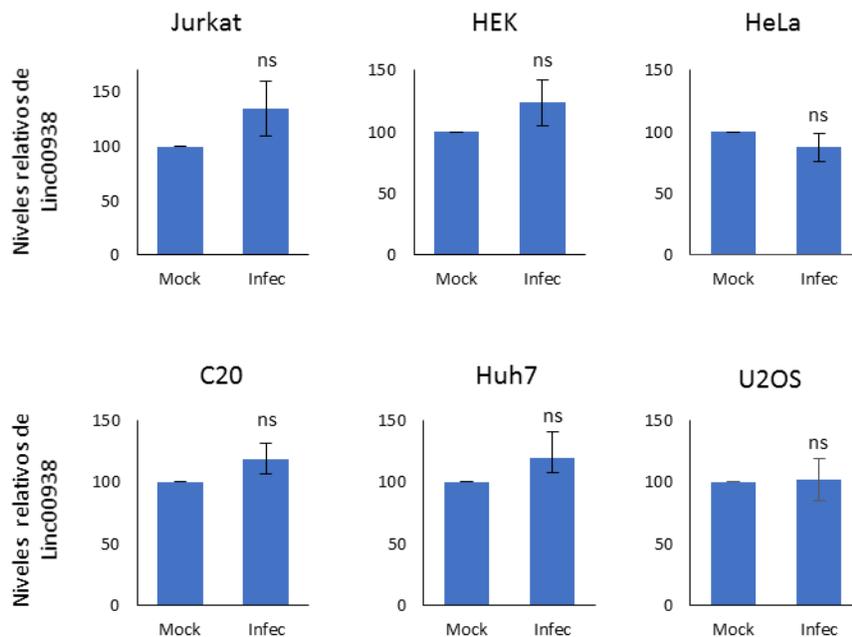


Figura 20. Expresión relativa de Linc00938 en células infectadas. Se infectaron 6 líneas celulares diferentes: Jurkat, HEK, HeLa, U2OS, C20 y Huh7 con pseudotipos virales de VIH-1, preparados con los vectores pNL4.3 Δ Env y pCMV-VSVg. Se cosecharon las células 20 hpi. Los niveles relativos de Linc00938 se midieron por RT-qPCR en células mock e infectadas y se normalizaron por los niveles de GAPDH. En este caso, se comparan los niveles de expresión de Linc00938 en células infectadas respecto a sus niveles en células mock. Los valores corresponden al promedio \pm SEM de 3 experimentos independientes, relativos a su control (NS: no significativo, t-test).

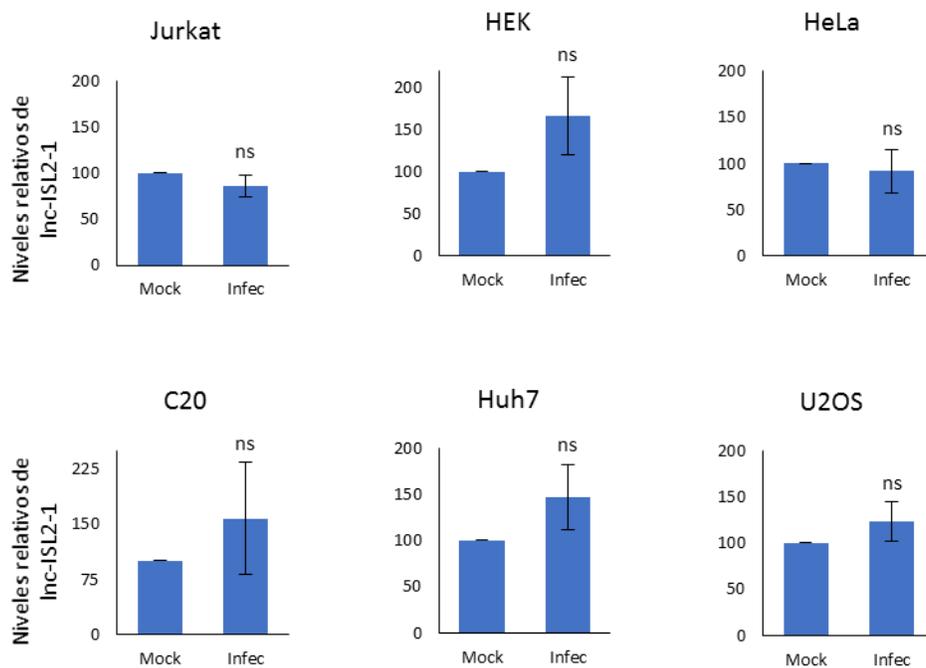


Figura 21. Expresión relativa de lnc-ISL2-1 en células infectadas. Se infectaron 6 líneas celulares diferentes: Jurkat, HEK, HeLa, U2OS, C20 y Huh7 con pseudotipos virales de VIH-1, preparados con los vectores pNL4.3 Δ Env y pCMV-VSVg. Se cosecharon las células 20 hpi. Los niveles relativos de lnc-ISL2-1 se midieron por RT-qPCR en células mock e infectadas y se normalizaron por los niveles de GAPDH. En este caso, se comparan los niveles de expresión de lnc-ISL2-1 en células infectadas respecto a sus niveles en células mock. Los valores corresponden al promedio \pm SEM de 3 experimentos independientes, relativos a su control (NS: no significativo, t-test).

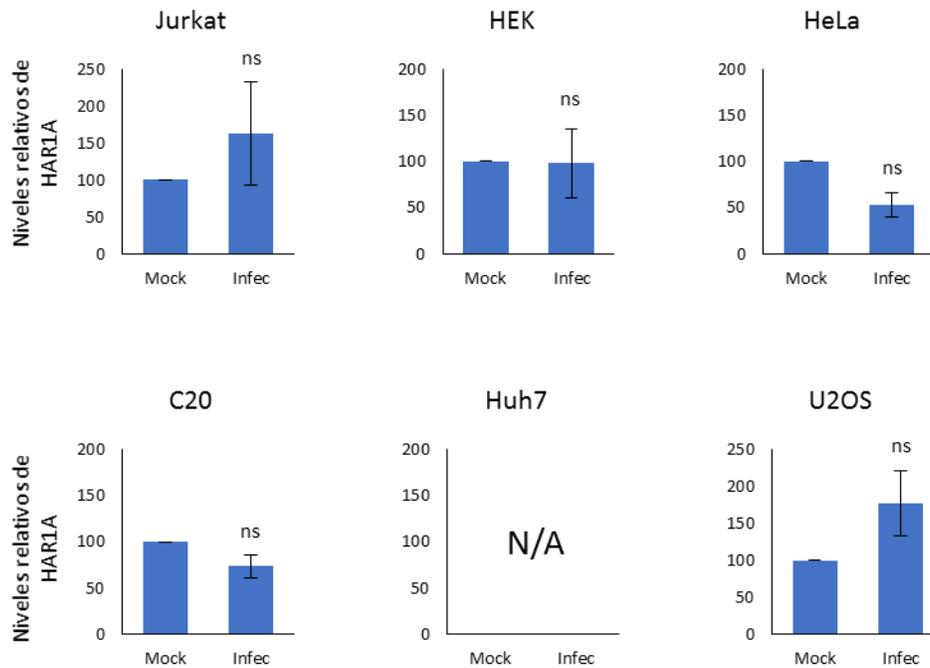


Figura 22. Expresión relativa de HAR1A en células infectadas. Se infectaron 6 líneas celulares diferentes: Jurkat, HEK, HeLa, U2OS, C20 y Huh7 con pseudotipos virales de VIH-1, preparados con los vectores pNL4.3 Δ Env y pCMV-VSVg. Se cosecharon las células 20 hpi. Los niveles relativos de HAR1A se midieron por RT-qPCR en células mock e infectadas y se normalizaron por los niveles de GAPDH. En este caso, se comparan los niveles de expresión de HAR1A en células infectadas respecto a sus niveles en células mock. En células Huh7 no se detectaron niveles de expresión de HAR1A tanto en células mock como en infectadas. Los valores corresponden al promedio \pm SEM de 3 experimentos independientes, relativos a su control (NS: no significativo, t-test).

4.3.2 Efecto de HAR1A sobre los niveles de Gag y del ARNg.

HAR1A es un ARNnc largo intergénico con un tamaño aproximado de 2700 pb. Algunos estudios han mostrado que HAR1A es un biomarcador de carcinoma hepatocelular, gliomas y cáncer oral. Específicamente, se han asociado bajos niveles de expresión de este ARNnc largo con un mal pronóstico en estos tipos de cáncer y se ha sugerido que HAR1A podría actuar como un supresor de tumores (Shi et al., 2019; Zou et al., 2017). Sin embargo, no se ha relacionado con la regulación de la replicación viral. Debido a los significativos bajos niveles de HAR1A en todas las líneas celulares estudiadas en esta tesis, se decidió evaluar el efecto de

éste sobre la replicación viral mediante la sobreexpresión de HAR1A y la detección de los niveles de la proteína viral Gag. Se eligieron 3 líneas celulares diferentes: i) HeLa, pues es un modelo permisivo ampliamente utilizado en investigación en VIH-1; ii) Huh7, pues en esta línea no se observó expresión de HAR1A y por último iii) C20, pues es una línea de microglías humanas que es naturalmente susceptible y permisiva a la infección VIH-1, por lo que tiene un sentido biológico relevante. Se transfectaron las distintas líneas celulares con 1 µg y 2 µg de pcDNA-HAR1A, que codifica a HAR1A bajo el control del promotor CMV. Se utilizó como control el vector pcDNA-GlobinRen, que codifica a la luciferasa de *Renilla*, pero carece de intrones por lo que tiene bajos niveles de traducción (Ricci et al., 2009). A las 24 hpt, se lavaron las células y se infectaron. A las 20 hpi se cosecharon las células y se realizó una extracción de proteínas totales para analizar mediante Western blot los niveles de Gag y por otro lado se realizó una extracción de ARN totales para evaluar los niveles de HAR1A y también del ARNg de VIH-1. A partir de los Western blot (Figuras 23A, 23B y 23C) se hizo un análisis de densitometría para cuantificar los niveles de Gag normalizados por GAPDH (Figuras 23D, 23E y 23F). En este caso, no se observaron cambios significativos en los niveles de Gag en ninguna de las 3 líneas celulares con ninguna de las dos cantidades de HAR1A transfectadas. Se confirmó la sobreexpresión de HAR1A en las 3 líneas celulares mediante RT-qPCR, donde se observó un aumento significativo de éste en todas las condiciones relativas al control (Figuras 23G, 23H y 23I). Por último, no se observó un cambio significativo en los niveles de ARNg de VIH-1 en ninguna de las 3 líneas celulares (Figura 24A, 24B y 24C). Los resultados anteriores muestran que, en las condiciones probadas, HAR1A no tiene un efecto sobre los niveles de Gag de VIH-1 y tampoco sobre los niveles del ARNg.

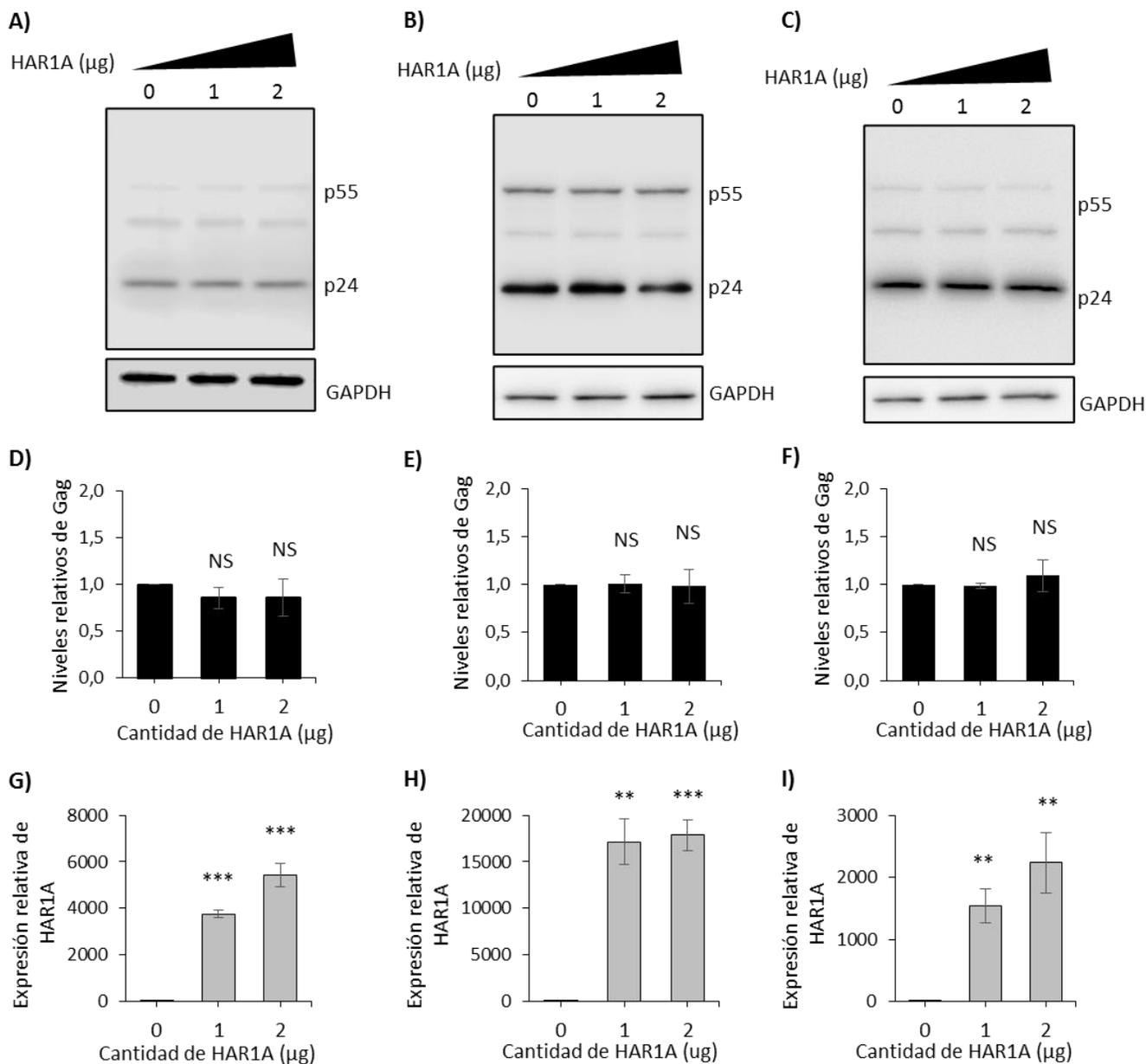


Figura 23. Efecto de la sobreexpresión de HAR1A sobre los niveles de la proteína viral Gag. Se sobreexpresó HAR1A en 3 líneas celulares mediante transfección con PEI, utilizando 1 µg y 2 µg del vector pcDNA-HAR1A. Como control se utilizó el vector pcDNA-Renilla. Se infectaron las células 24 hpt con virus pNL4.3-ΔEnv pseudotipados con VSVg. Se cosecharon las células 20 hpi. Los niveles de expresión de la proteína viral Gag fueron analizados mediante Western blot usando anticuerpo monoclonal anti-p24. Se muestra una imagen representativa para células HeLa (A), Huh7 (B) y C20 (C). Los niveles relativos de Gag se obtuvieron mediante densitometría de 3 réplicas biológicas en células HeLa (D), Huh7 (E) y C20 (F). Los niveles de HAR1A fueron medidos por RT-qPCR en células HeLa (G), Huh7 (H) y C20 (I). Los valores de expresión de HAR1A fueron normalizados por GAPDH. En todos los casos, los valores corresponden al promedio ± SEM de 3 experimentos independientes, relativos a su control (*: p<0,05; **: p < 0,01; ***: p < 0,001 y NS: no significativo, t-test).

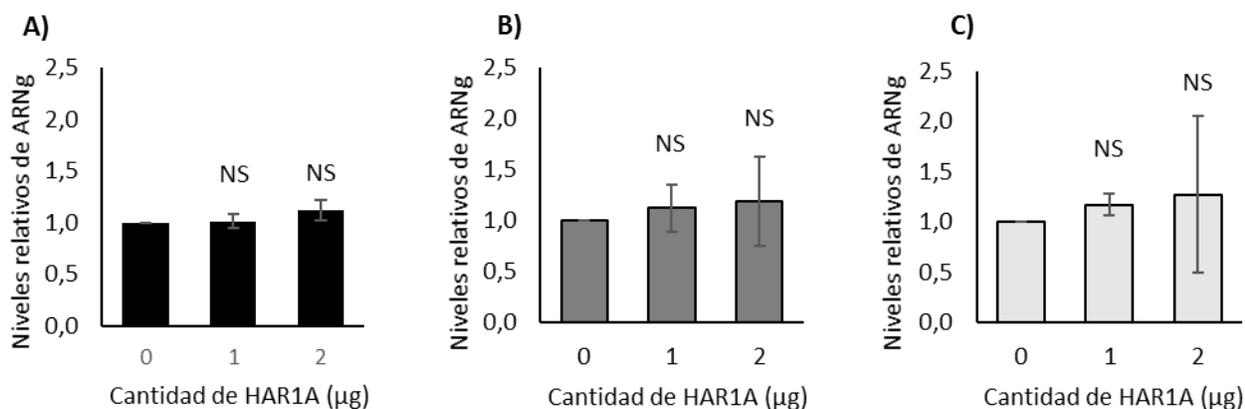


Figura 24. Efecto de la sobreexpresión de HAR1A sobre los niveles de ARNg de VIH-1. Se sobreexpresó HAR1A en 3 líneas celulares mediante transfección con PEI, utilizando 1 µg y 2 µg del vector pcDNA-HAR1A. Como control se utilizó el vector pcDNA-Renilla. Se infectaron las células 24 hpt con virus pNL4.3-ΔEnv pseudotipados con VSVg. Se cosecharon las células 20 hpi. Los niveles de expresión del ARNg de VIH-1 fueron medidos por RT-qPCR en células HeLa (A), Huh7 (B) y C20 (C). Los valores de expresión de ARNg fueron normalizados por GAPDH y corresponden al promedio ± SEM de 3 experimentos independientes, relativos a su control NS: no significativo, t-test).

5 DISCUSIÓN

Durante las etapas tardías del ciclo replicativo, el ARNg de VIH-1 cumple dos funciones relevantes: es el ARNm a partir del cual se sintetizan las proteínas virales Gag y Gag-Pol y es el genoma viral empaquetado en las nuevas partículas virales liberadas. Así, los factores celulares que forman parte de los mRNPs del ARNg han sido de gran interés, pues determinan su destino citoplasmático. Se han identificado decenas de proteínas que forman parte de los mRNPs de ARNg de VIH-1, sin embargo, hasta la fecha no se conocía ningún ARNnc largo que formara parte de estos complejos, a pesar de que han mostrado ser reguladores versátiles de la expresión génica a nivel transcripcional y postranscripcional. El objetivo del presente trabajo fue identificar ARNnc largos que forman parte de los mRNPs del ARNg de VIH-1 y determinar si alguno de ellos afecta la replicación del virus.

Estandarización de la preparación y purificación de sondas para aislar el ARNg de VIH-1 y sus ARNs asociados.

Con el objetivo de capturar el ARNg de VIH-1 se optimizó el protocolo de Purificación de ARN Antisentido publicado por McHugh et al., 2015 y se utilizó una sonda antisentido preparada de acuerdo con lo descrito en la sección 3.10. Tanto en células J-Lat activadas (modelo de latencia) como en células Jurkat infectadas, se logró enriquecer aproximadamente 6000 veces el ARNg relativo a GAPDH (Figuras 8A y 9C), mostrando una alta eficiencia y especificidad de la sonda en ambos modelos. Aún más, se logró detectar la presencia del tRNA^{Lys3} en las muestras de ARN capturado, indicando que además de capturar al ARNg es posible recuperar los ARNs asociados a él. En el año 2017, Knoener et al. realizaron una captura del ARNg de VIH-1 y mediante espectrometría de masas lograron identificar decenas de proteínas que interactúan directa o indirectamente con el ARNg e incluso validaron el efecto de algunas de ellas silenciándolas. Ellos lograron enriquecer alrededor de 40 veces el ARNg relativo a GAPDH

con una estrategia muy similar a la realizada en esta tesis. Utilizaron células Jurkat infectadas con pseudotipos virales y tratadas con formaldehído. Una de las grandes diferencias que podría explicar los altos niveles de enriquecimiento obtenidos en esta tesis es la naturaleza de la sonda. Knoener et al., utilizaron como sonda un oligonucleótido de 30 nt dirigido a un única región del ARNg y biotinilado sólo en un extremo. En cambio, en esta tesis se utilizó una sonda dirigida a la región Gag-Pol completa, fragmentada en tamaños de entre 50-300 nt y con un 50% de los UTPs biotinilados. Por un lado, el tamaño y la región contra la que está dirigida la sonda, permitiría que ésta pueda acceder a muchas de las regiones disponibles del ARNg y no a una sola, logrando capturar una mayor cantidad y diversidad de complejos. Además, al tener más moléculas de biotina, la recuperación con las perlas magnéticas acopladas a estreptavidina podría tener una mayor eficiencia.

La estrategia de preparación de sonda desarrollada en esta tesis puede ser utilizada potencialmente para cualquier ARN de secuencia conocida. Basta clonar parte de la secuencia del ARN de interés en un vector que permita la transcripción antisentido como los vectores comerciales pBlueScript y realizar la transcripción con algún nucleótido conjugado con biotina o una molécula que cumpla la misma función. Además, este tipo de sonda no sólo es útil para ensayos de captura como este, sino que también se puede utilizar para la detección subcelular de ARNs. Esta sonda permitió la detección de ARNg en un ensayo de Hibridación Fluorescente *In Situ* (FISH) (datos no mostrados), siguiendo la misma estrategia de preparación de la sonda, pero con digoxigenina en lugar de biotina.

La estrategia de captura permitió identificar ARNnc largos que forman parte de los mRNPs del ARNg de VIH-1, pero se puede optimizar.

Hasta la fecha se han caracterizado funcionalmente diez ARNnc largos que afectan la replicación de VIH-1, ya sea a nivel transcripcional o afectando el metabolismo de las células

infectadas. Sólo NEAT1 ha mostrado afectar la replicación del virus a nivel postranscripcional, sin embargo, no se ha demostrado una interacción directa o indirecta de éste con el ARNg. El presente trabajo es el primero en identificar ARNnc largos que forman parte de los mRNPs del ARNg de VIH-1. En particular, utilizando las herramientas bioinformáticas y el flujo de trabajo señalado en la Figura 12, se identificaron 4 ARNnc largos los cuales se intentaron validar mediante RT-qPCR sin éxito (Tabla 10).

El no haber logrado validar la presencia de ninguno de los 4 ARNnc largos mediante RT-qPCR puede deberse a diversos motivos. Una posibilidad es que los ARNs capturados no estén íntegros, dificultando la posibilidad de amplificarlos con partidores específicos. Esta idea surge al observar la cobertura del ARNg capturado. Como se puede apreciar en la Figura 14, en las muestras pulldown sólo se recupera la región Gag-Pol contra la que está diseñada la sonda, a pesar de que en las muestras input está la secuencia completa. De aquí es posible deducir que el ARNg podría estar fragmentado, y no sólo él, sino todos los ARNs asociados. La fragmentación de los ARNs ocurriría en los pasos iniciales de nuestra estrategia de captura, debido a la combinación del entrecruzamiento con 0,1% de formaldehído y la sonicación. Se ha observado que en estas condiciones las regiones de ARN más expuestas se rompen en fragmentos de distintos tamaños dependiendo de la intensidad de la sonicación y la concentración del formaldehído. Este fenómeno incluso se ha utilizado como alternativa para realizar *footprinting* en ensayos de inmunoprecipitación (Singh et al., 2014). Así, la estrategia de captura estandarizada en este trabajo podría optimizarse evaluando la integridad de los ARNs en distintas condiciones de sonicación y entrecruzamiento. En este trabajo se decidió utilizar formaldehído pues esta molécula es capaz de entrecruzar ARN con proteínas y también proteínas con proteínas a una distancia aproximada de 2Å (Hoffman et al., 2015), por lo que permitiría identificar una mayor población ARNs que forman parte de los mRNPs, incluyendo

aquellos unidos mediante más de una proteína. Sin embargo, esto trae como consecuencia la formación de grandes complejos macromoleculares, haciendo de la sonicación un paso necesario para la solubilidad de éstos. Otra posibilidad para evitar la fragmentación de ARNs es utilizar otras técnicas de entrecruzamiento, como la irradiación directa con luz UV a 254 nm o con la incorporación de nucleósidos fotoactivables como la 4SU e irradiación a 365 nm. Ambas estrategias entrecruzan solo ARNs con proteínas a una distancia cero, es decir, entrecruzan moléculas que interactúan directamente y además no generan reactividad entre proteínas (Castello et al., 2013; Pashev et al., 1991; Suchanek et al., 2005), por lo que no se forman grandes complejos macromoleculares y el paso de sonicación podría ser reducido o eliminado. Sin embargo, esta estrategia reduciría el rango de ARNs que podrían ser capturados.

Otro paso dentro de la estrategia de captura y análisis de resultados que podría optimizarse es la manera en que se calcula el enriquecimiento de ARNs. Este paso es crucial pues determina las conclusiones que se obtienen a partir del experimento realizado. Tal como se comentó en la sección 4.2.3, existen distintas maneras de calcular el enriquecimiento, como es el cálculo de la abundancia enriquecimiento relativo calculando la razón TPM pulldown/TPM input o utilizando otras herramientas bioinformáticas para cuantificar (HTSeq) y calcular el enriquecimiento (DESeq2). Ambas aproximaciones entregaron resultados diferentes (Tabla 6 y Tabla 10). A esto se suma la variabilidad asociada a la calidad de la muestra y la profundidad de la secuenciación por muestra. Una forma de reducir toda esta variabilidad es agregar un ARN externo, de secuencia y concentración conocidas, antes de la preparación de las librerías de ADNc. Esta estrategia se denomina *spike-in* y es usada para normalizar los datos entre muestras, incluso existe una mezcla de ARNs diseñados para este objetivo como el ERCC RNA Spike-In Mix (Invitrogen). Con la adición de este ARN control, se pueden realizar análisis de expresión diferencial o enriquecimiento más precisos, en especial en experimentos donde la

cantidad de material inicial entre muestras es variable (Locati et al., 2015; Lutzmayer et al., 2017).

La sobreexpresión de HAR1A no tuvo efecto en los niveles intracelulares de ARNg ni Gag.

Se decidió evaluar el efecto de alguno de los 4 ARNnc largos identificados como parte de los mRNPs del ARNg de VIH-1. Desafortunadamente no fue posible validar la presencia de ninguno de ellos mediante RT-qPCR, lo que pudo deberse a lo discutido anteriormente, así como también a una baja concentración de ellos en células Jurkat. A esto se suma el hecho de que esta línea celular es particularmente difícil de transfectar. Históricamente, han mostrado una muy baja eficiencia de transfección con distintos sistemas como lípidos o polímeros catiónicos (Olden et al., 2018). La técnica más recomendada es la electroporación, sin embargo, con esta estrategia la viabilidad celular se reduce significativamente. Por lo anterior, se evaluaron los niveles de los 4 ARNs seleccionados en otras líneas celulares. Al detectar muy bajos niveles de expresión relativa de HAR1A en comparación a los niveles en Jurkat (Figuras 19E y 19F), es que se decidió evaluar el efecto de la sobreexpresión de HAR1A en células HeLa, U2OS y C20. Sin embargo, no se observó ningún efecto en los niveles intracelulares de ARNg ni de Gag (Figuras 23 y 24).

El hecho de no haber observado un efecto de la sobreexpresión de HAR1A sobre los niveles Gag y ARNg intracelulares no significa necesariamente que este ARN no afecte alguna etapa del ciclo replicativo de VIH-1. Por un lado, es posible que basten unas pocas moléculas de HAR1A para que éste cumpla su función, por lo tanto, aumentar la cantidad mediante la sobreexpresión no generaría ningún efecto. En ese sentido, el silenciamiento es una mejor estrategia para evaluar el efecto que podría tener algún factor celular. En diversos trabajos, las estrategias de sobreexpresión y silenciamiento de ARNnc largos son utilizadas para evaluar su

función, sin embargo, sólo se han diseñado siRNAs para unos pocos ARNnc largos. Su diseño y optimización no es trivial, en particular para ARNnc pues son moléculas que ejercen su función como ARN y no como proteínas. Al momento de realizar esta tesis, no se habían descrito siRNAs para ninguno de los 4 ARNnc largos seleccionados. Sin embargo, recientemente se publicaron siRNAs que han mostrado silenciar exitosamente a HAR1A (Lee et al., 2021), por lo que sería una nueva alternativa para seguir ahondando en el efecto de este ARN.

Por otro lado, es posible que el efecto de HAR1A ocurra en alguna etapa del ciclo replicativo del virus que no haya sido evaluada en esta tesis. Como se mencionó al inicio de este trabajo, los ARNnc largos son capaces de regular la expresión génica a nivel transcripcional y postranscripcional, y existen diversos ejemplos que inhiben o activan la transcripción, *splicing*, traducción, degradación o la localización subcelular, mediante la interacción con ADN, ARN y/o proteínas. Un fenómeno que no se evaluó en esta tesis fue el empaquetamiento del ARNg y la formación de partículas virales. Puede que HAR1A afecte los niveles de ARNg empaquetado en las partículas virales liberadas, para lo cual sería necesario medir los niveles de ARN en el sobrenadante celular y normalizarlo por la cantidad de Gag. Este fenómeno se ha observado cuando se altera el estado de metilación del ARNg, sobreexpresando o silenciando las proteínas que lo metilan y demetilan (Pereira Montecinos, 2020). En ese sentido, una posibilidad es que HAR1A participe en el reclutamiento de esta maquinaria. Ya se sabe que un mecanismo común de los ARNnc largos es la regulación del estado de metilación del ADN. Además, se han reportado casos de ARNnc largos que regulan el estado de metilación de ARNs mediante la interacción con la demetilasa ALKBH5 (X. Wang et al., 2019) y con la metilasa METTL3 (Zhu et al., 2019). Por lo tanto, sería interesante identificar ARNnc largos que interactúen con la maquinaria de metilación y desmetilación en un

contexto de infección con VIH-1. Otro fenómeno que podría alterar el empaquetamiento del ARNg es una alteración en su localización subcelular. Esto ya se ha observado en VIH-1, donde el ARNnc largo NEAT1 secuestra a los ARNs virales con intrones en regiones subnucleares denominadas *paraspeckles*, impidiendo su exportación hacia el citoplasma (Zhang et al., 2013).

En esta tesis tampoco se evaluó la infectividad de las partículas virales al sobreexpresar HAR1A. Es posible que HAR1A altere la presencia de algún factor relevante para la infectividad de las nuevas partículas virales, promoviendo o inhibiendo la incorporación de proteínas virales y/o celulares. Este fenómeno también ha sido observado en VIH-1. Se ha reportado que 7SL, un ARNnc que es incorporado en las partículas virales, recluta al factor de restricción APOBEC3G (Zhang et al., 2010).

Como se puede apreciar, en este trabajo se desarrolló e implementó una estrategia de captura del ARNg de VIH-1 que mostró un alto nivel de enriquecimiento y, además, se identificaron de 4 ARNnc largos intergénicos que forman parte de los mRNPs del ARNg de VIH-1. El efecto sobre la replicación de VIH-1 que estos puedan tener y los mecanismos mediante los que ejerzan su función son aspectos interesantes para continuar indagando en futuras investigaciones.

6 CONCLUSIONES

- La estrategia de purificación de ARN con sondas antisentido 50% biotiniladas preparadas *in vitro* permite la captura y enriquecimiento de la región Gag-Pol del ARNg de VIH-1 de manera específica, lo cual fue evidenciado por RT-qPCR y también mediante análisis bioinformático de los resultados de RNA-seq.
- Las células Jurkat infectadas con pseudotipos virales de VIH-1 son un buen modelo para la captura del ARNg de VIH-1.
- Existen ARNnc largos asociados al ARNg de VIH-1, incluyendo ARNnc largos intergénicos. De estos, se identificaron cuatro ARNnc largos intergénicos con un enriquecimiento significativo, entre los que se encuentra HAR1A.
- HAR1A es un ARNnc largo que presenta una baja expresión en diversas líneas celulares relativo a su expresión en células Jurkat. Al sobreexpresar HAR1A no se observó un efecto sobre los niveles de la proteína viral Gag y tampoco sobre el ARNg de VIH-1.
- A pesar de lo anterior, no se descarta que exista un efecto de HAR1A sobre la replicación de VIH-1, el cual podría evidenciarse en otras condiciones experimentales, ya sea silenciando la expresión de HAR1A en células Jurkat o evaluando su efecto en otras etapas del ciclo replicativo.

7 REFERENCIAS

- Aillaud, M., & Schulte, L. N. (2020). Emerging roles of long noncoding rnas in the cytoplasmic milieu. *Non-Coding RNA*, 6(4), 1–14. <https://doi.org/10.3390/ncrna6040044>
- Ajamian, L., Abel, K., Rao, S., Vyboh, K., García-de-Gracia, F., Soto-Rifo, R., Kulozik, A. E., Gehring, N. H., & Mouland, A. J. (2015). HIV-1 recruits UPF1 but excludes UPF2 to promote nucleocytoplasmic export of the genomic RNA. *Biomolecules*, 5(4), 2808–2839. <https://doi.org/10.3390/biom5042808>
- Ajamian, L., Abrahamyan, L., Milev, M., Ivanov, P. V., Kulozik, A. E., Gehring, N. H., & Mouland, A. J. (2008). Unexpected roles for UPF1 in HIV-1 RNA metabolism and translation. *Rna*, 14(5), 914–927. <https://doi.org/10.1261/rna.829208>
- Askjaer, P., Jensen, T. H., Nilsson, J., Englmeier, L., & Kjems, J. (1998). The specificity of the CRM1-Rev nuclear export signal interaction is mediated by RanGTP. *Journal of Biological Chemistry*, 273(50), 33414–33422. <https://doi.org/10.1074/jbc.273.50.33414>
- Balasubramaniam, M., Pandhare, J., & Dash, C. (2018). Are microRNAs important players in HIV-1 infection? An update. *Viruses*, 10(3), 1–35. <https://doi.org/10.3390/v10030110>
- Baltz, A. G., Munschauer, M., Schwanhäusser, B., Vasile, A., Murakawa, Y., Schueler, M., Youngs, N., Penfold-Brown, D., Drew, K., Milek, M., Wyler, E., Bonneau, R., Selbach, M., Dieterich, C., & Landthaler, M. (2012). The mRNA-Bound Proteome and Its Global Occupancy Profile on Protein-Coding Transcripts. *Molecular Cell*, 46(5), 674–690. <https://doi.org/10.1016/j.molcel.2012.05.021>
- Bao, X., Guo, X., Yin, M., Tariq, M., Lai, Y., Kanwal, S., Zhou, J., Li, N., Lv, Y., Pulido-Quetglas, C., Wang, X., Ji, L., Khan, M. J., Zhu, X., Luo, Z., Shao, C., Lim, D. H., Liu, X., Li, N., ... Esteban, M. A. (2018). Capturing the interactome of newly transcribed RNA. *Nature Methods*, 15(3), 213–220. <https://doi.org/10.1038/nmeth.4595>
- Barat, C., Lullien, V., Schatz, O., Keith, G., Nugeyre, M. T., Gruninger-Leitch, F., Barre-Sinoussi, F., LeGrice, S. F. J., & Darlix, J. L. (1989). HIV-1 reverse transcriptase specifically interacts with the anticodon domain of its cognate primer tRNA. *EMBO Journal*, 8(11), 3279–3285. <https://doi.org/10.1002/j.1460-2075.1989.tb08488.x>
- Barichievy, S., Naidoo, J., Boullé, M., Scholefield, J., Parihar, S. P., Coussens, A. K., Brombacher, F., Sigal, A., & Mhlanga, M. M. (2018). Viral apoptosis evasion via the MAPK pathway by use of a host long noncoding RNA. *Frontiers in Cellular and Infection Microbiology*, 8(AUG), 1–16. <https://doi.org/10.3389/fcimb.2018.00263>
- Barichievy, S., Naidoo, J., & Mhlanga, M. M. (2015). Non-coding RNAs and HIV: Viral manipulation of host dark matter to shape the cellular environment. *Frontiers in Genetics*, 6(MAR), 1–11. <https://doi.org/10.3389/fgene.2015.00108>
- Barre-Sinoussi, F., Chermann, J., Rey, F., Nugeyre, M., Chamaret, S., Gruest, J., Dautet, C., Axler-Blin, C., Vezinet-Brun, F., Rouzioux, C., Rozenbaum, W., & Montagnier, L. (1983). Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science*, 220(4599), 868–871. <https://doi.org/10.1126/science.6189183>
- Barré-Sinoussi, F., Ross, A. L., & Delfraissy, J. F. (2013). Past, present and future: 30 years of HIV research. *Nature Reviews Microbiology*, 11(12), 877–883. <https://doi.org/10.1038/nrmicro3132>
- Barrera, A., Olguín, V., Vera-Otarola, J., & López-Lastra, M. (2020). Cap-independent translation initiation of the unspliced RNA of retroviruses. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, 1863(9), 194583. <https://doi.org/10.1016/j.bbagr.2020.194583>

- Bartel, D. P. (2009). MicroRNA Target Recognition and Regulatory Functions. *Cell*, 136(2), 215–233. <https://doi.org/10.1016/j.cell.2009.01.002>.MicroRNA
- Bartolomei, M. S., Zemel, S., & Tilghman, S. M. (1991). Parental imprinting of the mouse H19 Gene. *Nature*, 351(6322), 153–155. <https://doi.org/10.32388/sp99am>
- Bell, N. M., & Lever, A. M. L. (2013). HIV Gag polyprotein: Processing and early viral particle assembly. *Trends in Microbiology*, 21(3), 136–144. <https://doi.org/10.1016/j.tim.2012.11.006>
- Boliar, S., Gludish, D. W., Jambo, K. C., Kamng'ona, R., Mvaya, L., Mwandumba, H. C., & Russell, D. G. (2019). Inhibition of the lncRNA SAF drives activation of apoptotic effector caspases in HIV-1–infected human macrophages. *Proceedings of the National Academy of Sciences of the United States of America*, 116(15), 7431–7438. <https://doi.org/10.1073/pnas.1818662116>
- Bolinger, C., Sharma, A., Singh, D., Yu, L., & Boris-Lawrie, K. (2010). RNA helicase A modulates translation of HIV-1 and infectivity of progeny virions. *Nucleic Acids Research*, 38(5), 1686–1696. <https://doi.org/10.1093/nar/gkp1075>
- Bolinger, C., Yilmaz, A., Hartman, T. R., Kovacic, M. B., Fernandez, S., Ye, J., Forget, M., Green, P. L., & Boris-Lawrie, K. (2007). RNA helicase A interacts with divergent lymphotropic retroviruses and promotes translation of human T-cell leukemia virus type 1. *Nucleic Acids Research*, 35(8), 2629–2642. <https://doi.org/10.1093/nar/gkm124>
- Brasey, A., Lopez-Lastra, M., Ohlmann, T., Beerens, N., Berkhout, B., Darlix, J.-L., & Sonenberg, N. (2003). The Leader of Human Immunodeficiency Virus Type 1 Genomic RNA Harbors an Internal Ribosome Entry Segment That Is Active during the G₂/M Phase of the Cell Cycle. *Journal of Virology*, 77(7), 3939–3949. <https://doi.org/10.1128/jvi.77.7.3939-3949.2003>
- Briggs, J. A. G., Simon, M. N., Gross, I., Kräusslich, H. G., Fuller, S. D., Vogt, V. M., & Johnson, M. C. (2004). The stoichiometry of Gag protein in HIV-1. *Nature Structural and Molecular Biology*, 11(7), 672–675. <https://doi.org/10.1038/nsmb785>
- Briggs, J. A. G., Wilk, T., Welker, R., Kräusslich, H. G., & Fuller, S. D. (2003). Structural organization of authentic, mature HIV-1 virions and cores. *EMBO Journal*, 22(7), 1707–1715. <https://doi.org/10.1093/emboj/cdg143>
- Brockdorff, N., Ashworth, A., Kay, G. F., McCabe, V. M., Norris, D. P., Cooper, P. J., Swift, S., & Rastan, S. (1992). The product of the mouse Xist gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus. *Cell*, 71(3), 515–526. [https://doi.org/10.1016/0092-8674\(92\)90519-1](https://doi.org/10.1016/0092-8674(92)90519-1)
- Brown C, Ballabio A, Rupert JL, Lafreniere RG, Grompe M, Tonlorenzi R, & Willard HF. (1991). A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. *Nature*, 349(January), 38–44.
- Bryant, M., & Ratner, L. (1990). Myristoylation-dependent replication and assembly of human immunodeficiency virus 1. *Proceedings of the National Academy of Sciences of the United States of America*, 87(2), 523–527. <https://doi.org/10.1073/pnas.87.2.523>
- Buck, C. B., Shen, X., Egan, M. A., Pierson, T. C., Walker, C. M., & Siliciano, R. F. (2001). The Human Immunodeficiency Virus Type 1 gag Gene Encodes an Internal Ribosome Entry Site. *Journal of Virology*, 75(1), 181–191. <https://doi.org/10.1128/jvi.75.1.181-191.2001>
- Burdick, R. C., Li, C., Munshi, M. H., Rawson, J. M. O., Nagashima, K., Hu, W. S., & Pathak, V. K. (2020). HIV-1 uncoats in the nucleus near sites of integration. *Proceedings of the National Academy of Sciences of the United States of America*, 117(10), 5486–5493. <https://doi.org/10.1073/pnas.1920631117>

- Cabili, M., Trapnell, C., Goff, L., Koziol, M., Tazon-Vega, B., Regev, A., & Rinn, J. L. (2011). Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes and Development*, 25(18), 1915–1927. <https://doi.org/10.1101/gad.17446611>
- Caputi, M., Mayeda, A., Krainer, A. R., & Zahler, A. M. (1999). hnRNP A/B proteins are required for inhibition of HIV-1 pre-mRNA splicing. *EMBO Journal*, 18(14), 4060–4067. <https://doi.org/10.1093/emboj/18.14.4060>
- Carrieri, C., Cimatti, L., Biagioli, M., Beugnet, A., Zucchelli, S., Fedele, S., Pesce, E., Ferrer, I., Collavin, L., Santoro, C., Forrest, A. R. R., Carninci, P., Biffo, S., Stupka, E., & Gustincich, S. (2012). Long non-coding antisense RNA controls Uchl1 translation through an embedded SINEB2 repeat. *Nature*, 491(7424), 454–457. <https://doi.org/10.1038/nature11508>
- Carvajal, F., Vallejos, M., Walters, B., Contreras, N., Hertz, M. I., Olivares, E., Cáceres, C. J., Pino, K., Letelier, A., Thompson, S. R., & López-Lastra, M. (2016). Structural domains within the HIV-1 mRNA and the ribosomal protein S25 influence cap-independent translation initiation. *FEBS Journal*, 283, 2508–2527. <https://doi.org/10.1111/febs.13756>
- Castello, A., Fischer, B., Eichelbaum, K., Horos, R., Beckmann, B. M., Strein, C., Davey, N. E., Humphreys, D. T., Preiss, T., Steinmetz, L. M., Krijgsveld, J., & Hentze, M. W. (2012). Insights into RNA Biology from an Atlas of Mammalian mRNA-Binding Proteins. *Cell*, 149(6), 1393–1406. <https://doi.org/10.1016/j.cell.2012.04.031>
- Castello, A., Horos, R., Strein, C., Fischer, B., Eichelbaum, K., Steinmetz, L. M., Krijgsveld, J., & Hentze, M. W. (2013). System-wide identification of RNA-binding proteins by interactome capture. *Nature Protocols*, 8(3), 491–500. <https://doi.org/10.1038/nprot.2013.020>
- Chao, T. C., Zhang, Q., Li, Z., Tiwari, S. K., Qin, Y., Yau, E., Sanchez, A., Singh, G., Chang, K., Kaul, M., Karris, M. A. Y., & Ranaa, T. M. (2019). The long noncoding RNA HEAL regulates HIV-1 replication through epigenetic Regulation of the HIV-1 promoter. *MBio*, 10(5), 1–18. <https://doi.org/10.1128/mBio.02016-19>
- Chen, L., Chen, L., Zuo, L., Gao, Z., Shi, Y., Yuan, P., Han, S., Yin, J., Peng, B., He, X., & Liu, W. (2018). Short communication: Long noncoding RNA gas5 inhibits HIV-1 replication through interaction with miR-873. *AIDS Research and Human Retroviruses*, 34(6), 544–549. <https://doi.org/10.1089/aid.2017.0177>
- Chiu, I. M., Yaniv, A., Dahlberg, J. E., Gazit, A., Skuntz, S. F., Tronick, S. R., & Aaronson, S. A. (1985). Nucleotide sequence evidence for relationship of AIDS retrovirus to lentiviruses. *Nature*, 317(6035), 366–368. <https://doi.org/10.1038/317366a0>
- Choe, H., Farzan, M., Sun, Y., Sullivan, N., Rollins, B., Ponath, P. D., Wu, L., Mackay, C. R., LaRosa, G., Newman, W., Gerard, N., Gerard, C., & Sodroski, J. (1996). The β -chemokine receptors CCR3 and CCR5 facilitate infection by primary HIV-1 isolates. *Cell*, 85(7), 1135–1148. [https://doi.org/10.1016/S0092-8674\(00\)81313-6](https://doi.org/10.1016/S0092-8674(00)81313-6)
- Cimarelli, A., Sandin, S., Höglund, S., & Luban, J. (2000). Basic Residues in Human Immunodeficiency Virus Type 1 Nucleocapsid Promote Virion Assembly via Interaction with RNA. *Journal of Virology*, 74(7), 3046–3057. <https://doi.org/10.1128/jvi.74.7.3046-3057.2000>
- Clemson, C. M., Hutchinson, J. N., Sara, S. A., Ensminger, A. W., Fox, A. H., Chess, A., & Lawrence, J. B. (2009). An Architectural Role for a Nuclear Noncoding RNA: NEAT1 RNA Is Essential for the Structure of Paraspeckles. *Molecular Cell*, 33(6), 717–726. <https://doi.org/10.1016/j.molcel.2009.01.026>
- Coffin, J. M., Hughes, S. H., & Varmus, H. E. (1997). *Retroviral Virions and Genomes--Retroviruses*. Cold Spring Harbor Laboratory Press.

- Cohen, E. A., Dehni, G., Sodroski, J. G., & Haseltine, W. A. (1990). Human immunodeficiency virus vpr product is a virion-associated regulatory protein. *Journal of Virology*, *64*(6), 3097–3099. <https://doi.org/10.1128/jvi.64.6.3097-3099.1990>
- Collins, K. L., Chen, B. K., Kalams, S. A., Walker, B. D., & Baltimore, D. (1998). HIV-1 Nef protein protects infected primary cells against killing by cytotoxic T lymphocytes. *Nature*, *391*(22), 397–402. <https://doi.org/10.1038/255243a0>
- Cullen, B. R. (1991). Human Immunodeficiency Virus as a Prototypic Complex Retrovirus. *Journal of Virology*, *65*(3), 1053–1056.
- Dharan, A., Bachmann, N., Talley, S., Zwickelmaier, V., & Campbell, E. M. (2020). Nuclear pore blockade reveals that HIV-1 completes reverse transcription and uncoating in the nucleus. *Nature Microbiology*, *5*(9), 1088–1095. <https://doi.org/10.1038/s41564-020-0735-8>
- Djebali, S., Davis, C. A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., Tanzer, A., Lagarde, J., Lin, W., Schlesinger, F., Xue, C., Marinov, G. K., Khatun, J., Williams, B. A., Zaleski, C., Rozowsky, J., Röder, M., Kokocinski, F., Abdelhamid, R. F., ... Gingeras, T. R. (2012). Landscape of transcription in human cells. *Nature*, *489*(7414), 101–108. <https://doi.org/10.1038/nature11233>
- Dugré-Brisson, S., Elvira, G., Boulay, K., Chatel-Chaix, L., Moulard, A. J., & DesGroseillers, L. (2005). Interaction of Staufen1 with the 5' end of mRNA facilitates translation of these RNAs. *Nucleic Acids Research*, *33*(15), 4797–4812. <https://doi.org/10.1093/nar/gki794>
- Eddy, S. R. (2001). Non-coding RNA genes and the modern RNA world. *Nature Reviews Genetics*, *2*(12), 919–929. <https://doi.org/10.1038/35103511>
- Faghihi, M. A., Modarresi, F., Khalil, A. M., Wood, D. E., Sahagan, B. G., Morgan, T. E., Finch, C. E., St. Laurent, G., Kenny, P. J., & Wahlestedt, C. (2008). Expression of a noncoding RNA is elevated in Alzheimer's disease and drives rapid feed-forward regulation of β -secretase. *Nature Medicine*, *14*(7), 723–730. <https://doi.org/10.1038/nm1784>
- Fang, J., Kubota, S., Yang, B., Zhou, N., Zhang, H., Godbout, R., & Pomerantz, R. J. (2004). A DEAD box protein facilitates HIV-1 replication as a cellular co-factor of Rev. *Virology*, *330*(2), 471–480. <https://doi.org/10.1016/j.virol.2004.09.039>
- Feng, Y., Broder, C. C., Kennedy, P. E., & Berger, E. A. (1996). HIV-1 entry cofactor: Functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. *Science*, *272*(5263), 872–877. <https://doi.org/10.1126/science.272.5263.872>
- Fernandes, J. C. R., Acuña, S. M., Aoki, J. I., Floeter-Winter, L. M., & Muxel, S. M. (2019). Long non-coding RNAs in the regulation of gene expression: Physiology and disease. *Non-Coding RNA*, *5*(1). <https://doi.org/10.3390/ncrna5010017>
- Franke, E. K., Yuan, H. E. H., & Luban, J. (1994). Specific incorporation of cyclophilin a into HIV-1 virions. *Nature*, *372*(6504), 359–362. <https://doi.org/10.1038/372359a0>
- Freed, E. O. (2015). HIV-1 assembly, release and maturation. *Nature Reviews Microbiology*, *13*(8), 484–496. <https://doi.org/10.1038/nrmicro3490>
- Fröhlich, A., Rojas-Araya, B., Pereira-Montecinos, C., Dellarossa, A., Toro-Ascuy, D., Prades-Pérez, Y., García-de-Gracia, F., Garcés-Alday, A., Rubilar, P. S., Valiente-Echeverría, F., Ohlmann, T., & Soto-Rifo, R. (2016). DEAD-box RNA helicase DDX3 connects CRM1-dependent nuclear export and translation of the HIV-1 unspliced mRNA through its N-terminal domain. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, *1859*(5), 719–730. <https://doi.org/10.1016/j.bbagr.2016.03.009>
- Gallo, R., Wong-Staal, F., Montagnier, L., Haseltine, W. A., & Yoshida, M. (1988). HIV/HTLV gene

- nomenclature. *Nature*, 333(6173), 504. <https://doi.org/10.1038/333504a0>
- García-de-Gracia, F., Gaete-Argel, A., Riquelme-Barríos, S., Pereira-Montecinos, C., Rojas-Araya, B., Aguilera, P., Oyarzún-Arrau, A., Rojas-Fuentes, C., Acevedo, M. L., Chnaiderman, J., Valiente-Echeverría, F., Toro-Ascuy, D., & Soto-Rifo, R. (2021). CBP80/20-dependent translation initiation factor (CTIF) inhibits HIV-1 Gag synthesis by targeting the function of the viral protein Rev. *RNA Biology*, 18(5), 745–758.
- Giorgi, C., & Moore, M. J. (2007). The nuclear nurture and cytoplasmic nature of localized mRNPs. *Seminars in Cell and Developmental Biology*, 18(2), 186–193. <https://doi.org/10.1016/j.semcdb.2007.01.002>
- Gong, C., & Maquat, L. E. (2011). LncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 39 UTRs via Alu elements. *Nature*, 470(7333), 284–290. <https://doi.org/10.1038/nature09701>
- Gottlieb, M. S., Schroff, R., Schanker, H. M., Weisman, J. D., Fan, P. T., Wolf, R. A., & Saxon, Andrew. (1981). Pneumocystis carinii Pneumonia and Mucosal Candidiasis in Previously Healthy Homosexual Men. *The New England Journal of Medicine*, 305(24), 1425–1431.
- Guedán, A., Caroe, E. R., Barr, G. C. R., & Bishop, K. N. (2021). The role of capsid in HIV-1 nuclear entry. *Viruses*, 13(8), 1–14. <https://doi.org/10.3390/v13081425>
- Guttman, M., & Rinn, J. L. (2012). Modular regulatory principles of large non-coding RNAs. *Nature*, 482(7385), 339–346. <https://doi.org/10.1038/nature10887>
- Harrow, J., Frankish, A., Gonzalez, J. M., Tapanari, E., Diekhans, M., Kokocinski, F., Aken, B. L., Barrell, D., Zadissa, A., Searle, S., Barnes, I., Bignell, A., Boychenko, V., Hunt, T., Kay, M., Mukherjee, G., Rajan, J., Despacio-Reyes, G., Saunders, G., ... Hubbard, T. J. (2012). GENCODE: The reference human genome annotation for the ENCODE project. *Genome Research*, 22(9), 1760–1774. <https://doi.org/10.1101/gr.135350.111>
- Hartford, C. C. R., & Lal, A. (2020). When Long Noncoding Becomes Protein Coding. *Molecular and Cellular Biology*, 40(6), e00528-19.
- He, N., Liu, M., Hsu, J., Xue, Y., Chou, S., Burlingame, A., Krogan, N. J., Alber, T., & Zhou, Q. (2010). HIV-1 Tat and Host AFF4 Recruit Two Transcription Elongation Factors into a Bifunctional Complex for Coordinated Activation of HIV-1 Transcription. *Molecular Cell*, 38(3), 428–438. <https://doi.org/10.1016/j.molcel.2010.04.013>
- Hentze, M. W., Castello, A., Schwarzl, T., & Preiss, T. (2018). A brave new world of RNA-binding proteins. *Nature Reviews Molecular Cell Biology*, 19(5), 327–341. <https://doi.org/10.1038/nrm.2017.130>
- Hoffman, E. A., Frey, B. L., Smith, L. M., & Auble, D. T. (2015). Formaldehyde crosslinking: A tool for the study of chromatin complexes. *Journal of Biological Chemistry*, 290(44), 26404–26411. <https://doi.org/10.1074/jbc.R115.651679>
- Hu, G., Lou, Z., & Gupta, M. (2014). The long non-coding RNA GAS5 cooperates with the eukaryotic translation initiation factor 4E to regulate c-Myc translation. *PLoS ONE*, 9(9). <https://doi.org/10.1371/journal.pone.0107016>
- Hu, W., & Hughes, S. H. (2012). HIV-1 Reverse Transcription. *Cold Spring Harbor Perspectives in Medicine*, 2(10), a006882.
- Huan, C., Li, Z., Ning, S., Wang, H., Yu, X.-F., & Zhang, W. (2018). Long Noncoding RNA uc002yug.2 Activates HIV-1 Latency through Regulation of mRNA Levels of Various RUNX1 Isoforms and Increased Tat Expression. *Journal of Virology*, 92(9), 1–18. <https://doi.org/10.1128/jvi.01844-17>
- Huang, F., Zhang, J., Zhang, Y., Geng, G., Liang, J., Li, Y., Chen, J., Liu, C., & Zhang, H. (2015). RNA

- helicase MOV10 functions as a co-factor of HIV-1 Rev to facilitate Rev/RRE-dependent nuclear export of viral mRNAs. *Virology*, 486, 15–26. <https://doi.org/10.1016/j.virol.2015.08.026>
- Huang, Y., Mak, J., Cao, Q., Li, Z., Wainberg, M. A., & Kleiman, L. (1994). Incorporation of excess wild-type and mutant tRNA(3Lys) into human immunodeficiency virus type 1. *Journal of Virology*, 68(12), 7676–7683. <https://doi.org/10.1128/jvi.68.12.7676-7683.1994>
- Imam, H., Bano, A. S., Patel, P., Holla, P., & Jameel, S. (2015). The lncRNA NRON modulates HIV-1 replication in a NFAT-dependent manner and is differentially regulated by early and late viral proteins. *Scientific Reports*, 5, 1–10. <https://doi.org/10.1038/srep08639>
- Isel, C., & Karn, J. (1999). Direct evidence that HIV-1 Tat stimulates RNA polymerase II carboxyl-terminal domain hyperphosphorylation during transcriptional elongation. *Journal of Molecular Biology*, 290(5), 929–941. <https://doi.org/10.1006/jmbi.1999.2933>
- Ishaq, M., Marshall, H., & Natarajan, V. (2020). GADD34 attenuates HIV-1 replication by viral 5'-UTR TAR RNA-mediated translational inhibition. *Virology*, 540, 119–131. <https://doi.org/10.1016/j.virol.2019.11.010.GADD34>
- Ivanov, D., Kwak, Y. T., Guo, J., & Gaynor, R. B. (2000). Domains in the SPT5 Protein That Modulate Its Transcriptional Regulatory Properties. *Molecular and Cellular Biology*, 20(9), 2970–2983. <https://doi.org/10.1128/mcb.20.9.2970-2983.2000>
- Jacks, T., Powert, M. D., Masiarzt, F. R., Luciw, P. A., Barrt, P. J., & Varmus, H. E. (1987). Characterization of ribosomal frameshifting in HIV-1 gag-pol expression. *Nature*, 331(21), 280–283.
- Jacquet, S., Méreau, A., Bilodeau, P. S., Damier, L., Stoltzfus, C. M., & Branlant, C. (2001). A Second Exon Splicing Silencer within Human Immunodeficiency Virus Type 1 tat Exon 2 Represses Splicing of Tat mRNA and Binds Protein hnRNP H. *Journal of Biological Chemistry*, 276(44), 40464–40475. <https://doi.org/10.1074/jbc.M104070200>
- Jordan, A., Bisgrove, D., & Verdin, E. (2003). HIV reproducibly establishes a latent infection after acute infection of T cells in vitro. *EMBO Journal*, 22(8), 1868–1877. <https://doi.org/10.1093/emboj/cdg188>
- Jowett, J. B., Planelles, V., Poon, B., Shah, N. P., Chen, M. L., & Chen, I. S. (1995). The human immunodeficiency virus type 1 vpr gene arrests infected T cells in the G2 + M phase of the cell cycle. *Journal of Virology*, 69(10), 6304–6313. <https://doi.org/10.1128/jvi.69.10.6304-6313.1995>
- Kao, S. Y., Calman, A. F., Luciw, P. A., & Peterlin, B. M. (1987). Anti-termination of transcription within the long terminal repeat of HIV-1 by tat gene product. *Nature*, 330(6147), 489–493.
- Kapranov, P., Cheng, J., Dike, S., Nix, D. A., Duttagupta, R., Willingham, A. T., Stadler, P. F., Hertel, J., Hackermüller, J., Hofacker, I. L., Bell, I., Cheung, E., Drenkow, J., Dumais, E., Patel, S., Helt, G., Ganesh, M., Ghosh, S., Piccolboni, A., ... Gingeras, T. R. (2007). RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science*, 316(5830), 1484–1488. <https://doi.org/10.1126/science.1138341>
- Karn, J. (2011). The molecular biology of HIV latency: Breaking and restoring the Tat-dependent transcriptional circuit. *Curr Opin HIV AIDS*, 6(1), 4–11. <https://doi.org/10.1097/COH.0b013e328340ffbb>
- Karn, J., & Stoltzfus, C. M. (2012). Transcriptional and Posttranscriptional Regulation of HIV-1 Gene Expression. *Cold Spring Harbor Perspectives in Medicine*, 2(2), a006916. <https://doi.org/10.1016/B978-012095440-7/50028-7>
- Kazimierczyk, M., Kasproicz, M. K., Kasprzyk, M. E., & Wrzesinski, J. (2020). Human long noncoding RNA interactome: Detection, characterization and function. *International Journal of Molecular Sciences*, 21(3). <https://doi.org/10.3390/ijms21031027>

- Kennedy, E. M., Bogerd, H. P., Kornepati, A. V. R., Kang, D., Ghoshal, D., Marshall, J. B., Poling, B. C., Tsai, K., Gokhale, N. S., Horner, S. M., & Cullen, B. R. (2016). Posttranscriptional m6A Editing of HIV-1 mRNAs Enhances Viral Gene Expression. *Cell Host and Microbe*, *19*(5), 675–685. <https://doi.org/10.1016/j.chom.2016.04.002>
- Khong, A., & Parker, R. (2020). The landscape of eukaryotic mRNPs. *Rna*, *26*(3), 229–239. <https://doi.org/10.1261/rna.073601.119>
- Kleiman, L., Jones, C. P., & Musier-Forsyth, K. (2010). Formation of the tRNALys packaging complex in HIV-1. *FEBS Letters*, *584*(2), 359–365. <https://doi.org/10.1016/j.febslet.2009.11.038>
- Knoener, R. A., Becker, J. T., Scalf, M., Sherer, N. M., & Smith, L. M. (2017). Elucidating the in vivo interactome of HIV-1 RNA by hybridization capture and mass spectrometry. *Scientific Reports*, *7*(1), 1–16. <https://doi.org/10.1038/s41598-017-16793-5>
- Knoener, R., Evans, E., Becker, J. T., Scalf, M., Benner, B., Sherer, N. M., & Smith, L. M. (2021). Identification of host proteins differentially associated with hiv-1 rna splice variants. *ELife*, *10*, 1–32. <https://doi.org/10.7554/eLife.62470>
- Kretz, M., Sipsravili, Z., Chu, C., Webster, D. E., Zehnder, A., Qu, K., Lee, C. S., Flockhart, R. J., Groff, A. F., Chow, J., Johnston, D., Kim, G. E., Spitale, R. C., Flynn, R. A., Zheng, G. X. Y., Aiyer, S., Raj, A., Rinn, J. L., Chang, H. Y., & Khavari, P. A. (2013). Control of somatic tissue differentiation by the long non-coding RNA TINCR. *Nature*, *493*(7431), 231–235. <https://doi.org/10.1038/nature11661>
- Kula, A., Guerra, J., Knezevich, A., Kleva, D., Myers, M. P., & Marcello, A. (2011). Characterization of the HIV-1 RNA associated proteome identifies MatrIn 3 as a nuclear cofactor of Rev function. *Retrovirology*, *8*, 1–15. <https://doi.org/10.1186/1742-4690-8-60>
- Kulkarni, S., Lied, A., Kulkarni, V., Rucevic, M., Martin, M. P., Walker-Sperling, V., Anderson, S. K., Ewy, R., Singh, S., Nguyen, H., McLaren, P. J., Viard, M., Naranbhai, V., Zou, C., Lin, Z., Gatanaga, H., Oka, S., Takiguchi, M., Thio, C. L., ... Carrington, M. (2019). CCR5AS lncRNA variation differentially regulates CCR5, influencing HIV disease outcome. *Nature Immunology*, *20*(7), 824–834. <https://doi.org/10.1038/s41590-019-0406-1>
- Kutluay, S. B., & Bieniasz, P. D. (2010). Analysis of the initiating events in HIV-1 particle assembly and genome packaging. *PLoS Pathogens*, *6*(11). <https://doi.org/10.1371/journal.ppat.1001200>
- Kuzembayeva, M., Dilley, K., Sardo, L., & Hu, W. S. (2014). Life of psi: How full-length HIV-1 RNAs become packaged genomes in the viral particles. *Virology*, *454–455*(1), 362–370. <https://doi.org/10.1016/j.virol.2014.01.019>
- Lee, C. P., Ko, A. M. S., Nithyanantham, S., Lai, C. H., & Ko, Y. C. (2021). Long noncoding RNA HAR1A regulates oral cancer progression through the alpha-kinase 1, bromodomain 7, and myosin IIA axis. *Journal of Molecular Medicine*, *99*(9), 1323–1334. <https://doi.org/10.1007/s00109-021-02095-x>
- Lee, J. T., Davidow, L. S., & Warshawsky, D. (1999). Tsix, a gene antisense to Xist at the X-inactivation centre. *Nature Genetics*, *21*(4), 400–404. <https://doi.org/10.1038/7734>
- Lewis, P. F., & Emerman, M. (1994). Passage through mitosis is required for oncoretroviruses but not for the human immunodeficiency virus. *Journal of Virology*, *68*(1), 510–516. <https://doi.org/10.1128/jvi.68.1.510-516.1994>
- Li, C., Burdick, R. C., Nagashima, K., Hu, W. S., & Pathak, V. K. (2021). HIV-1 cores retain their integrity until minutes before uncoating in the nucleus. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(10). <https://doi.org/10.1073/pnas.2019467118>
- Li, H., Chi, X., Li, R., Ouyang, J., & Chen, Y. (2020). A Novel lncRNA, AK130181, Contributes to HIV-1 Latency by Regulating Viral Promoter-Driven Gene Expression in Primary CD4+ T Cells. *Molecular*

Therapy - Nucleic Acids, 20(June), 754–763. <https://doi.org/10.1016/j.omtn.2020.04.011>

- Li, J., Chen, C., Ma, X., Geng, G., Liu, B., Zhang, Y., Zhang, S., Zhong, F., Liu, C., Yin, Y., Cai, W., & Zhang, H. (2016). Long noncoding RNA NRON contributes to HIV-1 latency by specifically inducing tat protein degradation. *Nature Communications*, 7. <https://doi.org/10.1038/ncomms11730>
- Li, S., Hill, C. P., Sundquist, W. I., & Finch, J. T. (2000). Image reconstructions of helical assemblies of the HIV-1 CA protein. *Nature*, 407(6802), 409–413. <https://doi.org/10.1038/35030177>
- Lichinchi, G., Gao, S., Saletore, Y., Gonzalez, G. M., Bansal, V., Wang, Y., Mason, C. E., & Rana, T. M. (2016). Dynamics of the human and viral m(6)A RNA methylomes during HIV-1 infection of T cells. *Nature Microbiology*, 1(4), 16011. <https://doi.org/10.1038/nmicrobiol.2016.11>
- Lightfoote, M. M., Coligan, J. E., Folks, T. M., Fauci, A. S., Martin, M. A., & Venkatesan, S. (1986). Structural characterization of reverse transcriptase and endonuclease polypeptides of the acquired immunodeficiency syndrome retrovirus. *Journal of Virology*, 60(2), 771–775. <https://doi.org/10.1128/jvi.60.2.771-775.1986>
- Liu, H., Wu, X., Newman, M., Shaw, G. M., Hahn, B. H., & Kappes, J. C. (1995). The Vif protein of human and simian immunodeficiency viruses is packaged into virions and associates with viral core structures. *Journal of Virology*, 69(12), 7630–7638. <https://doi.org/10.1128/jvi.69.12.7630-7638.1995>
- Liu, J., Henao-Mejia, J., Liu, H., Zhao, Y., & He, J. J. (2011). Translational regulation of HIV-1 replication by HIV-1 rev cellular cofactors Sam68, eIF5A, hRIP, and DDX3. *Journal of Neuroimmune Pharmacology*, 6(2), 308–321. <https://doi.org/10.1007/s11481-011-9265-8>
- Locati, M. D., Terpstra, I., De Leeuw, W. C., Kuzak, M., Rauwerda, H., Ensink, W. A., Van Leeuwen, S., Nehrlich, U., Spaik, H. P., Jonker, M. J., Breit, T. M., & Dekker, R. J. (2015). Improving small RNA-seq by using a synthetic spike-in set for size-range quality control together with a set for data normalization. *Nucleic Acids Research*, 43(14), e89. <https://doi.org/10.1093/nar/gkv303>
- Loda, A., & Heard, E. (2019). Xist RNA in action: Past, present, and future. *PLoS Genetics*, 15(9), 1–17. <https://doi.org/10.1371/journal.pgen.1008333>
- Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, 15(12), 1–21. <https://doi.org/10.1186/s13059-014-0550-8>
- Lutzmayer, S., Enugutti, B., & Nodine, M. D. (2017). Novel small RNA spike-in oligonucleotides enable absolute normalization of small RNA-Seq data. *Scientific Reports*, 7(1), 1–6. <https://doi.org/10.1038/s41598-017-06174-3>
- Ma, L., Cao, J., Liu, L., Du, Q., Li, Z., Zou, D., Bajic, V. B., & Zhang, Z. (2019). Lncbook: A curated knowledgebase of human long non-coding rnas. *Nucleic Acids Research*, 47(D1), D128–D134. <https://doi.org/10.1093/nar/gky960>
- Maddon, P. J., Dalgleish, A. G., McDougal, J. S., Clapham, P. R., Weiss, R. A., & Axel, R. (1986). The T4 gene encodes the AIDS virus receptor and is expressed in the immune system and the brain. *Cell*, 47(3), 333–348. [https://doi.org/10.1016/0092-8674\(86\)90590-8](https://doi.org/10.1016/0092-8674(86)90590-8)
- Mailler, E., Bernacchi, S., Marquet, R., Paillart, J. C., Vivet-Boudou, V., & Smyth, R. P. (2016). The life-cycle of the HIV-1 gag–RNA complex. *Viruses*, 8(9), 1–19. <https://doi.org/10.3390/v8090248>
- Malim, M. H., & Emerman, M. (2008). HIV-1 Accessory Proteins-Ensuring Viral Survival in a Hostile Environment. *Cell Host and Microbe*, 3(6), 388–398. <https://doi.org/10.1016/j.chom.2008.04.008>
- Malim, M. H., Hauber, J., Le, S. Y., Maizel, J. V., & Cullen, B. R. (1989). The HIV-1 rev trans-activator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA. *Nature*, 338(6212), 254–257. <https://doi.org/10.1038/338254a0>

- Marchand, V., Santerre, M., Aigueperse, C., Fouillen, L., Saliou, J.-M., Van Dorsselaer, A., Sanglier-Cianféron, S., Branlant, C., & Motorin, Y. (2011). Identification of protein partners of the human immunodeficiency virus 1 tat / rev exon 3 leads to the discovery of a new HIV-1 splicing regulator, protein hnRNP K. *RNA Biology*, 8(2), 325–342. <https://doi.org/10.4161/rna.8.2.13984>
- Martin Stoltzfus, C. (2009). Regulation of HIV-1 Alternative RNA Splicing and Its Role in Virus Replication. In *Advances in Virus Research* (1st ed., Vol. 74, Issue 09, pp. 1–40). Elsevier Inc. [https://doi.org/10.1016/S0065-3527\(09\)74001-1](https://doi.org/10.1016/S0065-3527(09)74001-1)
- Mateu, M. G. (2009). The capsid protein of human immunodeficiency virus: Intersubunit interactions during virus assembly. *FEBS Journal*, 276(21), 6098–6109. <https://doi.org/10.1111/j.1742-4658.2009.07313.x>
- Mayeda, A., Sreaton, G. R., Chandler, S. D., Fu, X.-D., & Krainer, A. R. (1999). Substrate Specificities of SR Proteins in Constitutive Splicing Are Determined by Their RNA Recognition Motifs and Composite Pre-mRNA Exonic Elements. *Molecular and Cellular Biology*, 19(3), 1853–1863. <https://doi.org/10.1128/mcb.19.3.1853>
- Mcdougal, J. S., Kennedy, M. S., Sligh, J. M., Cort, S. P., Mawle, A., & Nicholson, J. K. A. (1986). Binding of HTLV-III/LAV to T4+ T cells by a complex of the 110K viral protein and the T4 molecule. *Science*, 231(4736), 382–385. <https://doi.org/10.1126/science.3001934>
- McHugh, C. A., Chen, C. K., Chow, A., Surka, C. F., Tran, C., McDonel, P., Pandya-Jones, A., Blanco, M., Burghard, C., Moradian, A., Sweredoski, M. J., Shishkin, A. A., Su, J., Lander, E. S., Hess, S., Plath, K., & Guttman, M. (2015). The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature*, 521(7551), 232–236. <https://doi.org/10.1038/nature14443>
- Monette, A., Ajamian, L., López-Lastra, M., & Mouland, A. J. (2009). Human immunodeficiency virus type 1 (HIV-1) induces the cytoplasmic retention of heterogeneous nuclear ribonucleoprotein A1 by disrupting nuclear import. Implications for HIV-1 gene expression. *Journal of Biological Chemistry*, 284(45), 31350–31362. <https://doi.org/10.1074/jbc.M109.048736>
- Monette, A., Valiente-Echeverría, F., Rivero, M., Cohen, É. A., Lopez-Lastra, M., & Mouland, A. J. (2013). Dual Mechanisms of Translation Initiation of the Full-Length HIV-1 mRNA Contribute to Gag Synthesis. *PLoS ONE*, 8(7), 1–12. <https://doi.org/10.1371/journal.pone.0068108>
- Moore, M. J. (2005). From birth to death: The complex lives of eukaryotic mRNAs. *Science*, 309(5740), 1514–1518. <https://doi.org/10.1126/science.1111443>
- Moore, M. J., & Proudfoot, N. J. (2009). Pre-mRNA Processing Reaches Back to Transcription and Ahead to Translation. *Cell*, 136(4), 688–700. <https://doi.org/10.1016/j.cell.2009.02.001>
- Müller-Mcnicoll, M., & Neugebauer, K. M. (2013). How cells get the message: Dynamic assembly and function of mRNA-protein complexes. *Nature Reviews Genetics*, 14(4), 275–287. <https://doi.org/10.1038/nrg3434>
- Noh, J. H., Kim, K. M., McClusky, W. G., Abdelmohsen, K., & Gorospe, M. (2018). Cytoplasmic functions of long noncoding RNAs. *Wiley Interdisciplinary Reviews: RNA*, 9(3), 1–15. <https://doi.org/10.1002/wrna.1471>
- Ocwieja, K. E., Sherrill-Mix, S., Mukherjee, R., Custers-Allen, R., David, P., Brown, M., Wang, S., Link, D. R., Olson, J., Travers, K., Schadt, E., & Bushman, F. D. (2012). Dynamic regulation of HIV-1 mRNA populations analyzed by single-molecule enrichment and long-read sequencing. *Nucleic Acids Research*, 40(20), 10345–10355. <https://doi.org/10.1093/nar/gks753>
- Olden, B. R., Cheng, Y., Yu, J. L., & Pun, S. H. (2018). Cationic polymers for non-viral gene delivery to human T cells. *Journal of Controlled Release*, 282(March), 140–147. <https://doi.org/10.1016/j.jconrel.2018.02.043>

- Olguín Araneda, V. (2020). *La Ribonucleoproteína Heterogénea Nuclear K (hnRNP K) es un Factor que Actúa en Trans (ITAF) Sobre los Sitios Internos de Entrada al Ribosoma (IRESes) de HIV-1, HTLV-1 y MMTV*. Pontificia Universidad Católica de Chile.
- Ono, A. (2010). Relationships between plasma membrane microdomains and HIV-1 assembly. *Biology of the Cell*, 102(6), 335–350. <https://doi.org/10.1042/bc20090165>
- Ono, A., Orenstein, J. M., & Freed, E. O. (2000). Role of the Gag Matrix Domain in Targeting Human Immunodeficiency Virus Type 1 Assembly. *Journal of Virology*, 74(6), 2855–2866. <https://doi.org/10.1128/jvi.74.6.2855-2866.2000>
- Paillart, J. C., Marquet, R., Skripkin, E., Ehresmann, C., & Ehresmann, B. (1996). Dimerization of retroviral genomic RNAs: Structural and functional implications. *Biochimie*, 78(7), 639–653. [https://doi.org/10.1016/S0300-9084\(96\)80010-1](https://doi.org/10.1016/S0300-9084(96)80010-1)
- Pashev, I. G., Dimitrov, S. I., & Angelov, D. (1991). Crosslinking proteins to nucleic acids by ultraviolet laser irradiation. *Trends in Biochemical Sciences*, 16(C), 323–326. [https://doi.org/10.1016/0968-0004\(91\)90133-G](https://doi.org/10.1016/0968-0004(91)90133-G)
- Pereira Montecinos, C. A. (2020). *Impacto de la Metilación de Adenosinas en el Destino Citoplásmico del ARN Genómico de VIH-1*. Universidad de Chile.
- Pollard, V. W., & Malim, M. H. (1998). The HIV-1 Rev Protein. *Medicine*, 52, 491–532. <http://www.ncbi.nlm.nih.gov/pubmed/9891806>
- Popov, S., Rexach, M., Zybarth, G., Railing, N., Lee, M. A., Ratner, L., Lane, C. M., Moore, M. S., Blobel, G., & Bukrinsky, M. (1998). Viral protein R regulates nuclear import of the HIV-1 pre-integration complex. *EMBO Journal*, 17(4), 909–917. <https://doi.org/10.1093/emboj/17.4.909>
- Postler, T. S., Pantry, S. N., Desrosiers, R. C., & Ghosh, S. (2017). Identification and characterization of a long non-coding RNA up-regulated during HIV-1 infection. *Virology*, 511(May), 30–39. <https://doi.org/10.1016/j.virol.2017.08.006>
- Purcell, D. F., & Martin, M. A. (1993). Alternative splicing of human immunodeficiency virus type 1 mRNA modulates viral protein expression, replication, and infectivity. *Journal of Virology*, 67(11), 6365–6378. <https://doi.org/10.1128/jvi.67.11.6365-6378.1993>
- Qu, D., Sun, W. W., Li, L., Ma, L., Sun, L., Jin, X., Li, T., Hou, W., & Wang, J. H. (2019). Long noncoding RNA MALAT1 releases epigenetic silencing of HIV-1 replication by displacing the polycomb repressive complex 2 from binding to the LTR promoter. *Nucleic Acids Research*, 47(6), 3013–3027. <https://doi.org/10.1093/nar/gkz117>
- Ramdas, P., Sahu, A. K., Mishra, T., Bhardwaj, V., & Chande, A. (2020). From Entry to Egress: Strategic Exploitation of the Cellular Processes by HIV-1. *Frontiers in Microbiology*, 11(December), 1–18. <https://doi.org/10.3389/fmicb.2020.559792>
- Ramos, H., Monette, A., Niu, M., Barrera, A., Brenda, L., Guizar, P., Pino, K., Desgroseillers, L., Moulard, A. J., & López-Lastra, M. (2021). The double-stranded RNA-binding protein, Stauf1, is an IRES-transacting factor regulating HIV-1 cap-independent translation initiation. *Nucleic Acids Research*, 1–19. <https://doi.org/https://doi.org/10.1093/nar/gkab1188>
- Ransohoff, J. D., Wei, Y., & Khavari, P. A. (2018). The functions and unique features of long intergenic non-coding RNA. *Nature Reviews Molecular Cell Biology*, 19(3), 143–157. <https://doi.org/10.1038/nrm.2017.104>
- Ray, R. M., & Morris, K. V. (2020). Long non-coding RNAs mechanisms of action in HIV-1 modulation and the identification of novel therapeutic targets. *Non-Coding RNA*, 6(1), 1–17. <https://doi.org/10.3390/NCRNA6010012>

- Ricci, E. P., Mure, F., Gruffat, H., Decimo, D., Medina-Palazon, C., Ohlmann, T., & Manet, E. (2009). Translation of intronless RNAs is strongly stimulated by the Epstein-Barr virus mRNA export factor EB2. *Nucleic Acids Research*, *37*(15), 4932–4943. <https://doi.org/10.1093/nar/gkp497>
- Ricci, E. P., Rifo, R. S., Herbreteau, C. H., Decimo, D., & Ohlmann, T. (2008). Lentiviral RNAs can use different mechanisms for translation initiation. *Biochemical Society Transactions*, *36*(4), 690–693. <https://doi.org/10.1042/BST0360690>
- Rice, A. P., & Mathews, M. B. (1988). Transcriptional but not translational regulation of HIV-1 by the tat gene product. *Nature*, *332*(6164), 551–553. <https://doi.org/10.1038/332551a0>
- Rinn, J. L., & Chang, H. Y. (2012). Genome regulation by long noncoding RNAs. *Annual Review of Biochemistry*, *81*, 145–166. <https://doi.org/10.1146/annurev-biochem-051410-092902>
- Rinn, J. L., & Chang, H. Y. (2020). Long Noncoding RNAs: Molecular Modalities to Organismal Functions. *Annual Review of Biochemistry*, *89*, 283–308. <https://doi.org/10.1146/annurev-biochem-062917-012708>
- Rinn, J. L., Kertesz, M., Wang, J. K., Squazzo, S. L., Xu, X., Bruggmann, S. A., Goodnough, L. H., Helms, J. A., Farnham, P. J., Segal, E., & Chang, H. Y. (2007). Functional Demarcation of Active and Silent Chromatin Domains in Human HOX Loci by Noncoding RNAs. *Cell*, *129*(7), 1311–1323. <https://doi.org/10.1016/j.cell.2007.05.022>
- Riquelme-Barrios, S., Pereira-Montecinos, C., Valiente-Echeverría, F., & Soto-Rifo, R. (2018). Emerging roles of N6-methyladenosine on HIV-1 RNA metabolism and viral replication. *Frontiers in Microbiology*, *9*(MAR), 1–11. <https://doi.org/10.3389/fmicb.2018.00576>
- Rivas-Aravena, A., Ramdohr, P., Vallejos, M., Valiente-Echeverría, F., Dormoy-Raclet, V., Rodríguez, F., Pino, K., Holzmann, C., Huidobro-Toro, J. P., Gallouzi, I. E., & López-Lastra, M. (2009). The Elavl-like protein HuR exerts translational control of viral internal ribosome entry sites. *Virology*, *392*(2), 178–185. <https://doi.org/10.1016/j.virol.2009.06.050>
- Robey, W. G., Safai, B., Oroszlan, S., Arthur, L. O., Gonda, M. A., & Gallo, R. C. (1985). Characterization of Envelope and Core Structural Gene Products of HTVL-III with Sera from AIDS Patients. *Science*, *228*(4699), 593–595.
- Ropers, D., Ayadi, L., Gattoni, R., Jacquenet, S., Damier, L., Branlant, C., & Stévenin, J. (2004). Differential effects of the SR proteins 9G8, SC35, ASF/SF2, and SRp40 on the utilization of the A1 to A5 splicing sites of HIV-1 RNA. *Journal of Biological Chemistry*, *279*(29), 29963–29973. <https://doi.org/10.1074/jbc.M404452200>
- Rosen, C. A., Sodroski, J. G., & Haseltine, W. A. (1985). Location of cis-acting regulatory sequences in the human T-cell leukemia virus type I long terminal repeat. *Cell*, *41*(3), 813–823. <https://doi.org/10.1073/pnas.82.19.6502>
- Sánchez-Velaz, N., Udofia, E. B., Yu, Z., & Zapp, M. L. (2004). hRIP, a cellular cofactor for Rev function, promotes release of HIV RNAs from the perinuclear region. *Genes and Development*, *18*(1), 23–34. <https://doi.org/10.1101/gad.1149704>
- Schwartz, O., Marechal, V., Le Gall, S., Lemonnier, F., & Heard, J.-M. (1996). Endocytosis of major histocompatibility complex class I molecules is induced by the HIV-1 Nef protein. *Nature*, *2*(3), 338–342.
- Sharma, A., Yilmaz, A., Marsh, K., Cochrane, A., & Boris-Lawrie, K. (2012). Thriving under stress: Selective translation of HIV-1 structural protein mRNA during Vpr-mediated impairment of eIF4E translation activity. *PLoS Pathogens*, *8*(3). <https://doi.org/10.1371/journal.ppat.1002612>
- Sheehy, A. M., Gaddis, N. C., Choi, J. D., & Malim, M. H. (2002). Isolation of a human gene that inhibits

- HIV-1 infection and is suppressed by the viral Vif protein. *Nature*, 418(6898), 646–650. <https://doi.org/10.1038/nature00969>. Published
- Shi, Z., Luo, Y., Zhu, M., Zhou, Y., Zheng, B., Wu, D., Wang, S., Xie, X., Lin, H., & Yu, X. (2019). Expression analysis of long non-coding RNA HAR1A and HAR1B in HBV-induced hepatocellular carcinoma in Chinese patients. *Lab Medicine*, 50(2), 150–157. <https://doi.org/10.1093/labmed/lmy055>
- Singh, G., Pratt, G., Yeo, G. W., & Moore, M. J. (2015). The clothes make the mRNA: Past and present trends in mRNP fashion. *Annual Review of Biochemistry*, 84, 325–354. <https://doi.org/10.1146/annurev-biochem-080111-092106>
- Singh, G., Ricci, E. P., & Moore, M. J. (2014). RIPiT-Seq: A high-throughput approach for footprinting RNA: Protein complexes. *Methods*, 65(3), 320–332. <https://doi.org/10.1016/j.ymeth.2013.09.013>
- Skripkin, E., Paillart, J. C., Marquet, R., Ehresmann, B., & Ehresmann, C. (1994). Identification of the primary site of the human immunodeficiency virus type 1 RNA dimerization in vitro. *Proceedings of the National Academy of Sciences of the United States of America*, 91(11), 4945–4949. <https://doi.org/10.1073/pnas.91.11.4945>
- Sobhian, B., Laguette, N., Yatim, A., Nakamura, M., Levy, Y., Kiernan, R., & Benkirane, M. (2010). HIV-1 Tat Assembles a Multifunctional Transcription Elongation Complex and Stably Associates with the 7SK snRNP. *Molecular Cell*, 38(3), 439–451. <https://doi.org/10.1016/j.molcel.2010.04.012>
- Soto-Rifo, R., Rubilar, P. S., Limousin, T., De Breyne, S., Décimo, D., & Ohlmann, T. (2012). DEAD-box protein DDX3 associates with eIF4F to promote translation of selected mRNAs. *EMBO Journal*, 31(18), 3745–3756. <https://doi.org/10.1038/emboj.2012.220>
- Soto-Rifo, R., Rubilar, P. S., & Ohlmann, T. (2013). The DEAD-box helicase DDX3 substitutes for the cap-binding protein eIF4E to promote compartmentalized translation initiation of the HIV-1 genomic RNA. *Nucleic Acids Research*, 41(12), 6286–6299. <https://doi.org/10.1093/nar/gkt306>
- Statello, L., Guo, C. J., Chen, L. L., & Huarte, M. (2021). Gene regulation by long non-coding RNAs and its biological functions. *Nature Reviews Molecular Cell Biology*, 22(2), 96–118. <https://doi.org/10.1038/s41580-020-00315-9>
- Strebel, K., Klimkait, T., & Martin, M. A. (1988). A Novel Gene of HIV-1, vpu, and its 16-Kilodalton Product. *Science*, 241(4870), 1221–1223.
- Suchanek, M., Radzikowska, A., & Thiele, C. (2005). Photo-leucine and photo-methionine allow identification of protein-protein interactions in living cells. *Nature Methods*, 2(4), 261–267. <https://doi.org/10.1038/nmeth752>
- Sundquist, W. I., & Krausslich, H.-G. (2012). HIV-1 Assembly, Budding, and Maturation. *Cold Spring Harbor Perspectives in Medicine*, 2(7), a006924.
- Sunwoo, H., Dinger, M. E., Wilusz, J. E., Amaral, P. P., Mattick, J. S., & Spector, D. L. (2009). Men ϵ/β nuclear-retained non-coding RNAs are up-regulated upon muscle differentiation and are essential components of paraspeckles. *Genome Research*, 19(3), 347–359. <https://doi.org/10.1101/gr.087775.108>
- Svitkin, Y. V., Pause, A., & Sonenberg, N. (1994). La autoantigen alleviates translational repression by the 5' leader sequence of the human immunodeficiency virus type 1 mRNA. *Journal of Virology*, 68(11), 7001–7007. <https://doi.org/10.1128/jvi.68.11.7001-7007.1994>
- Tazi, J., Bakkour, N., Marchand, V., Ayadi, L., Aboufirassi, A., & Branlant, C. (2010). Alternative splicing: Regulation of HIV-1 multiplication as a target for therapeutic action: Minireview. *FEBS Journal*, 277(4), 867–876. <https://doi.org/10.1111/j.1742-4658.2009.07522.x>

- Tirumuru, N., Zhao, B. S., Lu, W., Lu, Z., He, C., & Wu, L. (2016). N6-methyladenosine of HIV-1 RNA regulates viral infection and HIV-1 Gag protein expression. *ELife*, *5*(2016JULY), 1–20. <https://doi.org/10.7554/eLife.15528>
- Toro-Ascuy, D., Rojas-Araya, B., García-de-Gracia, F., Rojas-Fuentes, C., Pereira-Montecinos, C., Gaete-Argel, A., Valiente-Echeverría, F., Ohlmann, T., & Soto-Rifo, R. (2018). A Rev–CBP80–eIF4A1 complex drives Gag synthesis from the HIV-1 unspliced mRNA. *Nucleic Acids Research*, *46*(21), 11539–11552. <https://doi.org/10.1093/nar/gky851>
- Tripathi, V., Ellis, J. D., Shen, Z., Song, D. Y., Pan, Q., Watt, A. T., Freier, S. M., Bennett, C. F., Sharma, A., Bubulya, P. A., Blencowe, B. J., Prasanth, S. G., & Prasanth, K. V. (2010). The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Molecular Cell*, *39*(6), 925–938. <https://doi.org/10.1016/j.molcel.2010.08.011>
- Trypsteen, W., Mohammadi, P., Van Hecke, C., Mestdagh, P., Lefever, S., Saeys, Y., De Bleser, P., Vandesomepele, J., Ciuffi, A., Vandekerckhove, L., & De Spiegelaere, W. (2016). Differential expression of lncRNAs during the HIV replication cycle: An underestimated layer in the HIV-host interplay. *Scientific Reports*, *6*(October), 1–12. <https://doi.org/10.1038/srep36111>
- Trypsteen, W., White, C. H., Mukim, A., Spina, C. A., de Spiegelaere, W., Lefever, S., Planelles, V., Bosque, A., Woelk, C. H., Vandekerckhove, L., & Beliakova-Bethell, N. (2019). Long non-coding RNAs and latent HIV – A search for novel targets for latency reversal. *PLoS ONE*, *14*(11), 1–25. <https://doi.org/10.1371/journal.pone.0224879>
- UNAIDS. (2021). Fact Sheet 2021 Global HIV Statistics. *Ending the AIDS Epidemic*, June, 1–3.
- Urdaneta, E. C., Vieira-Vieira, C. H., Hick, T., Wessels, H. H., Figini, D., Moschall, R., Medenbach, J., Ohler, U., Granneman, S., Selbach, M., & Beckmann, B. M. (2019). Purification of cross-linked RNA-protein complexes by phenol-toluol extraction. *Nature Communications*, *10*(1), 1–17. <https://doi.org/10.1038/s41467-019-08942-3>
- Veronese, F. D., Copeland, T. D., Devico, A. L., Rahman, R., Oroszlan, S., Gallo, R. C., & Sarngadhaan, M. G. (1986). Characterization of Highly Immunogenic p66/p51 as the Reverse Transcriptase of HTLV-III/LAV. *Science*, *231*(4743), 1289–1291.
- Veronese, F. D., DeVico, A. L., Copeland, T. D., Oroszlan, S., Gallo, R. C., & Sarngadharan, M. G. (1985). Characterization of gp41 As the Transmembrane Protein Coded by the HTLV-III/LAV Envelope Gene. *Science*, *229*(4720), 1402–1405.
- Wang, K. C., Yang, Y. W., Liu, B., Sanyal, A., Corces-Zimmerman, R., Chen, Y., Lajoie, B. R., Protacio, A., Flynn, R. A., Gupta, R. A., Wysocka, J., Lei, M., Dekker, J., Helms, J. A., & Chang, H. Y. (2011). A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature*, *472*(7341), 120–126. <https://doi.org/10.1038/nature09819>
- Wang, T., Tian, C., Zhang, W., Luo, K., Sarkis, P. T. N., Yu, L., Liu, B., Yu, Y., & Yu, X.-F. (2007). 7SL RNA Mediates Virion Packaging of the Antiviral Cytidine Deaminase APOBEC3G. *Journal of Virology*, *81*(23), 13112–13124. <https://doi.org/10.1128/jvi.00892-07>
- Wang, X., Zhang, J., & Wang, Y. (2019). Long noncoding RNA GAS5-AS1 suppresses growth and metastasis of cervical cancer by increasing GAS5 stability. *American Journal of Translational Research*, *11*(8), 4909–4921.
- Welker, R., Kottler, H., Kalbitzer, H. R., & Kräusslich, H. G. (1996). Human immunodeficiency virus type 1 Nef protein is incorporated into virus particles and specifically cleaved by the viral proteinase. *Virology*, *219*(1), 228–236. <https://doi.org/10.1006/viro.1996.0240>
- Wilén, C. B., Tilton, J. C., & Doms, R. W. (2012). HIV: Cell binding and entry. *Cold Spring Harbor Perspectives in Medicine*, *2*(8), 1–14. <https://doi.org/10.1101/cshperspect.a006866>

- Wilk, T., Gross, I., Gowen, B. E., Rutten, T., De Haas, F., Welker, R., Krausslich, H.-G., Boulanger, P., & Fuller, S. D. (2001). Organization of Immature Human Immunodeficiency Virus Type 1. *Journal of Virology*, 75(2), 759–771. <https://doi.org/10.1128/JVI.75.2.759>
- Wu, H., Yang, L., & Chen, L. L. (2017). The Diversity of Long Noncoding RNAs and Their Generation. *Trends in Genetics*, 33(8), 540–552. <https://doi.org/10.1016/j.tig.2017.05.004>
- Wu, S. R., Löving, R., Lindqvist, B., Hebert, H., Koeck, P. J. B., Sjöberg, M., & Garoff, H. (2010). Single-particle cryoelectron microscopy analysis reveals the HIV-1 spike as a tripod structure. *Proceedings of the National Academy of Sciences of the United States of America*, 107(44), 18844–18849. <https://doi.org/10.1073/pnas.1007227107>
- Yap, K. L., Li, S., Muñoz-Cabello, A. M., Raguz, S., Zeng, L., Mujtaba, S., Gil, J., Walsh, M. J., & Zhou, M. M. (2010). Molecular Interplay of the Noncoding RNA ANRIL and Methylated Histone H3 Lysine 27 by Polycomb CBX7 in Transcriptional Silencing of INK4a. *Molecular Cell*, 38(5), 662–674. <https://doi.org/10.1016/j.molcel.2010.03.021>
- Yasuda-Inoue, M., Kuroki, M., & Ariumi, Y. (2013). Distinct DDX DEAD-box RNA helicases cooperate to modulate the HIV-1 Rev function. *Biochemical and Biophysical Research Communications*, 434(4), 803–808. <https://doi.org/10.1016/j.bbrc.2013.04.016>
- Yedavalli, V. S. R. K., & Jeang, K. T. (2010). Trimethylguanosine capping selectively promotes expression of Rev-dependent HIV-1 RNAs. *Proceedings of the National Academy of Sciences of the United States of America*, 107(33), 14787–14792. <https://doi.org/10.1073/pnas.1009490107>
- Yoon, J. H., Abdelmohsen, K., & Gorospe, M. (2013). Post-transcriptional gene regulation by long noncoding RNA. *Journal of Molecular Biology*, 425(19), 3723–3730. <https://doi.org/10.1016/j.jmb.2012.11.024>. Post-transcriptional
- Yoon, J. H., Abdelmohsen, K., Srikantan, S., Yang, X., Martindale, J. L., De, S., Huarte, M., Zhan, M., Becker, K. G., & Gorospe, M. (2012). LincRNA-p21 Suppresses Target mRNA Translation. *Molecular Cell*, 47(4), 648–655. <https://doi.org/10.1016/j.molcel.2012.06.027>
- Zapp, M. L., & Green, M. R. (1989). Sequence-specific RNA binding by the HIV-1 Rev protein. *Nature*, 342(6250), 714–716. <https://doi.org/10.1038/342714a0>
- Zhang, Q., Chen, C. Y., Yedavalli, V. S. R. K., & Jeang, K. T. (2013). NEAT1 long noncoding RNA and paraspeckle bodies modulate HIV-1 posttranscriptional expression. *MBio*, 4(1), 1–9. <https://doi.org/10.1128/mBio.00596-12>
- Zhang, W., Du, J., Yu, K., Wang, T., Yong, X., & Yu, X.-F. (2010). Association of Potent Human Antiviral Cytidine Deaminases with 7SL RNA and Viral RNP in HIV-1 Virions. *Journal of Virology*, 84(24), 12903–12913. <https://doi.org/10.1128/jvi.01632-10>
- Zhang, X., Hong, R., Chen, W., Xu, M., & Wang, L. (2019). The role of long noncoding RNA in major human disease. *Bioorganic Chemistry*, 92(August), 103214. <https://doi.org/10.1016/j.bioorg.2019.103214>
- Zhu, L., Zhu, Y., Han, S., Chen, M., Song, P., Dai, D., Xu, W., Jiang, T., Feng, L., Shin, V. Y., Wang, X., & Jin, H. (2019). Impaired autophagic degradation of lncRNA ARHGAP5-AS1 promotes chemoresistance in gastric cancer. *Cell Death and Disease*, 10(6). <https://doi.org/10.1038/s41419-019-1585-2>
- Zila, V., Margiotta, E., Turoňová, B., Müller, T. G., Zimmerli, C. E., Mattei, S., Allegretti, M., Börner, K., Rada, J., Müller, B., Lusic, M., Kräusslich, H. G., & Beck, M. (2021). Cone-shaped HIV-1 capsids are transported through intact nuclear pores. *Cell*, 184(4), 1032-1046.e18. <https://doi.org/10.1016/j.cell.2021.01.025>

Zou, H., Wu, L. X., Yang, Y., Li, S., Mei, Y., Liu, Y. Bin, Zhang, L., Cheng, Y., & Zhou, H. H. (2017). lncRNAs PVT1 and HAR1A are prognosis biomarkers and indicate therapy outcome for diffuse glioma patients. *Oncotarget*, 8(45), 78767–78780. <https://doi.org/10.18632/oncotarget.20226>