

UCH-FC
Biotecnología
B826
c.1



UNIVERSIDAD DE CHILE FACULTAD DE CIENCIAS ESCUELA DE PREGRADO

**IDENTIFICACIÓN DEL SISTEMA VIRAL
DEL TIPO dsRNA PRESENTE EN
*Xanthophyllomyces dendrorhous***

Tesis

entregada a la

Universidad de Chile

en cumplimiento parcial de los requisitos

para optar al grado de

Ingeniera en Biotecnología Molecular

Por

Natalia Bravo Ayala

Agosto 2011

Director de Tesis Dr. Marcelo Baeza Cancino

Co-director de Tesis Dr. Víctor Cifuentes Guzmán



INFORME DE APROBACIÓN SEMINARIO DE TÍTULO

Se informa a la Escuela de Pregrado de la Facultad de Ciencias, que el Seminario de Título presentado por la candidata:

NATALIA ANDREA BRAVO AYALA

**“Identificación del sistema viral del tipo dsRNA presente en
Xanthophyllomyces dendrorhous”**

Ha sido aprobado por la Comisión evaluadora y revisora, como requisito parcial, para optar al título profesional de Ingeniera en Biotecnología Molecular.

Dr. Marcelo Baeza Cancino
Director del Seminario de Título

Dr. Victor Cifuentes Guzmán
Co-Director del Seminario de Título

Comisión Revisora y Evaluadora

Dra. Margarita Carú Marambio
Presidenta

Dra. Claudia Stange Klein
Evaluadora



Santiago de Chile, Octubre de 2011

A mis papas, Ange y Migue



Nací un 12 de Diciembre del año 1985, en la ciudad de Santiago. En el año 1990, cuando mi familia terminó de completarse, nos mudamos a Rancagua, lugar donde pase la mayor parte de mi infancia. En esta ciudad, además curse mi enseñanza básica en el colegio Grecia el cual se caracterizó por inculcar en mi el interés de observar mi alrededor. En el año 2000 regresé a Santiago para terminar mi enseñanza media, y a pesar que todos los test vocacionales indicaban que debía estudiar arte, en el año 2005 ingresé a la Facultad de Ciencias de la Universidad de Chile en donde he pasado los últimos 6 años aprendiendo y aplicando lo que me gusta. Hoy 3 de Octubre del año 2011 me titulo de Ingeniera en Biotecnología Molecular.

Agradecimientos

Finalizada esta etapa quiero agradecer a las personas que han estado conmigo para apoyarme y acompañarme en cada uno de mis momentos.

Agradezco a Marcelo, mi tutor y profesor, por dejarme trabajar en su grupo *Killer*, los consejos y la guía. También agradezco a el profe Víctor por ser parte de este proceso y permitirme seguir en avanzando en el.

Imposible no mencionar a las personas que conocí en este camino. Mis fieles compañeros Valentina, Nicolás, Chichi, Bryan, Mary, Lucho, Chilo, Ine, Celeste, Pao, Claudio, Poroto, Yury, Cristóbal y el Papö, por el tiempo, la buena onda y disposición de siempre estar ahí con una sonrisa. Y los que aparecieron en el camino, los monos Mario y JuanMa, Nicole y César, Pame, Koke, PaPablo, Lorelia y Gaby.

A mi grupo de laboratorio, los que aún están y los que ya tomaron otro rumbo, en especial a Salva, mamita Jenny, Rafael, Vicho, MarioC3, tío Andrés y Mauricio, no sólo por la guía en trabajo de lab, sino también por la alegría y las conversaciones extra laboratorio.

A mis amigos, los no-biólogos, que siempre intentan entender lo que les explico, Yokito y Pk, Faby, Barbari, Pauli, Rulo, Tota, Goyo, Barbaritha, Tomy, Matías, Emi, Maxi, Marie, Pablo y Andrés.

Y por sobre todo agradecer a mi familia, por su esfuerzo, comprensión y apoyo en todo. A mis papas por no poner reparo en mi decisión de estudiar esta carrera y a mis hermanos por el apoyo y cariño.

INDICE



Índice.....	iv
Índice de figuras.....	vii
Índice de Tablas.....	ix
Abreviaturas.....	x
Resumen.....	xi
Abstract.....	xiii
1. INTRODUCCION.....	1
1.1. Totiviridae	2
1.1. Sistema viral en <i>Saccharomyces cerevisiae</i>	5
1.1.1. Replicación viral	8
1.2. <i>Xanthophyllomyces dendrorhous</i>	12
2. HIPOTESIS	14
3. OBJETIVOS	14
3.1. Objetivo General.....	14
3.2. Objetivos Específicos.....	14
4. MATERIALES Y MÉTODOS	15
4.1. Cepas.	15
4.2. Plasmidios.	15
4.3. Cultivo de microorganismos.	15
4.3.1. Cultivo de <i>E. coli</i>	15
4.3.2. Cultivo de <i>X. dendrorhous</i>	15
4.4. Técnicas de Biología Molecular.....	16
4.4.1. Extracción de RNA total de <i>X. dendrorhous</i>	16
4.4.2. Extracción de DNA plasmidial.	16
4.4.3. Electroforesis de en geles de agarosa.....	17
4.4.4. Purificación de fragmentos de DNA a partir de geles de agarosa.	17
4.4.5. Ligación de DNA o dsRNA.	17
4.4.6. Digestiones Enzimáticas.	18
4.4.7. Reacción de Transcripción Reversa.....	19
4.4.8. Formación de segunda hebra de cDNA.....	20
4.4.1. Amplificación por PCR.....	22
4.4.2. Preparación de células electrocompetentes de <i>E. coli</i>	22
4.4.3. Electrotransformación de <i>E. coli</i>	22
4.5. Secuenciación.....	23
4.6. Análisis bioinformático de secuencias.	23
5. RESULTADOS	24
5.1. Purificación de los dsRNA.....	24
5.2. Clonamiento de las moléculas de dsRNA	24

5.2.1.	Clones de moléculas de mayor tamaño mediante RT-PCR.....	26
5.2.2.	Modificación al Método de Amplificación por Partidor Único.....	27
5.3.	Análisis Bioinformático	29
5.3.2	L1-dsRNA-A.....	36
5.3.3.	L2-dsRNA.....	36
6.	DISCUSION	48
6.1.	Clonamiento de las moléculas de dsRNA	48
6.1.1.	dsRNA de mayor tamaño	48
6.1.2.	dsRNA de menor tamaño.....	49
6.2.	Análisis Bioinformático	50
7.	CONCLUSIONES.....	52
8.	BIBLIOGRAFIA	53
9.	ANEXO.....	57

INDICE DE FIGURAS

Figura 1: Organización y corrimiento del marco de lectura en la traducción del genoma viral de <i>S. cerevisiae</i> (L-dsRNA).	6
Figura 2: Ciclo de replicación viral en <i>S. cerevisiae</i>	9
Figura 3: Modelo que remarca la importancia en la eficiencia en el cambio del marco de lectura.....	11
Figura 4: Representación esquemática de la estrategia para la síntesis y amplificación de los cDNAs a partir de dsRNA.....	21
Figura 5: Electroforesis de extracción de dsRNA desde la cepa UCD 67-385 de <i>X. Dendrorhous</i>	25
Figura 6: Secuencia del alineamiento de los clones de la molécula L1 dsRNA-B.....	30,31
Figura 7: Alineamiento de la secuencia traducida de la molécula L1-dsRNA con la base de datos de nucleótidos traducidos (blastx) de Genebank.	32
Figura 8: Sitio de corrimiento del marco de lectura en la molécula L-B 385 descritos por FS Finder.....	34
Figura 9: Alineamiento de las secuencias aminoacídicas, de diversos totivirus, de los ocho motivos conservados definidos en las RdRp virales.....	35
Figura 10: Secuencia nucleotídica y traducida de la molécula L1 dsRNA-A.....	37,38
Figura 11: Sitio de corrimiento del marco de lectura en la molécula L-A 385 descrito por FS Finder.....	39
Figura 12: Secuencia de la molécula L2-dsRNA obtenida por el alineamiento de los clones.....	40
Figura 13: Alineamiento de la secuencia traducida de la molécula L2-dsRNA con la base de datos de nucleótidos traducidos (blastx) de Genebank.	42

Figura 14: Secuencia de la molécula S1-A dsRNA de 730 pb obtenida por el alineamiento de los clones.....	44
Figura 15: Secuencia de la molécula S1-B dsRNA de 797 pb obtenida por el alineamiento de los clones	45
Figura 16: Secuencia de la molécula S2-A dsRNA de 694 pb obtenida por el alineamiento de los clones	46
Figura 17: Secuencia de la molécula S2-B dsRNA de 694 pb obtenida por el alineamiento de los clones	47

INDICE DE TABLAS

Tabla 1: Totivirus con secuencias nucleotídicas completas publicadas.....	3
Tabla 2: Detalle del nombre de los clones obtenidos para la molécula L2- dsRNA.....	27
Tabla 3: Clones obtenidos a partir de la reacción de RT-PCR de la molécula S1- dsRNA.....	28
Tabla 4: Clones obtenidos a partir de la reacción de RT-PCR de la molécula S2- dsRNA.....	28



ABREVIATURAS



cDNA	DNA copia
CML	Corrimiento del marco de lectura
EGE	Elementos Genéticos Extracromosómicos
dsDNA	DNA de doble hebra
dsRNA	RNA de doble hebra
DTT	Ditiotreitol
EDTA	Acido etilendiaminotetracético
LB	Medio de cultivo Luria Bertani
ORF	Marco abierto de lectura
RdRp	RNA polimerasa dependiente de RNA
RT-PCR	Reacción de transcripción reversa acoplada a PCR
SDS	Dodecil sulfato de sodio
ssRNA	DNA de hebra simple
TAE	Amortiguador tris-acetato-EDTA
TE	Amortiguador tris-EDTA
Tris	Tris(hidroximetil)aminometano
UV	Luz ultravioleta
X-gal	5-bromo-4-cloro-3-indolil- β D-galactopiranosido
VLP	Partícula tipo virus
YM	Medio de cultivo completo

RESUMEN

Micovirus es la denominación dada a los virus encontrados en hongos filamentosos y levaduras, cuyo material genético principalmente es de tipo RNA de doble hebra (dsRNA) encapsidado en partículas tipo virus. A diferencia de virus de mamíferos, no poseen un ciclo infectivo extracelular y en general son funcionalmente crípticos en el huésped. El sistema *killer* de *Saccharomyces cerevisiae*, fenotipo dado por la presencia de micovirus, es el mayormente caracterizado genética y bioquímicamente. Este sistema está compuesto por dos virus con genoma de dsRNA, que codifican para las proteínas virales estructurales y para una toxina, la cual es exportada y produce la muerte de las células sensibles.

Xanthophyllomyces dendrorhous es una levadura de importancia biotecnológica en la cual se han encontrado moléculas de dsRNA. En cepas de esta levadura se ha descrito la presencia de hasta cuatro dsRNAs de diferente tamaño (L1, L2, S1, S2), pero se desconoce su identidad y su clasificación como virus. El objetivo principal de la presente memoria de título fue definir si estos dsRNA corresponden a micovirus y su probable estructura genómica.

Los análisis de los resultados obtenidos para los dsRNA de mayor tamaño, sugieren que estos corresponderían a virus del género Totivirus. El estudio molecular muestra que hay dos secuencias de la molécula dsRNA L1, sugiriendo que existirían dos virus con similar tamaño pero con diferente secuencia, conformando así el sistema viral de *X. dendrorhous* UCD 67-385 en L1-A, L1-B y L2. El análisis de su secuencia muestra que sus genomas están estructurados de la misma forma, con dos ORF, uno cercano al extremo 5' que codificaría para la proteína de cápside viral y el segundo ubicado hacia el extremo 3' que codificaría para una RNA polimerasa dependiente de RNA. En

L1-B dsRNA el análisis bioinformático de la secuencia traducida evidencia que estos dos ORFs, ubicados en distinto marco de lectura, están sobrepuestos en 114 pares de bases. En esta sobreposición de ORF se encontraron estructuras características de corrimiento de marco de lectura como lo son un 'pseudoknot' y un sitio resbaladizo, de modo que la RdRp sólo se sintetizaría como una proteína de fusión. Estas estructuras no se encuentran en L1-A dsRNA ya que los ORFs descritos en ella están en igual marco y por tanto no existe una sobreposición de los mismos. En la molécula L2, el 5' ORF presenta mayor identidad con virus F de *Black raspberry* y, a pesar que los ORFs están en distinto marco, no se encontraron estructuras típicas para el corrimiento en el marco de lectura.

Del análisis dsRNA de menor tamaño, se observó variación respecto de sus secuencias, con largos desde 707 a 730 para S1 y de 694 a 739 para S2. En todos se encontró 1 ORF. En la comparación de la secuencia nucleotídica y de la secuencia traducida de los ORFs con la base de datos no se encontró moléculas similares.

Resumiendo, el análisis de todas estas secuencias sugiere que los dsRNAs de mayor tamaño presentes en *X. dendrorhous*, cepa UCD-385, pertenecen a virus de la familia *Totiviridae*, existiendo una coinfección viral en esta levadura.

ABSTRACT

Mycoviruses are viruses found in filamentous fungi and yeasts, whose genetic material is mainly double stranded RNA (dsRNA) encapsidated in virus like particles. Unlike mammalian viruses, mycoviruses don't have an extracellular infective cycle and are, in most cases, functionally cryptic in the host. The killer system of *S. cerevisiae*, a phenotype encoded by mycoviruses, is the most characterized at genetic and biochemical level. This system is composed by two dsRNA viruses that encode for structural viral proteins and for an extracellular toxin that produces death of sensitive cells.

In the biotechnologically important yeast *Xanthophyllomyces dendrorhous* the presence of dsRNA with different sizes (L1, L2, S1, S2) has been described in some strains. The identity and classification of these dsRNAs remains unknown. The main goal of this study was to define if these dsRNAs belong to mycoviruses and determine their possible genomic structure.

Analysis of the results obtained for dsRNAs of larger size, suggests that these dsRNA belongs to the Totivirus genus. Interestingly, two sequences were obtained for dsRNA L1, suggesting that there are two viruses of similar size but different sequences. Therefore the L1-A, L1-B and L2 dsRNAs conform the *X. dendrorhous* UCD 67-385 viral system. The structure of this three genomes is similar showing a 5'-ORF that encodes for the viral capsid protein and a 3'-ORF that encodes for a RNA- dependent RNA polymerase. In L1-B the ORFs are in different reading frames, overlapping in 114 base pairs. In this overlapping region characteristic structures for frameshifting such as pseudoknot and slippery site were found, so the RdRp would be synthesized as a fusion protein. A similar organization was found for ORFs of L2 dsRNA, but typical structures for

frameshifting were not found. Instead, the ORFs of L1-A dsRNA are in the same reading frame and not overlapped.

In the analysis of small size dsRNA sequence variation was observed, with lengths going from 707 to 730 bp for S1 and lengths from 694 to 739 bp for S2. Only one ORF was found in each small size dsRNA studied and they showed no homology for nucleotide or translated sequences when analyzed in the database .

In summary, bioinformatic analysis of L1-A, L1B and L2 dsRNAs of *X. dendrorhous* UCD-385 strain suggest they belong to the Totiviridae virus family, so therefore there should be a multiple dsRNA infection in this yeast so this dsRNAs could enter the cell.

1. INTRODUCCION

Los virus infectan todo tipo de organismos y tipos celulares. Esta versatilidad es dependiente de las diferencias tanto en la morfología de su cápside proteica y presencia o no de manto, así como en el tipo de material genético que poseen. En particular, aquellos que han sido encontrados en hongos filamentosos y levaduras se les ha denominado micovirus. El material genético de los micovirus puede ser de DNA de simple (ssDNA) o doble hebra (dsDNA), circular o lineal, o de RNA de doble hebra (dsRNA) con genoma encapsidado en partículas tipo virus (VLP). Estos últimos no poseen una fase infectiva extracelular en su ciclo de replicación, por lo que no son considerados infecciosos, y son transmitidos a la descendencia celular o a ciertas células, por apareamiento de estas (Schmitt y Breinig, 2002). A pesar de ser infecciones latentes en su hospedero, su presencia no se ha relacionado de manera clara con alguna expresión fenotípica, con la excepción de la hipovirulencia o de las toxinas 'killer' en ciertas especies de levaduras (Compel y col., 1999; Nuss, 2005). Aún así, se consideran de un gran éxito puesto que son frecuentes en todas las taxas de hongos.

El Comité Internacional de Taxonomía de Virus (ICTV) realiza una clasificación de los virus en base a sus características físicas y el hospedero a infectar. Los virus con genoma de RNA se han clasificado en diez familias, siete de las cuales corresponden a virus de RNA de doble hebra (dsRNA). La descripción de las familias se realizó según la presencia y características de una cápside que proteja su genoma y la segmentación del dsRNA. Las familias descritas son: *Birnaviridae*, *Chrysoviridae*, *Cystoviridae*, *Hypoviridae*, *Reoviridae*, *Partitiviridae* y *Totiviridae*. Dentro de estos, se incluyen virus de importancia médica, veterinaria y agropecuaria que afectan a vertebrados, invertebrados, plantas, hongos y procariontes. La mayoría de estos virus, aunque no todos, tienen una estructura de

partícula viral de forma icosaédrica, con un diámetro entre los 30 y 90 nm, y en su genoma contiene el único gen que es común en todos los virus de RNA: el gen para una RNA polimerasa dependiente de RNA (RdRp) (Bruenn, 1993). Esto da lugar a muchas similitudes, no solo en estrategias de replicación, sino también en propiedades estructurales y bioquímicas, lo que provee evidencia de la existencia de un ancestro común. De estas familias, sólo *Hypoviridae*, *Partitiviridae* y *Totiviridae* tienen como hospedero levaduras y hongos filamentosos, siendo *Totiviridae* la familia mejor descrita. Los miembros de esta familia se caracterizan por tener un genoma no segmentado que codifica para dos proteínas esenciales para su replicación y mantenimiento (Ghabrial, 1998). Pertenecen a esta familia los virus presentes en *Botryotinia fuckeliana*, *Epichloe festucae*, *Magnaporthe oryzae*, *Saccharomyces cerevisiae* y *Ustilago maydis*.

1.1. Familia *Totiviridae*

Los totivirus que infectan levaduras, hongos y protozoos son los únicos virus de dsRNA que presentan genoma lineal no segmentado de un tamaño entre 4,6 y 7 kb. El virión contiene a una hebra doble de RNA que no tiene 5'-cap. Algunos virus tienen un dsRNA adicional que codifica para una toxina proteica y que es encapsidado separadamente del genoma viral. (Fauquet y cols., 2005). Se ha descrito y anotado la secuencia nucleotídica completa de veinticuatro virus pertenecientes a la familia *Totiviridae* (Tabla 1). La organización de su genoma evidencia la presencia de dos ORFs, generalmente sobrepuestos. Un ORF codifica una proteína de cápside (Gag) ubicado hacia el extremo 5' y el otro ORF próximo hacia el extremo 3' codifica para una RdRp. La traducción del mensajero que codifica para la polimerasa sólo se produce por medio de un mecanismo conservado que involucra la síntesis de una proteína de fusión. En la familia *Totiviridae* se proponen tres mecanismos básicos en la traducción de la RdRp. El

Tabla 1: Totivirus con secuencias nucleotídicas completas publicadas.

Virus	Acrónimo	Nº acceso	Tamaño (pb)
<i>Black raspberry virus F</i>	BRVF	NC009890.1	5077
<i>Botryotinia fuckeliana totivirus 1</i>	BFTV1	NC009224.1	5261
<i>Conithyrium minitans</i>	CmRV	AF527633	4975
<i>Eimeria brunett RNA virus 1</i>	EbRV1	AF356189	5358
<i>Epichloe festucae virus 1</i>	EFV1	AM261427	5109
<i>Giardia lamblia virus</i>	GLV	L13218	6100
<i>Giardia lamblia virus cepa China</i>	GLV-C	AF525216	6273
<i>Gremmeniella abietina RNA virus L1</i>	GaRVL1	AF337175	5133
<i>Gremmeniella abietina RNA virus L2</i>	GaRVL2	NC005965.1	5129
<i>Helicobasidium mompa No.17 L</i>	HMN17	NC005074.1	5207
<i>Helminthosporium victoriae 190S virus</i>	Hv190SV	U41345	5179
<i>Leshmania RNA virus 1-1</i>	LRV1-1	M92355	5284
<i>Leshmania RNA virus 1-4</i>	LRV1-4	U01899	5283
<i>Leshmania RNA virus 2-1</i>	LRV2-1	U32108	5241
<i>Magnaporthe oryzae virus 2</i>	MaOV2	NC010246.1	5193
<i>Saccharomyces cerevisiae virus L-A</i>	ScV-L1	J04692	4579
<i>Saccharomyces cerevisiae virus L-a</i>	ScV-La	U01060	4615
<i>Sphaeropsis sapinea RNA virus 1</i>	SsRv1	AF038665	5163
<i>Sphaeropsis sapinea RNA virus 2</i>	SsRv2	AF039080	5202
<i>Thielaviopsis basicola dsRNA virus 1</i>	CeRV1	AY561500	5310
<i>Trichomonas vaginalis virus</i>	TVV	U08999	4647
<i>Trichomonas vaginalis virus II</i>	TVVII	AF127178	4674
<i>Trichomonas vaginalis virus 3</i>	TVV3	AF325840	4844
<i>Ustilago maydis virus H1</i>	UmVH1	NC003823	6099

primero es un mecanismo hipotético de salto del ribosoma usado por el virus 2-1 de RNA presente en *Leshmania* para la producción de una proteína de fusión capsídeo-RdRp (Scheffter y col., 1994; Scheffter y col., 1995). El segundo mecanismo está basado en diferentes tipos de corrimiento en el marco de lectura del ribosoma. El virus de *Giardia lamblia* (Wang y col., 1993) y *Saccharomyces cerevisiae* L-A (Dinman y col., 1991) producen una proteína de fusión Gag-RdRp por medio de un mecanismo en el cambio de marco de lectura en -1 durante la traducción. En cambio el virus 1-1 de RNA presente en *Leshmania* y el virus *Trichomonas vaginalis* cepa T1 probablemente usan un mecanismo en el cambio de marco de lectura +1 (Ghabrial, 1998) (Tai y Ip, 1995). Ambos mecanismos necesitan la presencia de un sitio resbaladizo heptamérico y una estructura de 'pseudoknot' en la zona de sobreposición (Diamond y col., 1989; Tzeng y col., 1992). El tercer mecanismo usado para traducir la RdRp es un acoplamiento hipotético de término-reinicio, descrito en *Helminthosporium victoriae* 190S (Soldevila y Ghabrial, 2000). En el 5' UTR se postula la presencia de un sitio funcional de entrada interna al ribosoma (IRES) que dirige la traducción del ORF de gag por un mecanismo cap-independiente (Huang y Ghabrial, 1996). Además sitios que son esenciales para la encapsidación, replicación y corrimiento en el marco de lectura han sido definidos en el genoma viral.

El virión está formado por una única proteína de peso molecular entre los 73 y 88 kDa, que en ciertos casos se le ha descrito actividad quinasa o actividad catalítica. La RNA polimerasa está presente como proteína de fusión Gag-Pol, en donde el dominio Pol de la proteína tiene tres sitios de actividad de unión de RNA de hebra simple. Cerca de la cuarta parte del N-terminal de la polimerasa es necesaria para el ensamblaje del virus y la encapsidación de la hebra positiva de

RNA y por tanto, uno de estos sitios de unión a RNA está en esta zona. El virión no presenta lípidos y no se ha reportado la presencia de carbohidratos.

1.1. Sistema viral en *Saccharomyces cerevisiae*.

Los virus de levaduras mejor caracterizados son los presentes en la levadura *S. cerevisiae*, responsables del fenotipo 'killer' descrito en la década del 60'. Este fenotipo corresponde a la producción una proteína de bajo peso molecular, señalada como una toxina, que es responsable de la actividad antifúngica (Schmitt y Breinig, 2002). Los viriones son partículas icosaédricas, de un diámetro de 39 nm, que contienen una molécula de dsRNA de 4,6 kb, denominado ScV-L (Wickner, 1996). El genoma viral codifica principalmente para una proteína estructural de la cápside viral de 76 kDa denominada Gag, y también para una RNA Polimerasa dependiente de RNA (RdRp) de 180 kDa (Fig 1). Se evidencia además, la presencia de un segundo dsRNA, de tamaño entre los 1,6-1,8 kb (ScV-M), denominado virus "Killer", puesto que codifica para una preprotoxina, precursora de la toxina, que participa también en la inmunidad contra la misma. ScV-M utiliza las proteínas generadas a partir de la expresión de los genes codificados en ScV-L para poder formar partículas virales y en base a esto es que se le considera un virus satélite (Schmitt y Breinig, 2006). Aún ocupando las misma proteínas, este segundo dsRNA se encapsida en partículas independientes a las que contienen al genoma de la partícula denominada L-A, puesto que no está unido físicamente a él (Preisig y col., 1998). Fundamentado en esto y en la evidencia de encontrar más de una molécula ScV-M dentro de la partícula viral, se observa que la capacidad de contenido de la partícula está determinada por la estructura de la proteína de cápside y no por el tamaño del genoma.

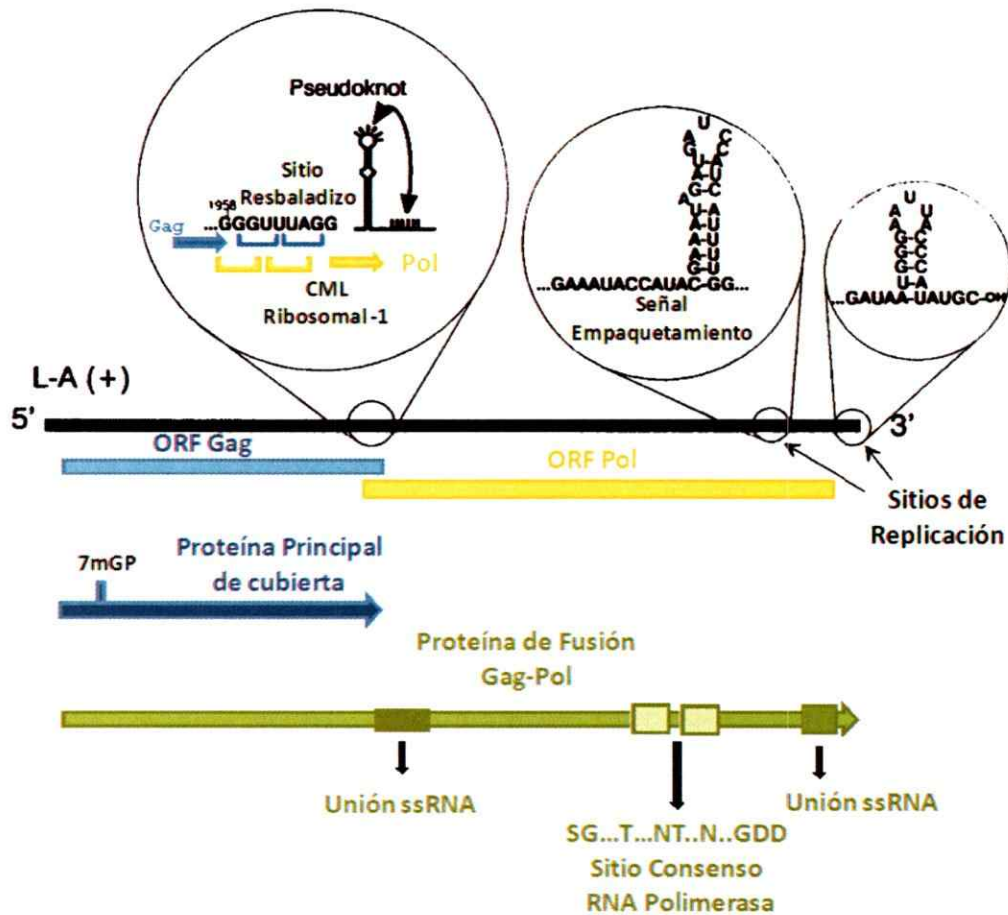


Figura 1: Organización y corrimiento del marco de lectura en la traducción del genoma viral de *S. cerevisiae* (L-dsRNA). En celeste se presenta el ORF codificante para la proteína de la cápside, Gag, y en amarillo, el ORF codificante del segmento Pol del a proteína de fusión Gag-Pol (Adaptado de Wickner, 1996)

En el genoma de ScV-L-A existen dos marcos abiertos de lectura, sobrepuestos en 130 nucleótidos en donde el 3'-ORF está en un marco -1 en relación al 5'-ORF (Diamond y col., 1989). El ORF cercano al 5' codifica para una polimerasa en la cual existen tres dominios de unión a RNA de hebra simple, cada uno necesario para la propagación del virus.

Uno de estos dominios se encuentra en la parte central de la molécula (residuos aminoacídicos 506 a 546). Este dominio es inusual puesto que sólo se detecta luego de la eliminación de una parte del C-terminal de la polimerasa. Se ha postulado que este motivo forma parte de los elementos requeridos por la maquinaria de replicación viral. Esta replicación viral exige que la polimerasa se una al RNA viral de hebra simple, a través de los motivos de unión, y luego lo suelte (Ribas y col., 1994; Routhier y Bruenn, 1998). Los otros dos motivos se encuentran próximos al extremo carboxilo. Estos forman un "loop", en donde el primero está involucrado en el reconocimiento de la proteína de la capsida y el empaquetamiento del RNA y el segundo es necesario para la replicación de la molécula una vez dentro de la partícula viral. Por tanto estas regiones son consideradas como una región de alto grado de identidad dentro de los genomas de virus de dsRNA ya que su supresión elimina la capacidad de encapsidación (Ribas y col., 1994; Wickner, 1996). Como es de esperar, los residuos altamente conservados en la hebra positiva de los dsRNA son estrictamente requeridos para la propagación de los satélites, y es por esto que ScV-M presenta doble señal de empaquetamiento (Routhier y Bruenn, 1998).

Con respecto a la proteína Gag, en *S. cerevisiae* se requieren 120 monómeros de ésta para formar cada partícula viral. Esta proteína tiene un rol estructural en la formación de los compartimientos que protegen el genoma de dsRNA de otros componentes celulares, y además cumple un rol catalítico ya que puede remover una estructura 5'cap desde RNAs mensajeros celulares y adjuntarlo

a su propia secuencia aminoacídica en el residuo de histidina 154. Esta actividad catalítica sería un mecanismo de distracción parcial al sistema de exoribonucleasa celular para que los RNAs virales puedan ser traducidos (Masison y col., 1995).

1.1.1. Replicación viral

Dentro de las partículas virales es donde ocurre la síntesis de ambas cadenas de RNA, tanto la positiva como la negativa, pero en diferentes puntos del ciclo de replicación viral.

La cadena positiva de RNA es expulsada de la partícula viral, sirviendo tanto como RNA mensajero para la síntesis de nuevas proteínas virales, y como genoma viral. Estas partículas contienen una hebra positiva, y mediante la acción de la RdRp se produce la síntesis de la hebra negativa (paso de replicación) para formar el dsRNA, completando el ciclo (Fig. 2). El ciclo de replicación de los dsRNAs satélites que codifican para proteínas de la toxina (M1, M2) o dsRNA no completos, denominados defectivos de L-A (como X-dsRNA) es similar al de ScV-L, con la excepción que no todos los transcritos de la hebra positiva son liberados de la partícula.

La proteína RdRp es expresada solamente como una proteína de fusión formada por un evento de corrimiento de marco de lectura (CML) a nivel ribosómico en -1, evento que ocurre alrededor del 2% de la traducción, cuando el ribosoma está en la región de sobreposición de los dos ORF. Para que ocurra este evento de CML se necesita la presencia de dos estructuras en el +RNA: un sitio resbaladizo y la formación de un loop. El sitio resbaladizo es una secuencia de 7 nucleótidos de la forma X XXY YYZ, donde X es cualquier nucleótido, pero las tres bases deben ser la misma, Y sólo puede ser A o T (U) y Z no puede ser G. Esta secuencia permite a los RNAs de transferencia re-aparear en el sitio A y P del ribosoma, debido a la presencia de las bases en Y que permiten un desapareamiento del RNA en el sitio

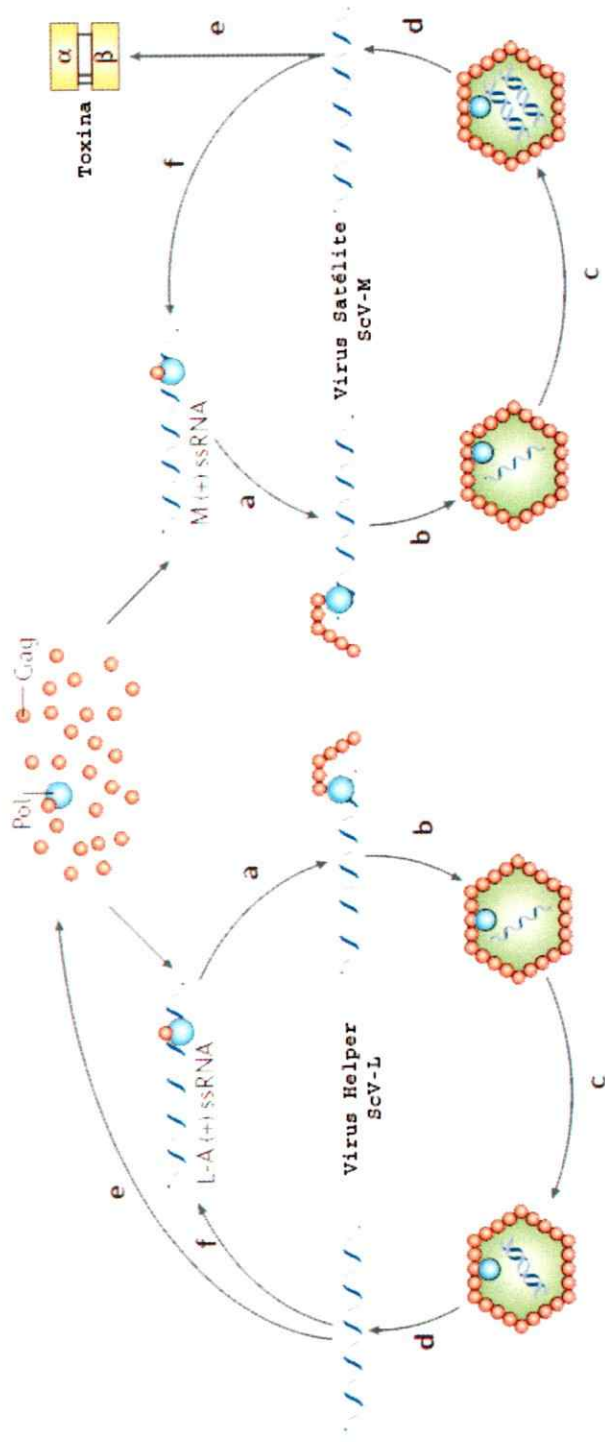


Figura 2: Ciclo de replicación viral en *S. cerevisiae*: Tanto el virus 'helper' (ScV-L) como el virus 'killer' (ScV-M), compiten por las proteínas víricas, Gag y Gag-Pol, codificadas por ScV-LA. (a) Unión de ssRNA a la proteína de fusión Gag-Pol para su encapsidación, (b) ensamblaje de la partícula viral, (c) paso de replicación en donde ocurre la síntesis de la hebra negativa de ssRNA, (d) expulsión de la hebra positiva de ssRNA hacia el citosol en donde ocurrirá (e) la traducción de las proteínas virales codificadas en este genoma y también (f) la transcripción de la hebra de ssRNA (modificado de Schmitt y Breinig., 2006).

A. El nuevo apareamiento ocurre en un marco -1 con respecto a las bases ya traducidas (Farabaugh, 1996). El segundo elemento estructural necesario es generalmente un pseudoknot de RNA que corresponde a una estructura de loop, el cual se une a una secuencia del 3' a la base del tallo (Tzeng y col., 1992). La presencia de esta estructura promueve la detención de los ribosomas con el fin de promover el desapareamiento de los tRNA. De esta forma al retroceder en una base el ribosoma, se comienza a expresar un nuevo marco de lectura que da lugar a la proteína gag-pol. Puesto que la secuencia de la polimerasa incluye los dominios necesarios para el empaquetamiento viral de la hebra positiva de ssRNA y suponiendo que el dominio Gag de la proteína de fusión asegura que la polimerasa esté en la partícula viral mediante la interacción proteína gag-gag, la existencia de la proteína de fusión garantizaría el empaque del genoma viral y la inclusión de la polimerasa de RNA en la partícula (Dinman y Wickner, 1992; Wickner, 1996).

La eficiencia del CML-1 determina la tasa de proteínas Gag y Gag-pol producidas en la célula. Así, la concentración de proteína Gag es de 50 veces más que la concentración de la proteína Gag-Pol. Un exceso de Gag y por lo tanto una disminución de Gag-Pol tiene como consecuencia un ensamblaje de partículas virales incompletas, puesto que se pueden formar partículas cerradas antes que la proteína de fusión Gag-Pol se una a un ssRNA. Esto resulta en la disminución gradual de los niveles de ScV-M y su eventual pérdida. Al contrario, una disminución en la concentración de la proteína Gag y un exceso de Gag-Pol provoca un aumento en el ensamblaje de partículas virales incompletas, pero pocas partículas completas si es que la unión del ssRNA no es un prerrequisito para la polimerización (Fig. 2) (Dinman y Wickner, 1992).

Ya que el sistema viral presente en *S. cerevisiae* fue el primero en ser caracterizado, sentó las bases para descubrir y estudiar sistemas virales similares en otras levaduras y hongos filamentosos. Tal es el caso de *Ustilago Maydis*

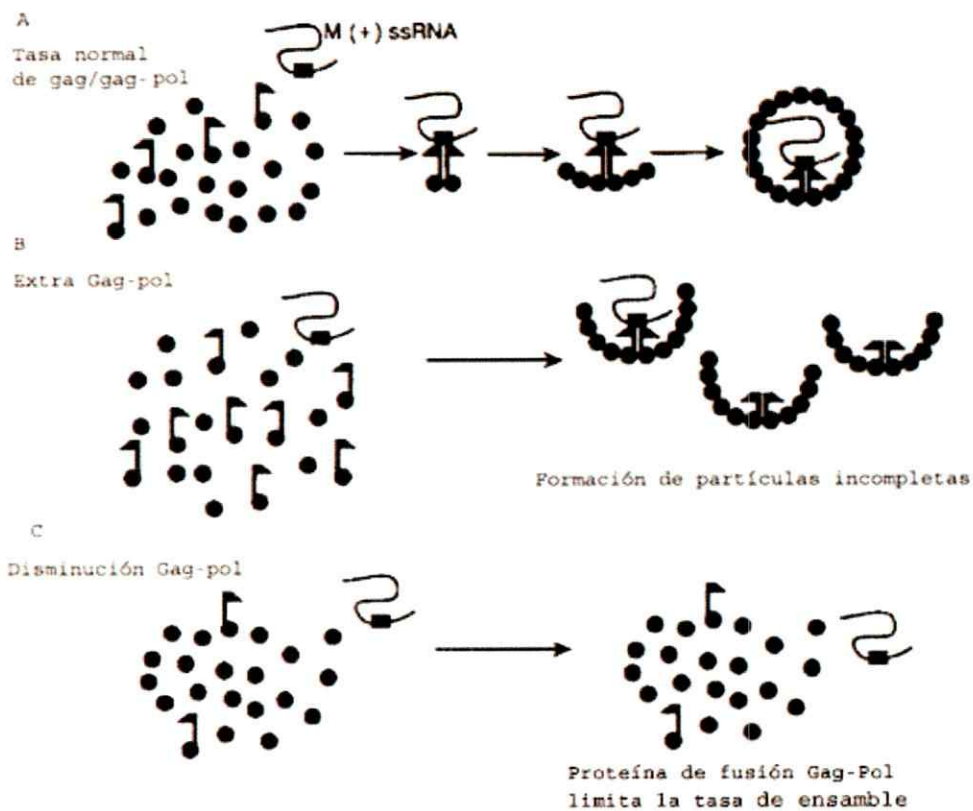


Figura 3: Modelo que remarca la importancia en la eficiencia en el cambio del marco de lectura. (A) Una correcta tasa de proteínas gag/gag-pol permiten que los dímeros de gag-pol una a un ssRNA produciendo el ensamble correcto de partículas virales. (B) Un exceso de gag-pol resulta en la formación de partículas incompletas y pocas partículas completas. (C) Una disminución de gag-pol tiene como efecto en un ensamblaje de partículas virales incompletas, puesto que se forman partículas cerradas antes que la proteína de fusión Gag-Pol se una a un ssRNA (Modificado de Dinman y Wickner, 1992).

(Koltin y Day, 1976) o *Botrytis cinerea* (Castro y col., 1999), ambas de importancia económica al ser hongos que afectan al maíz y los viñedos respectivamente. Bajo esta línea, existen otras levaduras de importancia biotecnológica e industrial, tal como *Xanthophyllomyces dendrorhous*, en la que ya se ha descrito la presencia de cuatro elementos extracromosómicos (EGEs) del tipo dsRNA en la cepa UCD 67-385 y que es necesario avanzar en su caracterización.

1.2. *Xanthophyllomyces dendrorhous*.

X. dendrorhous es una levadura basidiomicete de importancia biotecnológica debido a su producción de carotenoides, tales como fitoeno, licopeno, β -caroteno y principalmente astaxantina (cerca del 85% total de los pigmentos celulares), siendo este último ocupado en la industria acuícola como colorante de salmones para generar la pigmentación naranja de estos y en la industria farmacéutica como antioxidante. En esta levadura se ha descrito la presencia de elementos genéticos extracromosómicos del tipo dsRNA (Castillo y Cifuentes, 1994; Pfeiffer y col., 1996), que presenta un polimorfismo tanto en tamaño como en número. Entre las distintas cepas aisladas desde diferentes sitios geográficos, tales como Japón, Rusia, Alaska, Finlandia, Argentina y Chile, se ha descrito la presencia desde cero hasta cuatro moléculas de dsRNA, las cuales han sido denominadas L1-dsRNA (5kb), L2-dsRNA (3,7 kb), M-dsRNA (1,2 kb), S1-dsRNA (0,9 Kb) y S2-dsRNA (0,8Kb) (Baeza y col., 2009).

El trabajo inicial en los dsRNA de *X. dendrorhous* fue relacionar el origen de las cepas con el número de moléculas encontradas (polimorfismo). Aún cuando se logró generar una relación, experimentos de hibridación de sondas, las cuales fueron generadas mediante transcripción reversa con marca radioactiva, evidenciaron la variabilidad existente en las distintas moléculas según su origen.

Trabajos anteriores mostraron la presencia, a nivel citoplasmático, de VLPs de un tamaño aproximado de 40 nm. Pfeiffer y cols (1996) además indican que existe una copurificación de estas VLPs con los dsRNA, sugiriendo que L1-dsRNA forma parte de estas partículas. Por otro lado, en un intento por caracterizar estas moléculas, se logró clonar parcialmente los genomas L1 y L2 de la cepa UCD 67-385 (Sanhueza M., 2008).

En el presente trabajo se propone caracterizar completamente las cuatro moléculas presentes en la cepa UCD 67-385. Esto se realizará mediante la clonación de estos dsRNA y el análisis, tanto de su secuencia nucleotídica, como de la secuencia aminoacídica traducida *in silico*.

2. HIPOTESIS

Los dsRNA presentes en *X. dendrorhous* corresponden a genomas virales pertenecientes a la familia *Totivirus*.

3. OBJETIVOS

3.1. Objetivo General

Identificar y caracterizar la estructura genómica de las moléculas de dsRNA presentes en *X. dendrorhous*.

3.2. Objetivos Específicos

- Purificación de cada uno de los de dsRNAs presentes en la cepa UCD 67-385 de *X. dendrorhous*.
- Optimización del protocolo de clonación de dsRNA
- Clonación y secuenciación de cada uno de los dsRNAs.
- Análisis bioinformático de las secuencias obtenidas.

4. MATERIALES Y MÉTODOS

4.1. Cepas y plasmidios.

La cepa UCD 67-385 de *X. dendrorhous* se utilizó para la extracción de dsRNA. La cepa electrocompetente de *E. coli* DH5 α se utilizó para la clonación de los dsRNAs usando como vector de clonación los plasmidios pBluescript SK2 (pBS) y pCR-Blunt II-TOPO.

4.2. Partidores.

En la Tabla A1 del anexo se muestran los nombres y descripción de los partidores que se utilizaron en este trabajo.

4.3. Cultivo de microorganismos.

4.3.1. Cultivo de *E. coli*.

E. coli se cultivó a 37°C en agitación constante por al menos 18 h, en tubos de 12 ml con 6 ml de caldo LB, que contiene 1% de triptona, 0,5% de extracto de levadura y 0,5% de NaCl. En el caso de cultivo en placa se agregó al medio LB, Bacto-Agar al 1,5%. Para el reconocimiento de clones transformantes y que poseen un plasmidio recombinante, se utilizó medio LB-agar suplementado con ampicilina (100 μ g/ml) y X-Gal (30 μ g/ml) o LB-agar suplementado con kanamicina (50 μ g/ml) según el vector utilizado.

4.3.2. Cultivo de *X. dendrorhous*.

X. dendrorhous se cultivó medio YM que contiene 0,3% de extracto de levadura, 0,3% de extracto de malta, 0,5% de bactopectona y 2% de glucosa, en matraces de 250 ml con 100 ml de medio. Los cultivos se crecieron por 5 días en agitación constante a 22 °C.

4.4. Técnicas de Biología Molecular.

4.4.1. Extracción de RNA total de *X. dendrorhous*.

El cultivo de células se centrifugó a 2700g y el pellet de células se lavó dos veces con amortiguador TE (25mM Tris-HCl, 10 mM EDTA, pH 8) y resuspendido en 3 ml del mismo amortiguador. A un volumen de 100 µl de células se adicionó 400 µl de amortiguador TE, 250 µl de microesferas de vidrio de 0,5 mm de diámetro (BioSpec Products Inc., USA) y un volumen de fenol ácido a pH 4,0. La mezcla se agitó por 3 min en un equipo Mini-Beadbeater-16 (BioSpec) y luego se centrifugó a 10.000 g durante 10 min. Luego de extraer la fase acuosa, ésta se sometió a una extracción orgánica con un volumen de fenol ácido y posteriormente con un volumen de cloroformo:alcohol isoamílico (24:1). Los RNAs presentes en la fase acuosa se precipitaron agregando dos volúmenes de isopropanol e incubados a -20°C por 16 h, y luego se centrifugó a 10.000 g durante 10 min. El pellet se resuspendió en 20 µl de agua libre de nucleasas. Para degradar el DNA restante en la extracción de dsRNA se utilizó 1 U de RNasa H por cada µg de ácido nucleico presente. Para degradar el ssRNA presente en la muestra se utilizó 1 U de Nucleasa S1 por µg de ácido nucleico presente. Esta muestra de dsRNA se guardó a -20°C hasta su utilización.

4.4.2. Extracción de DNA plasmidial.

La extracción de DNA plasmidial desde *E. coli* se realizó según lo descrito por Sambrook y cols. (2001) y además mediante el kit de extracción de DNA Plasmidial (Axygene), de acuerdo a las recomendaciones del fabricante. El DNA fue guardado a -20° C hasta su utilización.

4.4.3. Electroforesis de en geles de agarosa.

Las muestras de dsRNA, productos de PCR y DNA plasmidial se prepararon agregando el amortiguador de carga BCP (100mM Rojo cresol, 8,3% de glicerol). La electroforesis se realizó en geles de agarosa con una concentración de agarosa entre 1 y 1,5% con Bromuro de Etidio (0,5 µg/ml) en amortiguador TAE (44,5 mM Tris, 44,5 mM ácido bórico, 1mM EDTA, pH 8.0), siendo este último también el amortiguador de corrida. Como marcadores de peso molecular se utilizaron λ HindIII (Fermentas), 100 bp DNA Ladder (New England Biolabs) y 1Kb DNA Ladder (New England Biolabs).

4.4.4. Purificación de fragmentos de DNA a partir de geles de agarosa.

Las bandas de dsRNA y DNA se recuperaron desde los geles de agarosa, mediante cortar el fragmento de interés, y utilización de el Kit de purificación de DNA UltraClean™ 15 (MO BIO) siguiendo las instrucciones del fabricante.

4.4.5. Ligación de DNA.

Las reacciones de ligación se realizaron siguiendo las recomendaciones del proveedor de la DNA ligasa o kit. En el caso de pBlueScriptSK se utilizó amplicón en una proporción molar 3:1 respecto del vector, con 2U de DNA ligasa (Fermentas) en un volumen final de 10 µl, siendo incubadas durante 16h. La reacción se incubó a 22°C en el caso de extremos cohesivos y a 15°C los extremos romos. La reacción se dializó durante 20 min en membranas de nitrocelulosa (Millipore) contra agua destilada, y fueron guardadas a -20°C hasta su utilización.

4.4.6. Digestiones Enzimáticas.

4.4.6.1. Digestión con RNasa H.

Para degradar el dsRNA luego de la reacción de RT-PCR se utilizó 1U de RNasa H (New England Biolabs) incubando durante 30 min a 37°C en amortiguador 20mM Tris HCl, KCl 0,1 M, MgCl₂ 10mM, mM DTT 0,1mM, sacarosa al 5%, pH 7,5, en un volumen final de 20 µl.

4.4.6.2. Digestión con Nucleasa S1.

Para degradar el ssRNA presente en la extracción de dsRNA, o para asegurar extremos romos en los dsRNA antes de ligar el adaptador, se utilizó 1U de Nucleasa S1 (New England Biolabs) por µg de ácidos nucleicos e incubó durante 40 min a 37°C en amortiguador acetato de potasio 30 mM pH 4,5, NaCl 0,3M, sulfato de Zinc 1mM y glicerol al 5%, en un volumen final de 20 µl.

4.4.6.3. Digestión con Endonucleasas de Restricción

Para obtener el vector pBlueScriptSK lineal y con extremos romos, este se incubó con 1 U de por la enzima *EcoRV* (Fermentas) por cada µg de DNA en buffer Tris-HCl10mM, MgCl₂10 mM, KCl 100 mM y 0,1 mg/ml BSA, pH 8,5. Se ocupó 1U de enzima por cada µg de DNA.

Para linearizar el vector pBlueScriptSK con el fin de poder determinar el tamaño del inserto clonado, se incubó con la enzima *NotI* (New England Biolabs) en buffer Tris-HCl 50mM, MgCl₂ 10 mM, NaCl 100 mM y 0,1 mg/ml BSA, pH 7,5. Se ocupó 1U de enzima por cada µg de DNA.

4.4.6.1. Digestión con DNasa I

Para degradar el DNA restante en la extracción de dsRNA se utilizó 1 U de enzima por cada µg de ácido nucleico presente. Se incubó durante 10 min a 37°C

en amortiguador Tris HCl 10 mM, pH 7,5, MgCl₂ 2,5 mM, CaCl₂ 0,5 mM en un volumen final de 50 µl.

4.4.7. Reacción de Transcripción Reversa

4.4.7.1. Reacción de Transcripción Reversa (RT)

Las reacciones de RT realizadas para obtener el cDNA a partir de los dsRNA fueron realizadas con la enzima Transcriptasa Reversa M-MLV (Invitrogen). Las reacciones de RT se realizaron en un volumen final de 20 µl. La reacción contenía entre 300 y 500 ng de dsRNA purificado, DMSO al 20%, 2 µM de cada partidor o hexanucleótidos de secuencia aleatoria, amortiguador 5X (Tris-HCl 250 mM (pH 8.3), KCl 375 mM, MgCl₂ 15 mM), DTT 0.1M, 40U de Inhibidor de RNasa (New England Biolabs), dNTPs 0,2 mM y 200U de enzima. Se incubó a 94°C por 8 min el dsRNA junto con los partidores, en presencia de DMSO y se enfrió rápidamente en hielo. La reacción de RT se realizó adicionando dNTPs, la enzima, amortiguador de enzima, DTT e inhibidor de RNasa y se incubó a 37°C durante 90 min, para luego detener la reacción incubando a 72°C por 15 min.

4.4.7.2. Modificación de la técnica de amplificación por partidor único (M-SPAT)

En las reacciones de SPAT realizadas, en el caso de las moléculas de mayor tamaño para obtener la parte final de cada una, y en el caso de las moléculas de menor tamaño para obtener el clon completo, se utilizó la enzima Transcriptasa Reversa Máxima™ (Fermentas).

El oligonucleótido NBA1 fue ligado al dsRNA purificado previamente. En esta reacción se utilizó 2U de T4 DNA ligasa (Fermentas) junto con 10U de T4 RNA ligasa2 (New England Biolabs) en un volumen final de 20 µl, en presencia de

polietilenglicol (PEG) 4000 al 50% a 37°C durante 20 h (Fig. 4). Luego de esto, reacciones de RT se realizaron en un volumen final de 20 µl. Esta reacción contenía entre 100 y 200 ng de dsRNA purificado unido al adaptador NBA1, DMSO al 20%, 2 µM de partidor NBA-1, 2 µM de partidor interno en caso de ser requerido, amortiguador 5X (250 mM Tris-HCl (pH 8.3), KCl 375 mM, MgCl₂ 15 mM, DTT 50mM), 40U de Inhibidor de RNasa (New England Biolabs), dNTP 0,2 mM y 200U de enzima. Se incubó a 94°C por 8 min el dsRNA junto con los partidores, en presencia de DMSO y se enfrió rápidamente en hielo. La reacción de RT se realizó adicionando dNTPs, la enzima, amortiguador de enzima inhibidor de RNasa y se incubó a 55°C durante 90 min, para luego detener la reacción incubando a 85°C por 5 min. Luego de esto, se completó los extremos de las hebras de cDNA mediante la incubación de estas con la enzima *Taq* DNA polimerasa a 72°C por 10 min. Esta reacción se realizó en un volumen final de 25 µl y contenía amortiguador de PCR 1X pH 8,4 (Tris-HCl 200 mM, KCl 500 mM), amortiguador BCP (Rojo cresol 100mM, glicerol al 8,3%), de MgCl₂ 2 mM, dNTP 0,2 mM, 1 unidad de enzima y un volumen definido de la reacción de transcripción reversa.

4.4.8. Formación de segunda hebra de cDNA

Para sintetizar la segunda hebra del cDNA se utilizó la enzima DNA polimerasa I (New England Biolabs). Las reacciones se realizaron en un volumen final de 40 µl. Esta reacción contenía los 20 µl de la reacción de RT, Amortiguador 2 NEB 1X (Tris HCl 10mM, NaCl 50 mM, MgCl₂ 10 mM , DTT 1 mM , pH 7.9), dNTP 0,2 mM y 50U de DNA Pol I. La segunda hebra fue reparada con 10U de la enzima T4 DNA ligasa (Invitrogen) y completada en sus extremos mediante el uso de 5U de Fragmento Klenow (New England Biolabs).

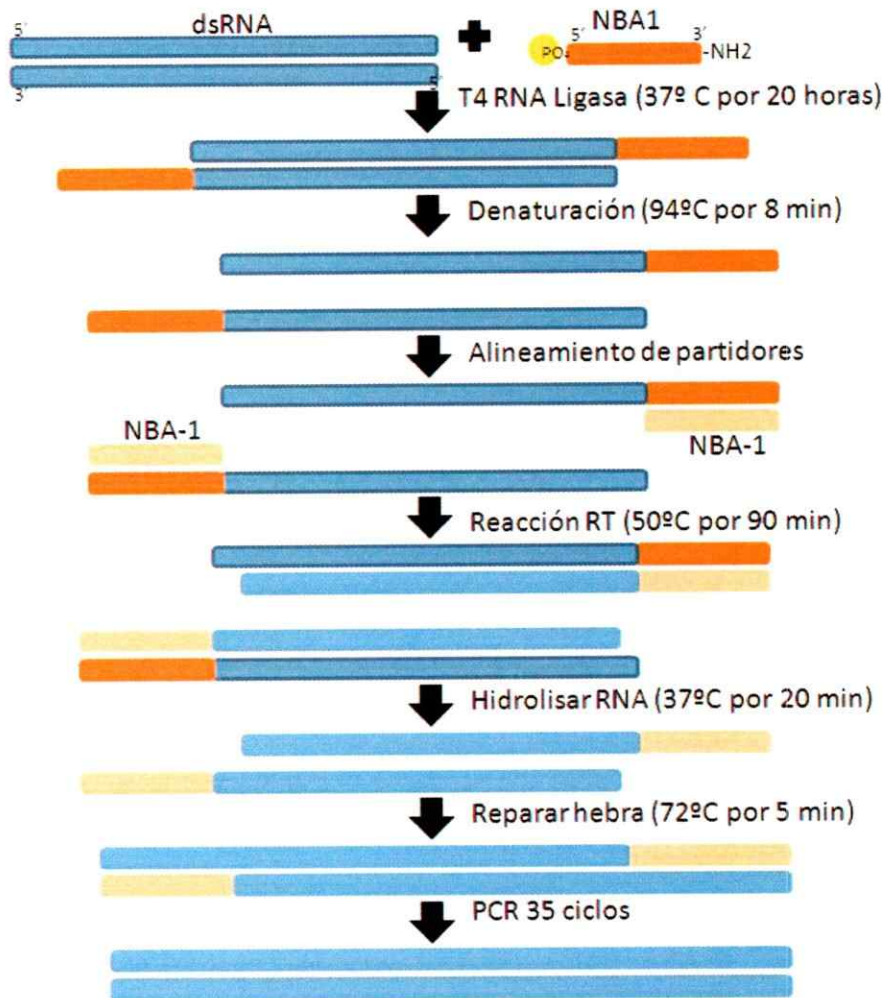


Figura 4: Representación esquemática de la estrategia para la síntesis y amplificación de los cDNAs a partir de dsRNA.

4.4.1. Amplificación por PCR

Las reacciones de PCR se realizaron en un volumen final de 25 μ l. Esta reacción contenía amortiguador de PCR 1X pH 8,4 (Tris-HCl 200 mM, KCl 500 mM), amortiguador BCP (Rojo cresol 100mM, glicerol al 8,3%), $MgCl_2$ 2 mM, dNTP 0,2 mM, 1 μ M de cada partidor, entre 5 y 100 ng de DNA y 1 U de *Taq* DNA polimerasa. Se utilizó un termociclador Multigene (Labnet), y el programa fue: desnaturalización inicial a 94 °C por 3 min, 35 ciclos de desnaturalización a 94 °C por 30 s, alineación de partidores a 55°C por 30 s y elongación a 72 °C por 3 min, y por último elongación final a 72°C por 10 min. En algunos casos, la temperatura de alineación de partidores fue modificada para aumentar la especificidad de la reacción.

4.4.2. Preparación de células electrocompetentes de *E. coli*.

E.coli se cultivó en 500 ml de medio LB con ácido nalidíxico (15 μ g/ml) usando como inocúlo 5 ml de un cultivo de 12 h y se incubó a 37°C con agitación constante hasta alcanzar una densidad óptica a 560 nm entre 0,4 y 0,6 (fase exponencial temprana media). Las células se colectaron por centrifugación a 4.000 g durante 15 min a 4 °C. El sobrenadante se eliminó y las células se lavaron dos veces con agua destilada estéril fría. El primer lavado se realizó con 1000 ml y el segundo con 500 ml. Las células se suspendieron en 20 ml de glicerol 10 % (v/v) a 4 °C, se centrifugaron y luego resuspendieron en 3 ml de glicerol frío al 10 % (v/v). Alícuotas de 40 μ l de células se almacenaron a -80 °C hasta su utilización.

4.4.3. Electrotransformación de *E. coli*.

Los 40 μ l de células electrocompetentes se descongelaron y se le agregó 4 μ l de DNA (30 ng) y la mezcla se electroporó en cubetas de 2mm. Las condiciones de electroporación fueron: 25 μ F, 200 Ω y 2,5 kV. Luego de la electroporación, se

adicionó 1 ml de medio LB y la suspensión de células se transfirió a un tubo Eppendorf e incubó a 37°C por 1 h. Se sembraron alícuotas de 100 µl en placas Petri con LB- agar -ampicilina-X-Gal en el caso de transformación con el vector pBS o en placas placa Petri con LB-agar LB-Kanamicina en el caso de la transformación con pCR-Blunt II-TOPO. Las placas se incubaron a 37 °C hasta el desarrollo de colonias. Se seleccionaron las colonias blancas, las que fueron analizadas mediante PCR.

4.5. Secuenciación.

Para secuenciar los fragmentos clonados, se utilizó el kit de secuenciación "DNA Sequencing Kit Big Dye™ Terminator v3.1" (Applied Biosystems), y el equipo 3100 Avant Genetic Analyser (Applied Biosystems). Se utilizó los partidores universales M13 o partidores de secuencia específica para secuenciar cada clon. Las secuencias se realizaron al menos dos veces.

4.6. Análisis bioinformático de secuencias.

Las secuencias se analizaron con el programa ContigExpress del paquete de programa Vector NTI v10 y la herramienta de búsqueda de alineamientos mediante la comparación de secuencia tBlastx. La predicción de posibles sitios de CML se realizó con la herramienta FSFinder 2.

5. RESULTADOS

5.1. Purificación de los dsRNA

Estudios realizados con anterioridad determinaron que la cepa UCD 67-385 de *X. dendrorhous* contiene elementos extracromosómicos correspondientes a RNA de doble hebra (Castillo y Cifuentes, 1994; Pfeiffer y col., 1996). Se realizaron extracciones de RNA, las que fueron tratadas con nucleasa S1 y DNAsa para degradar RNAs de hebra simple y DNAs presentes. Estas purificaciones fueron analizadas mediante electroforesis en geles de agarosa. Se observó la presencia de cuatro moléculas de dsRNA, dos de mayor tamaño, denominadas L1 (5 Kb) y L2 (3,4 Kb) y dos moléculas de menor tamaño, denominadas S1 (0,9 Kb) y S2 (0,7 Kb) (Fig. 5).

5.2. Clonamiento de las moléculas de dsRNA

Con el fin de identificar la secuencia y por tanto conocer la información codificada en estas moléculas se propuso el clonamiento de estas. Trabajos realizados anteriormente (Sanhueza, 2008) consistentes en la construcción de una genoteca mediante el uso de hexanucleótidos de secuencia aleatoria permitieron obtener un alineamiento total para la molécula L1 de 3,6 Kb y en el caso de la molécula L2 un alineamiento de 1,3 Kb. Por tanto, fue necesario plantear una nueva estrategia para completar la clonación de las moléculas de dsRNA y así obtener la secuencia nucleotídica de cada molécula.

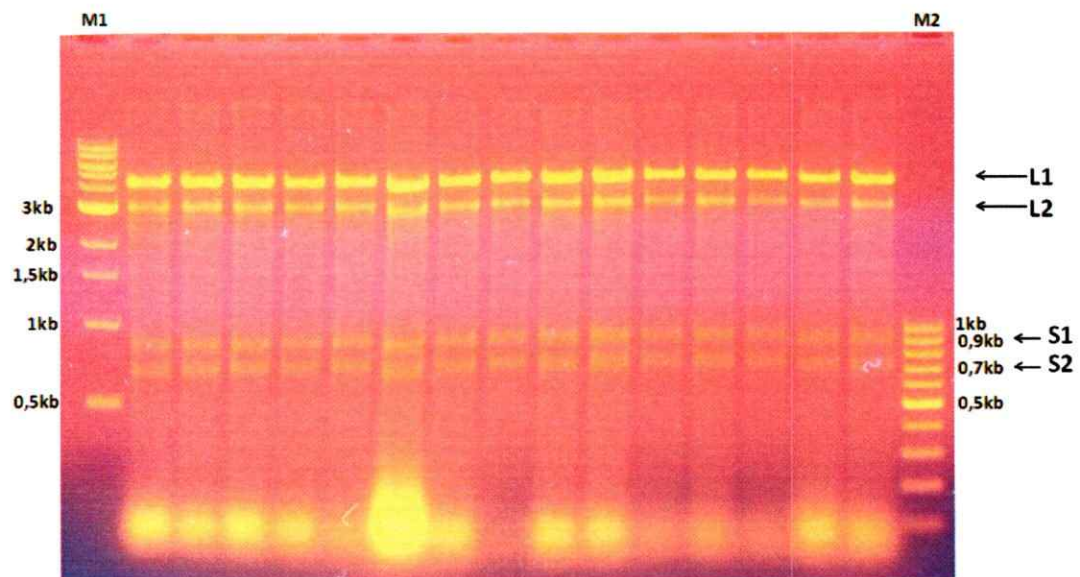


Figura 5: Electroforesis de extracción de dsRNA desde la cepa UCD 67-385 de *X. dendrorhous*. M1: Marcador de peso molecular 1kb; M2: Marcador de peso molecular 100bp; L1: L1-dsRNA (5kb); L2: L2-dsRNA (3,4kb); S1: S1-dsRNA (0,9kb); S2: S2-dsRNA (0,7kb).

5.2.1. Clones de moléculas de mayor tamaño mediante RT-PCR

5.2.1.1. L1-dsRNA

En el caso de la molécula L1, existían clones en la genoteca cuya secuencia no se encontraba en el alineamiento final de 3,6 Kb. A cada uno de estos clones se les realizó un blastx. Uno de ellos, denominado L5, presentó un grado de similitud en su secuencia traducida con una proteína de la cápside del virus de *S. cerevisiae*. Por tanto se diseñaron partidores a partir de esta zona y mediante una reacción de transcripción reversa seguida de PCR (RT-PCR), clonar el fragmento. El clon obtenido fue analizado mediante PCR de colonia y posteriormente secuenciado. Este clon de 928 pb, denominado clon LIV, alineó con la secuencia obtenida previamente y con clones que no ensamblaban anteriormente. Así se obtuvo un alineamiento final con este método de 4950 pb, denominado L1-B-dsRNA.

Durante el desarrollo inicial de la genoteca con partidores aleatorios se encontraron más de nueve clones de cDNA diferentes que no alineaban (LXVI, LXVII, L8.2, L11, L18, L51, L53, L55, L71) con la secuencia final para L1-B-dsRNA, lo que sugiere que estos clones no fueron derivados desde la molécula L1-B-dsRNA. Al realizar reacciones de RT-PCR con partidores específicos de la zona y usados para extender este cDNA parcial, se encontraron clones que alinearon entre sí de acuerdo a su secuencia, obteniendo un alineamiento parcial de 4400 kb, siendo esta otra molécula de dsRNA designada como L1-A-dsRNA.

5.2.1.2. L2-dsRNA

El último alineamiento realizado para la molécula L2 mostró identidad con la molécula L1 en 1327 pb. Por tanto se decidió ocupar partidores diseñados a partir de la molécula L1 para realizar RT-PCR y de esta manera comprobar si es que se

mantenía la identidad de ambas moléculas. Los clones obtenidos como consecuencia de esta estrategia se detallan en la tabla 2.

Tabla 2: Detalle del nombre de los clones obtenidos para la molécula L2-dsRNA.

Clon	Tamaño fragmento
M1	714
M8	561
M43	700
M45	630

El conocimiento de la secuencia de estos clones permitió obtener un alineamiento final de 3065 pb presentando identidad con la molécula de mayor tamaño L1-B-dsRNA.

5.2.2. Modificación al Método de Amplificación por Partidor Único

5.2.2.1. Clonación moléculas dsRNA de menor tamaño

Las reacciones de RT-PCR con hexanucleótidos de secuencia aleatoria realizadas con las moléculas S1 y S2 no permitieron obtener grandes resultados. De esta forma, se buscó una nueva metodología. En el año 1992, Lambden y cols. desarrollaron un método para amplificar el genoma de rotavirus mediante el uso de un partidor único (Lambden y col., 1992). Mediante la unión de un oligonucleótido adaptador de DNA (NBA1) al dsRNA, seguida de un RT-PCR realizado sólo con el partidor de secuencia reversa complementaria al adaptador (NBA-1) se logró obtener clones completos de cDNA de tales moléculas.

5.2.2.2. S1-dsRNA y S2-dsRNA

Los clones obtenidos para la molécula S1-dsRNA se detallan en la tabla 3.

Tabla 3: Clones obtenidos a partir de la reacción de RT-PCR de la molécula S1-dsRNA.

Clon	Tamaño fragmento(pb)
S1 Kan1	730
S1 Kan3	797
S1 Kan4	730
S1 Kan7	797

Por igualdad de secuencia se obtuvieron dos grupos de clones: unos de 730 pb y unos de 979 pb.

Los clones obtenidos para la molécula S2-dsRNA se detallan en la tabla 4.

Tabla 4: Clones obtenidos a partir de la reacción de RT-PCR de la molécula S2-dsRNA.

Clon	Tamaño fragmento(pb)
S2 Kan1	694
S2 Kan5	739

Por igualdad de secuencia se obtuvieron dos grupos de clones: unos de 694 pb y unos de 739 pb.

5.2.2.3. RACE (Rapid Amplification of cDNA ends) de moléculas de mayor tamaño

Para poder completar el conocimiento de la secuencia, tanto en el extremo 5' como en el extremo 3' de las moléculas de dsRNA de mayor tamaño se modificó el método de amplificación por partidor único para poder obtener las secuencias finales. De esta forma, una vez ligado el adaptador (NBA1) al extremo de la molécula, se realizó una reacción de RT-PCR con un partidor interno específico del dsRNA junto al partidor NBA-1 y se analizó la secuencia de los clones obtenidos luego de una transformación. Aquellos clones que en su secuencia se encontrara la secuencia del adaptador se alinearon con la secuencia parcial conocida de cada

dsRNA. De esta forma se logró obtener el extremo 3' de la molécula L1-dsRNA, con un tamaño final de 5491 (Fig. 6). En este clon, denominado L31 se encontró también el codón de término para el ORF descrito.

5.3. Análisis Bioinformático

A partir del alineamiento final para cada molécula de mayor tamaño se buscaron los posibles ORFs presentes y su homología observada en la base de datos Genbank para secuencias traducidas (blastx).

5.3.1 L1-dsRNA-B

Para la molécula de 5491 pb (Fig. 6) se encontraron dos ORFs. Un ORF hacia el extremo 5' de la molécula, en el marco de lectura +1, que abarca desde el nucleótido número 28 hasta el 1957, y codificaría para una proteína de 643 aminoácidos. Un segundo ORF se encontró hacia el extremo 3' de la molécula, en el marco de lectura +3, que se encuentra desde la posición 1995 hasta 4815 pb, y codificaría para una proteína de 940 aminoácidos. Esta molécula presenta un 3' UTR de 519 pb y un 45% de contenido GC. El alineamiento de la secuencia traducida fue comparado con la base de datos se obtuvo alta identidad, con valores de E cercanos a 1×10^{-95} , con secuencias asociadas a RNAs polimerasas dependientes de RNA de sistemas virales, en el marco de lectura +3. La mayor identidad se encontró con el virus F de *Black raspberry* y para el virus L-A de *Saccharomyces cerevisiae*. Además en menor grado de identidad se encontró, en el marco de lectura +1, la proteína de la capsida del virus L-A de *Saccharomyces cerevisiae* (Fig 7). Estas zonas de identidad coinciden con los ORFs predichos anteriormente.

```

- | Y I Y I R L T I G R I G P S R C M L E L P P V Y R I S A E F A L R H
1 | GTGATTGTAA TACGACTCAC TATAGGGGGA ATTGGGCOCT CTAGATGCAT GCTCGAGCTG CCGCCAGTGG TGAGGATATC TGCAGAAATC GOCCTTAGAC
- | H R V V T R A V M E L G G Y K Y A S N L E L A T H F Q V A K A K F D
101 | ATAGGGTTTG GACCAGGGCA GTAATGGAGT TAGGAGGTGT CAAGTACGCT TCAAACCTTAG AGCTAGTAC CCACITTCOA GTGGCGAAAG CTAATTTGA
- | D V I K P Y A R S N F Y G Y N K K F I D E S G Y Y D P L R A I D E Y
201 | GTCATCAAA CCAGTGGCCA GGTCCAACTT TTACGGCTAT AACAGAAAT TTAITGATGA AAGCGGTGTC TACGACCCGC TCAGAGCTAT AGATGAGTAT
- | S R T Y P G M R L G R D D L R A L S L T D N K T D S Q E A Y I Y N M
301 | TCCAGGACCG TACCTGGTAT GAGACTCGGA AGAGATGATT TACGGGCGCT TTCTCTGACG GATAATAAGA CCGACAGTCA AGAGGGCTAT ATATATACA
- | M L Y S V L K A R L Y Y D M K G S D N K F T Y K Y S S F K D T H V G
401 | TGCTASTATC CTGGCTGAAG GCTAGGCTTT ACGTGGACAT GAAAGTTTCT GATAACAAGT TTACASTCAA ATATTCTTCA TTTAAGGACA CGCACGTTGG
- | G Y D I N D S Y G I E S Y E Y N L G P P N P T G E A A I S L E F R D
501 | GTACGATATT AACGACTCCT ACGGCATAGA GTCACTGAG GTCAATCTAG GACCGCCAA CCTACGGGA GAAGCAGTCA TCAGCTTGGG GTTCAGAGAC
- | T N N F V S K P Y Y L K Y S N N S M E Q G S F Y L A H Y L G S N G T
601 | ACTAATAACT TGTGGTCCA ACCCTATGTC TTGAAATACT CTAATAACAG TATGGAACAA GGATCCTTTT ATCTGGCACA CGTCTAGGA TCTAATGGCA
- | T S G L S A D I Q I D A L D F N E L L L D P V G S F P T G A F N I F
701 | CTTCGGACT GTACGCGAC ATACAGATTG ACGCTTGGG CTTCACAGAA TTGCTTTTGG ACCCGTGGG GTCCTTCCCT ACTGGAGCT TCAATATATT
- | F A D F V A K P N V I V L W I M D Y Y R L N R V E Q E F A S A F E L
801 | CGTGTATTTT TGGGCTAAAC GTAACGTAAT CTGGCTTGG ATAATGGATT ATGTGGGCT CAACAGATG GAGCAAGAT TTGCTCAGC TTTTGAATTG
- | L G A L A T Q P L P S Y H E S I L V S K S R T Y Y N M S K F S P T R
901 | CTGGAGCTT TAGCAACTCA ACCCTCCTCC CTCTACCAAG AGATATTTT GTGGTCAAAG TCGAGAAGT TAGTCAACAT GTCAAGATT TCGCCACAA
- | R A R Y P A N L T G E P N V H D L N A Q Q F T F D E E K S P A G F I
1001 | GAGCTAGAT GCGAGCCAC TTGACGGGTG AACCTAAGT ACAATGCTG AACGACAC AGTTCACITT CGACGAGAA AAGTCCCTG CTGGTTTTAT
- | I T A S A Y L N Y A F V I G I Y G M Y S N F A E D C S D V T D A F I
1101 | AACGCGAGC GCGGTGTTAA ATTACGCTT TTGGATAGGC ATTACGGAA TGGTCTCAA TTTTCAGAA SACTGTTCG ATTGGACTGA TGGCTTCATA
- | S S D A E L G I L S T Y E A R P A M I S L V T G K E T N S C F S N N
1201 | TCATCAGACG CGAAGCTGG AATAGTGTCA ACAGTTGAAG CCAGACTGC AATGATAAGT CTGTAAAGG GGAAGAGAC CAACAGTGC TTTTCCAACA
- | N C F L T Y D L S G M Y G Y K Q L I Y D E K I D P N H P G Y I L F D
1301 | ATTGCTTCTT GACTTATGAT CTGCTGGGA TGTATGGAT TAAACAGCTA ATAGTGGATG AGAAGATAGA TCCCAACCAC CCAGGGGTA TACTTTTTGA
- | D T Y P A F V S G S L L M G A Y A T D Y P V L K H L E P H Q H I K Y
1401 | CACTGTGCA GCTTGTGAT CAGGCTCACT ACTCATGGGA GCGGTGGCAA CTGATTACC AGTGTGAAA CACTAGAAC CACCCAACA CATTAAAGT
- | E R D G L L G A R E A A M L A N S Y R L F G N D V I I E H F R S A E
1501 | GAGCGAGACG GCTTGTGGG GCTAGAGAA GCGCAATGT TGGTAAACG TTATAGACTT TTTGGCAGC ATGTATTAT CGAATTTTC AGATCAGCAG
- | E Y Y P T Y A N S E E C V T A T Y E L F G R T R T F D I M R V S S S
1601 | AGGTGTACC AACTTACGCA AACAGTGAAG AGTGTGTCAC AGCCACATAC GAATTTTGG GTAGAAGTAA AACATTGAT ATAATGCGT TTTCTAGTTC
- | S Y K R T G R T Y E I P D A T H V R T Y G E C K L T M D M P V L A Y
1701 | ATACAAGAG ACAGGCGGA CTACGAAAT AACTGACGC ACTCACTTC GTACTACGG TGAGTGAACA CTCACATGG ACATGCCAT ACTGGCGCTC
- | C G V K Q R K T V H R P K M M L D S R R Y A T K F M Y G A T S G F E
1801 | TGTGGTGGG AACAAAGAAA GACGTTGCAC AGAAGTAAA TGTGCTGGA CTCCAGCGA GTAGCAACA AGTTCATGTT AGGTGCGACT TCTGGTTTCG
- | E K T R F A Y Y N R R N V M A Q G F H E A K A E M A P A L P L V R G
1901 | AAAAAAGTAT ATTGCGATC TATAACAGAA GGAAGTTTAT GCGCAGGGT TTTCAAGAG CAAGAAGGGA AATGGTCCA GCACCTCCTC TGGTAAGAGG
- | M Y T F
- | G S A Y S T A T K N L L Y F T I G P
2001 | TTTGGTGTG TCCAGGCA CAGAAACTT GCTATATTC ACGTAGGAC CGTAAGGCT GCATTAAACA GTGCATAAC TGGCAAAGAA ATGGTTACTT

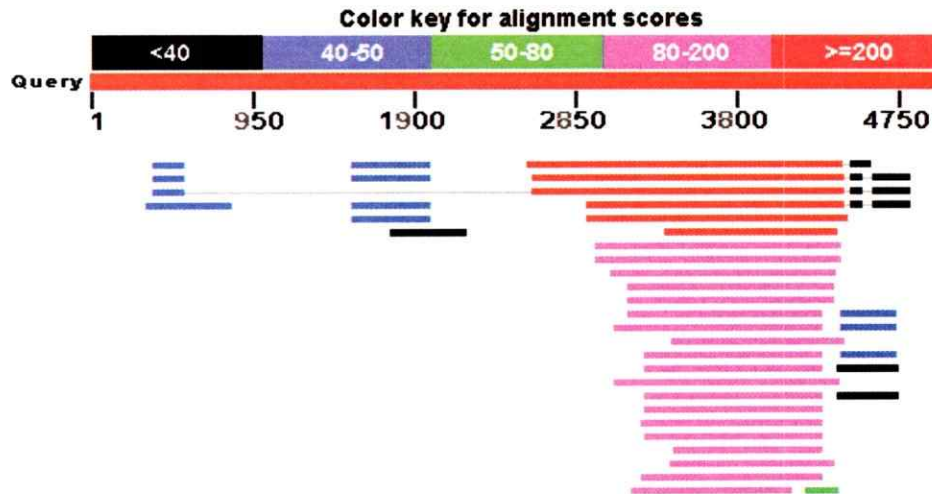
```

```

-> F M N P N L F Y S Y D V T P M V N Y S K Y S F E E Y L E R G Y P S S
2101 TCATGAACCC AAATCTCTT STACTTATG ATGTACTCC GATGGTCAAC GTAAGTAAAG TAAGCTTGA AGAAGTGTTA GAGAGGGAT ACCCGTCATC
-> S L P Y H G V Y P M Y S G S L F L S A N A S D V A S Q C H L E Q G M
2201 TTGCGCCGTA CACGGCGTGG TGCCAAATGTT GTCCAGATCC CTCTCTTGT CCGCTAACGC ATCAGACGTT GCCTCACAGT GTCACTTTGA ACAGGGCATG
-> K Y T Y D E Y G T V S P E D A L R V A Q M Y R M F G H E L E I R S E
2301 TAGGTCAAGG TAGATGAATA TGCCACTGTG AGCCAGAAAG ATGGCGTTAG GGTGGCTCAA ATGTATAGAA TGTTTGGTCA CGAAGTTGAA ATAGATCGG
-> E K T E E I H E L F A P Y Q E C V I Y P S A L L Y N T R D T D R L K
2401 AAAGACTGA GGAATACAT GAAGTGTTCG CTCGGTGC AAGTGGTGC ATTACCCCTT CGGCACACT TTACACACAG AGGGACACAG ATAGATTAA
-> K L V S S L R R P G R S S T I P D Y S A L T A G Q T I T I D Y T I P
2501 GCAAGTGTG TCAGTGGGA GACCTGGCGG CAGTAGCATT ATACAGACG TTTCGGCTCT GACCGTGGG CAACAATCA CCATCGATTA CAGATACCT
-> R I G M H Y F N K R V L P I A P S M Y L P Q R K R E I T F R Y E G H
2601 AGGATAGGCA TGCACTACTT CAACAAAGG GTGCTCCCA TAGCTCCGTC AATGGTATTA CTTCAAAGGA AAAGGGAAT AACGTTGAGG GTCAAAGGCC
-> H T V L H K Y K M N A S N V T R P V K D F H G G E I Q Y A P L L P Y
2701 ATACAGTCTT CCATAAGTGC AAAATGAAGC CTTCAAAGCT GACAAGACT GTGAGGATT TTCAAGGAGG CGAAATACAG GTGGCACCTT TGTTGCCAGT
-> Y L R N A R Y P I Q Y N Q T S E E D Y T A E D Y E
2801 TCTAAGGAAT GCTGCTGTC CAATCCAACT CAATCAGACT TGGAGGAGG ACGTAAGTGC CGAAGATGA GAGTAGAACT ATAAAACCT CAAATACAG
-> Y Q P C E G E E S T
2901 GGAAGTACA CATCGGACTA TGTATGCAA TACGGACAC TATATAACAG GAATTCATC ATAGGGCTTA ACGTGCAGCC ATGTGAAGT GAGGAAGCA
-> T G N R V E F A D R Q I P L T I S L T S R G N I E A Y P F S E A Q Y
3001 GGGACACAG ATGGAGTGT GCGACACAGC AAATTCCTT AAGGATATCT CTAACAAGCC GGGGAACAT TGAAGCATA CCGTTTCAG AAGCTCAGTA
-> V A F L D R E S S N K F G I F E A S K S Y Y G V T M S G L V I A N D
3101 TGCTTCTG GACCGGAAA GCTCAAATA ATTCGCATA TTGAGGCAA GTAAAAGCT GTATGGGTA ACAATGCTG GGTGGTGTAT AGCTAATGT
-> K A K F F F A S I K Q Y L L T Q S K A M L I A M T R H F S S L Y E S
3201 AAGCTAAAT TTCTCTGC CTCATCAAG CAATATTGC TACCCAACT TAAAGCAATG TTAATAGCA GACACAGCA CTTCTCTAG CTTGACGGA
-> S Y A Y D D P T S P V N L F R D R F P G R S E N E A M T F K Q M E E
3301 TCTAGCTTA TGATGATCT TACGACCGG TAAACCTATT CAGACAGACA TTCCCGGTA GAAGCGAAA TGAAGCGAT ACATTCAAGC AAATGGAGGA
-> E L P K S K I S G A H H F T P M Q Y L K S I G R G R L E R A I Q
3401 GCTGCGAAG TCTAAATTT CAGAGCGCA CCAATTTTAC TTCAACCTA TGCAAGTCTT GAAATCGATA GGGAGGGA GACTGGAAAG GCGATCCAG
-> A S R L P K D A E M T F A T G M L L V Y T S I S D Q M A E L I R G A
3501 GCATCAGGCA TACCAAGGA CCGCAAAATG ACGTTGCA CAGGGATCT ACTGTGATC ACATCTATA GTGATCAGT GGCAGACTA ATCAGAGGGG
-> A G I F K C D S Y R E Y Y K E M K Q L S Y E A K S L Q N L Y T D D L
3601 CCGGAATATT CAATGTGAT AGTGTGAGG AATAGTAAA AGAATGAAA CAGTATCAC TGGAGCGAA GTGTGCAA AACCTGTGA CCGATGACT
-> L R T Y F E L E Y L Y N R I D G V Y D V E K E K E N R Q S Y N Y T N
3701 AAGACAGT TTGAGTATG AGGTGCTGT AATAGATTC GACGGTATG TAGACTGGG GAAGAGAGG GAAACAGAC AGTCAATAA CBTGCAAA
-> I X D S D Y F R S A C K I F E D A K G I G R R P K S F S Y V A
3801 ATAAAGATT CAGATGTTT TAGATCTGG TGAAGATAT TTGAAGAGC CAAAGTATA GCGCGCGCC CGAAGAGCTT CAGTTGGAG AGCTATTGG
-> A N R V Q V S A A G S I H S Q Y P R D M E Y V I R D Q Q S L K N K F
3901 CTAACAGT GCAATGAGT GCTGCGGCA GCATACACT ACAATATCC AGAGCAATG AGTACGTAAT TAGGACCAA CAGTCACTA AAACAAGTT
-> F I T I S N M P K C T Y D Y F S D R E P Q V Q G V S S T K Y E V G K
4001 TATACTATA TCCAAATGC CCAATGTAC AGTAGACTC TTTTCGCA GGGAGCGCA GGTCAAGGT TGCTGCTCA CGAAGTATG ATGGGGCA
-> Q R A I Y G T D L T S Y Y L S N F A F Y N C E N Y L P N Q F P V G R
4101 CAAGAGCAA TATATGGAC AGATCTGACT AGCTAGTAC TGCTAATTT CCGTTTTAC AATGTGAGA ACGTACTACC AAATCAATTT CCACTGGAC
-> R D A N D E N V Y N R Y S G Y L N N R M P F C L D F E D F N S G H S
4201 GAGTGCACA TACGAAAAC GTGGTGAACA GATTTTCGG AGTGTAAAT AACAGATGC CTTTCTGCT TGACTTTGAG GACTTCAATA GTCAACACAG
-> S S G N M K A V I Y A Y I E T F I D C L T P E Q E Q A A M V T A A S
4301 CTGGGCAAC ATGAGGCTG TATATAGC TTATAGCAG ACTTTCATG ATGCTTAC TCGGAAACA GAGCAAGCGG CAATGTGGAC CGCAGCTCA
-> L D K Q I I N D N V G T K T Y E S K G T L L S G V R L T T F M N S
4401 TTAGACAAAC AGATAATCAA TGACAACGTT GGAACAGA CTAAGTATG ATCTAAAGG ACGTACTTT CCGGGTGGAG GCTGACACT TTCTGAACA
-> S Y L N Y I Y T T K L A A E E K R P G D S L H N G D D Y G Y R S
4501 GTGTGTTAA CTATATTTT ACACGAAAT TAGCAGCAGA AGAAAAGCT CCGGGTATA GCTTGCACA CGGAGAGAT GTCTTAATAG GAGTGGGT
-> S M A L P Q R C M G N A I K Y N V R M Q S S K C A V G A I A E F L R
4601 AATGGCTGT CCGCAGAGT GTATGCAAAA CCGCAATAAA TATAAGTGC GAATGCAAG CAGCAAATG GCGTGGGTG CAATGGCAGA ATCTCTGGG
-> I D H K Q G G N G Q Y L S R A V A T M V H S R I E S R Y S T D I R D
4701 ATAGACACA AACAGGGAG CAAGCGGAC TACTATCCA GAGCAGTGC CACAATGTA CATTGAGAA TTGAAGTAG AGTCTGACT GACATAGAG
-> D L Y Q S M E N R F A D A K G R G M A N D I I S G L R E Q Y Y M R Q
4801 ACTTAGTACA ATCAATGAA AACAGATTC GTAGCGAAA AGGCAGAGG ATGGGAAAG ACATCATAT CCGATTGCGG GAACAGTACT ACATGAGGCA
-> G S Y L
4901 GAGTGTCTG TGACACCGC GTAGAGGACA TACTTGTAT AAAGATGCG CACAGATTS TAGGAGGAT AAGCGAGAA AAAGACTGA AAATGGGTG
5001 TTGTATAACA TOCCAACCTA GAGCGCAGAA GAAAGTGGC AATACCTGAC TTGCGGGAG TAAATGAATA TGCAACGAA ATACACAAAG CATTAAAAA
5101 AAATGTGCA ATTAAGACTA TATGTGACAG ATTATAGCAG GCTACATAG AAGCTGTCTC AATCAAAGAT CCGAAGATGA AAATCTTAC GGAACACAGA
5201 GATCAGTGTG TGTAAAGT ACGCAGAATA TACAAAGCAG ATAAGGTAG TCAATTGTCA CAGAATTAG GTAAAGCAGC CTTATAGGA TTCCATTGG
5301 AAGTGTGGG CAGAGAAATG CCGATGCAA CCATACACAC AATAGTAAAT ACTTCAAAA GGCCTAGG CTAATATAG CACATATAT AAATGTGAC
5401 AAAAGACCAG CAGGTCCAT ACAAACATC

```

**Figura 6: Secuencia del alineamiento de los clones de la molécula L1 dsRNA-
B. En azul se muestra la secuencia aminoacídica para los ORFs propuestos.**



Color	Acceso	Descripción	Marco	E-value
Rojo	YP_001497151.1	RdRp [Black raspberry virus F]	+3	2e-94
Rojo	AAA50321.1	RNA polimerasa [<i>Saccharomyces cerevisiae</i> L-A]	+3	7e-91
Rojo	NP_620493.1	Proteína de unión a ssRNA [<i>Saccharomyces cerevisiae</i> L-A]	+3	4e-90
Rojo	NP_620495.1	RNA polimerasa [<i>Saccharomyces cerevisiae</i> L-A]	+3	2e-89
Rojo	NP_042581.1	RNA polimerasa [<i>Saccharomyces cerevisiae</i> L-a]	+3	6e-84
Rojo	ACA61232.1	RNA polimerasa [<i>Ribes</i> virus F]	+3	7e-73
Rosa	CAJ29959.1	RdRp [<i>Amasya cherry disease-associate</i>]	+3	2e-37
Rosa	CAJ57274.1	RdRp propuesta [<i>Cherry chlorotic rusty spot</i>]	+3	2e-36
Rosa	CAJ29958.1	RdRp propuesta [<i>Amasya cherry disease-associate</i>]	+3	6e-32
Rosa	NP_620730.2	RdRp propuesta [<i>Trichomonas vaginalis</i> virus]	+3	9e-22
Rosa	AAA62868.1	RdRp propuesta [<i>Trichomonas vaginalis</i> virus 1]	+3	9e-22
Azul	AAA50320.1	Proteína de la cápside [<i>Saccharomyces cerevisiae</i> L-A]	+1	0.23
Azul	NP_620494.1	Proteína principal de la cápside [<i>Saccharomyces cerevisiae</i> L-A]	+1	0.30

Figura 7: Alineamiento de la secuencia traducida de la molécula L1-dsRNA con la base de datos de nucleótidos traducidos (blastx) de Genebank. Esquema con las identidades donde se indica en distintos colores los puntajes obtenidos. En la tabla se observa las principales identidades obtenidas mediante este alineamiento.

En base a esto, se buscó en la región cercana al inicio de la polimerasa, los motivos necesarios para un corrimiento en el marco de lectura. Para esto se utilizó la aplicación bioinformática FSFinder 2 (Moon y col., 2004) la cual busca una zona donde se encuentre un sitio resbaladizo y la formación de una horquilla de 4 a 11 nucleótidos río abajo a modo de región regulatoria. La búsqueda mostró la presencia de 4 posibles sitios resbaladizos, pero en sólo uno de ellos se predice la presencia de una estructura secundaria del tipo pseudoknot (Fig. 8). Este sitio resbaladizo se encuentra a partir de la base 1851, compuesto por tres guaninas y cuatro uracilos, y posee la secuencia: G GGU UUU. Seis nucleótidos río abajo, en la base 1864 se encuentra la región GCA A que formaría un pseudoknot con una región 65 nucleótidos más adelante, a partir de la base 1933. Esta zona regulatoria se encuentra entre los dos ORFs propuestos, evidenciando así una sobreposición de estos en 114 nucleótidos, desde el nucleótido 1843 hasta el nucleótido 1957.

Con respecto a la RNA polimerasa, Bruenn (1993) identificó 8 motivos conservados de las RpRd los cuales se han asociado a un rol importante en la estabilización de la estructura terciaria y en la actividad de la proteína indicando que el motivo 5 y 6 son indispensables (Routhier y Bruenn, 1998). Por tanto se buscó la presencia de estos motivos en la proteína que codifica para la polimerasa. La figura 9 muestra el alineamiento múltiple de la secuencia aminoacídica de los 24 miembros de la familia *totiviridae* de genoma conocido, con el ORF descrito para la polimerasa del sistema viral de la cepa UCD 67-385 de *X. dendrorhous*. Se observó un alto grado de similitud de esta polimerasa con los virus de *Saccharomyces cerevisiae* y *Black Raspberry* ambos pertenecientes al género *totivirus* y en un menor grado con el virus H1 de *Ustilago maydis*, del género *totivirus*, y con el virus de *Giardia lamblia* y el virus *Giardia lamblia* cepa china, del género *giardivirus*.



Figura 8: Sitio de corrimiento del marco de lectura en la molécula L-B 385 descritos por FS Finder. (A) Identificación de los dos ORFs descritos y su sobreposición de 132 nucleótidos evidenciada por un trazo horizontal negro. En rosado se observa la posición del sitio resbaladizo. (B) Estructura predicha para el corrimiento en el marco de lectura. En amarillo se muestra el sitio resbaladizo y en verde las bases nucleotídicas involucradas en la formación de la estructura secundaria pseudoknot.

5.3.2 L1-dsRNA-A

De la molécula L-A se logró obtener la secuencia nucleotídica de 4.592 pb (Fig. 10). Desde la base 1175 hasta el 3' final de esta molécula existe una similitud con la secuencia conocida de la molécula L1-dsRNA-B.

Esta similitud con L1-dsRNA-B se inicia desde la base 1917 y es 77 pb río arriba del codón de inicio del ORF para la RNA polimerasa, por tanto, las características de esta son las mismas a lo descrito para L1-dsRNA-B, con la excepción que el marco de lectura abierto está en +2. Esta molécula presenta un 3' UTR de 513 pb y un 5' UTR de al menos 98 pb y un 45% de contenido GC. El alineamiento comparado con la base de datos nuevamente mostró identidad con la proteína de la capsida del virus L-A de *S. cerevisiae*, en el marco de lectura +2, esta vez con valores de E-value mayores de 0,89.

Mediante la aplicación bioinformática FS Finder 2 (Moon y col., 2004) se buscó una aproximación para la presencia de los elementos para un cambio en el marco de lectura. Aún cuando se encuentran tres posibles combinaciones, sólo una se encuentra previa al codón de inicio de RdRp. El sitio resbaladizo se encuentra a partir de la base 1836, compuesto por tres guaninas, tres uracilos y una timina, y posee la secuencia: G GGU UUT. Seis, nueve y diez nucleótidos río abajo, en la base 1849, 1852 y 1853 respectivamente se encuentran las secuencias GCAG, GGTC, GTCA que podrían formar un pseudoknot con una región 57, 52 y 45 nucleótidos más adelante, a partir de la base 1910. Esta zona regulatoria se encuentra entre los dos ORFs propuestos, pero a diferencia de la anterior, ambos ORFs están en igual marco, por tanto no existe una sobreposición de los mismos (Fig. 11).

5.3.3. L2-dsRNA

En el caso de la molécula L2 se logró secuenciar 3902 pb (Fig 12). El alineamiento fue comparado con la base de datos Genebank el cual evidenció identidad, con valores de E de 2×10^{-98} , secuencias asociadas a RNAs polimerasas

M

1 AAATTCATT GAAAGTATTG TTAAGACAGA CTACGAGGGA TCTTCCAGC CTTTGTTCAA GGACGGCAGA TTCACTCTGA TCAACAACA GACTGCAATG
 -2 Y D M L G R Q Y D Y N S E L A S N F N Y C G K K P E V L H N Y G E A

101 GTAGACATGT TGGGAGACAA ATATGATGTC AACTCTGAGT TAGCTTCGAA CTTCAACGTG TGTGGTAAAA AACCGAAGT TTTACACAAAC GTTGGTGAAG
 -2 A D F S G L N R K Y F D D T N T Y S P L A A I N E F T K F Y P G Y R

201 CGGATTTTTC GGGACTCAAC AGGAAGTACT TTGATGACAC CAATACCTAC TCTCCGTTGG CAGCCATAAA CGAGTTCAC TAAATTCGTC CGTGGTGAAG
 -2 R L T K G E V D A F A K S Q I Y E D S Q E A Y I L N M L I S V F K A

301 ACTTACTAAG GGCAGAGTGG ACGCTTTTGC AAAAAGTCAA ATCTATGAGG ACTCCCAAGA GGCCTACATC TTGAATATGC TTATTTGCTG GTTTAAAGCT
 -2 L L Y E D T M S K D N V L H V K Q S G Y Q D S H V K A E Y G G V T G

401 CTGTTSTATG AGGCACTAA CTCTAAAGAC AATGACTGTC ACGTGAAGCA AAGCGGATAC CAAGACTGTC ACGTCAAAGC GGAATATGGA GGTGTAAACG
 -2 G D Y Y H E I N M G P P V D D M P E V M M M V R S K D N Y V D R P Y

501 GAGATGTTGT CCACGAAATA AACATGGGGC CTCGSGTTGA TACATGCCCC GAAGTAATGA TGATGGTGGG AAGTAAGGAC AATTATTGGG ACAGACCATA
 -2 Y Y L R F D N R S S A Q Y T F Y I T H C F G R D G T S P L N V D I H

601 CGTGCTAAGA TTGATAAATA GGTCTTCGGC TCAATACACT TTCTACATAA CTCATTGTTT TGGCAGGSAI GGCACATCTC CGCTTAACTG GGACATTCAC
 -2 I P S Y D F D Q M L F E F A S G A M R S I T D P A A L P V C K S G T

701 ATACCGASTG TAGACTTCGA CCAGATGTTA TTCGAACCGC CTCGCGGGC TATGAGATCA ATCACCAGCC CGCGGGCACT CCCTTGGTGC AAATCCGGCA
 -2 T L F G V I K D Y V T L N R V E R A F S A A F E T L T A I A F T P M

801 CCTTATTCGG CTGGATCAA GACTAGCTTA CTCCTAACAG AGTAAAGCGG GCTTTTATG CAGCAITTTGA GACTTTGACT GCTATAGCAT TCACTCCAA
 -2 M P S Y Q E S L Y V D K A I T Q Y V I A K F A P C R A K I P S N L E

901 GCCATCTTAT CAAGAGASTT TGGTGGGGA CAAAGCGATA ACGCAAGTGG TGATTGCAA AFTTGTCCG TGTAGAGCTA AAATACCTTC TAACTTGGAA
 -2 G E A M Y T D M D A H D F V L D E T K T P R K A L F S G A M I N Y L

1001 GGTGAGGCTA TGGTACAGA CATGGATGCG CACGATTTG TACTGGACGA GACTAAAAC CCGAGGAAAG CACTGTTTTC TGGAGCCATG ATTAACATC
 -2 L A F M G L H A I L S N Y A S R H E N V R S A F L H S H E E L A I L

1101 TAGCTTTTAT GGGCCTACAT GCGATACCTT CAAATTACGC TTCTAGGCAC GAAAATTGGA GATCAGCCTT TTTACACAGC CACGAAGAAC TTGCTATACT
 -2 L H D R T Y R A A L T S V I T G K E M Y T F M N P N L F Y S Y D V T

1201 TCACGATAGG ACCGTAAGGG CTCGATTAC CAGTGTGATA ACTGGCAAAG AAATGGTAC TTTGATGAA CCAATCTCT TTGTATCTTA TGAATGACT
 -2 P M Y N V S K V S F E E Y L E R G Y P S S L P V H G V V P M Y S G S

1301 CCGATGGTCA AGTAAGTAA AGTAAGCTTC GAAGAGTGT TAGAGAGGGG ATACCCGTCA TCTTTGCCG TACACGGCGT GGTGCCAATG GTGTCAGGAT
 -2 S L F L S A N A S D V A S Q C H L E Q G M K V T V D E Y G T V S P E

1401 CCTCTCTCTT GTCCGCTAAC GCAICAGAGC TTGCTCACA GTGTCACTT GAACAGGGA TGAAGGTCAC GGTAGATGAA TATGGCATG TGAGCCAGA
 -2 E D A L R Y A Q M Y R M F G H E L E I R S E K T E E I H E L F A P Y

1501 AGATGGCTT AGGGTGGCTC AATGTATAG AATGTTTGT CACGAACTTG AAATAAGATC GGAAAAGACT GAGGAATAC ATGAATGTT CGCTCCGGT
 -2 Q E C V I Y P S A L L Y N T R D T D R L K L V S S L R R P G R S S T

1601 CAAGAGTGGC TCATTIACCC TTGCGACTA CTTTACAACA CAGGAGACAG AGATAGATA AAGCTAGTGT CGTCACTGCG GAGACTGGC CGCAGTAGCA
 -2 T I P D Y S A L T A G Q T I T I D Y T I P R I G M H Y F N K R V L P

1701 CTATACAGA CGTTTGGGCT CTGACCGCTG GACAAACAAT CACCATGAT TACACGATAC CTAGGATAGG CATGCACTAC TTAACAACA GGGTGTCC
 -2 P I A P S M V L P Q R K R E I T F R Y K G H T Y L H K V K M N A S N

1801 CATAGCTCCG TCAATGGTAT TACCTCAAG GAAAGGGAA ATAAGTTC GGGTCAAAG CCATACAGT TCCATAAGG TCAAAATGAA CGCTTCAAC
 -2 V T R P V K D F H G G E I Q Y A P L L P Y L R N A R V P I Q V N Q T

1901 GTGACAAGC CTGTGAAGG TTTTACCGA GCGAAATAC AGTGGCAC TTTGTTGCC GTTCTAAGG ATGCTCGTGT GCCAATCAA GTCAATCAGA
 -2 T S E E D V T A E D V E

2001 CTTCCGAGGA GGCAGTACT GCGAAGATG TAGAGTAGA CTATAAACA TCCAAATACA CCGGAAGTGA CACATCGGAC TATCTATCGC AATACGACA

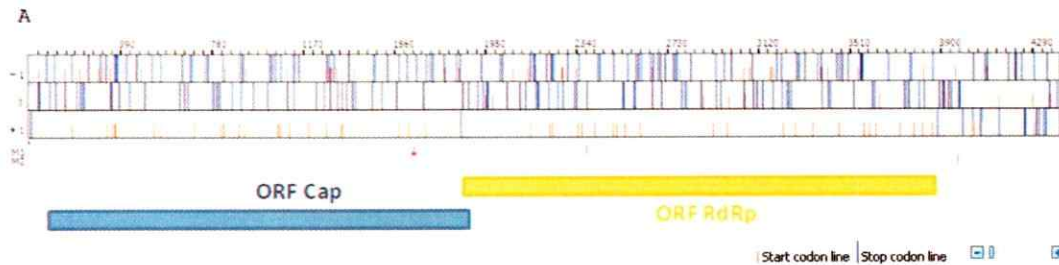
V Q P C E G E E S T G N R V E F A D R Q I P

```

2101 ACTATATAAC AGGAATTCAT TCATAGGGCT TAACGTGCAG CCATGTGAAG GTGAGGAAAG CACGGGCAAC AGATGGGAGT TTGCCGACAG ACAAAATCCC
-2 L T I S L T S R G N I E A V P F S E A Q Y A F L D R E S S N K F G I
2201 TTAACGATAT CTCTAACCAAG CCGGGGAAAC AITGAAGCAG TACCGTITTC AGAAGCTCAG TATGCCITCC TGGACCGGGA AAGCTCAAAI AAAITGGGCA
-2 I F E A S K S Y Y G V T M S G L V I A N G K A K F F F A S I K Q Y L
2301 TATTCGAGGC AAGTAAAAGC GTGTATGGGG TAACAATGTC TGGGTTGGTG ATAGCTAATG GTAAGGCTAA ATTTTCTTC GCCTCAATCA AGCAATATTI
-2 L L T Q S K A M L I A M T R H F S S L Y E S Y A Y D D P T S P V N L
2401 GCTAACCCAA TCTAAAGCAA TGTTAATAGC CATGACCAGA CACTTCTCTA GCCTGTACGA GAGCTACGCT TATGATGATC CTACGTACCC GGTAAACCTA
-2 F R D R F P G R S E N E A M T F K Q M E E L P K S K I S G A H H F H
2501 TTCAGAGACA GATTCGCCGG TAGAAGCGAA AATGAAGCGA TGACATTCNA GCAAATGGAG GAGCTGCCGA AGTCTAAAAT TTCAGGAGCG CACCATTTTC
-2 H F T P M Q V L K S I G R G R L E R A I Q A S R L P K D A E M T F A
2601 ACTTCACACC TATGCAAGTG CTGAATCGA TAGGGAGAGG GAGACTGAA AGGGGATCC AGGCATCAGC GCTACCAAAG GAGCGCGAAA TGACGTTTGC
-2 A T G M L L V Y T S I S D Q M A E L I R G A G I F K C D S V R E Y V
2701 CACAGGATG CTACTGTGGT ACACATCTAT TAGTGATCAG ATGSCAGAAC TAATCAGAGG CCGCGGAATA TTCAGTGTG ATAGTGTGAG GGAATACGTA
-2 K E M K Q L S V E A K S L Q N L Y T D D L R T V F E L E V L V N R I
2801 AAAGAAATGA AACAGCTATC AGTGGAGGCG AAGTCGCTGC AAAACCTCGT GACCGATGAC CTAAGAACAG TATTCGAGT AGAGTGTGTC GTAATATGAA
-2 I D G V V D V E K E K E N R Q S V N Y T N I K D S D V F R S A C K I
2901 TCGACGGTGT AGTAGACTGG GAGAAAGAGA AGGAAAACAG ACAGTCASTA AACGTGACAA ACATAAAGGA TTCAGATGTG TTTAGATCTG CGTGTAGAT
-2 I F E D A K G I G R R P K S F S V E S Y V A N R V Q V S A A G S I H
3001 AITGAGAGC GCGAAGGTA TAGGCCGGCG CCGAAGAGC TTCAGTGGG AGAGCTAATG GGCTAACAGG TGGCAATGGA GTGCTGGGG CAGCATACAC
-2 S Q Y P R D M E Y V I R D Q Q S L K N K F I T I S N M P K C T V D Y
3101 TCACAATATC CAGAGACAT GAGTACGTA ATTAGGACCC AACACTCACT TAAAAACAG TTTATACTA TATCCAATAT GCGCAATGT ACAGTAGACT
-2 Y F S D R E P Q Y Q G V S S T K Y E V G K Q R A I Y G T D L T S Y V
3201 ACTTTTCGGA CAGGAGCGCG CAGGTGCAAG GTTGGTCTC AACGAAGTAT GAATGGGGCA AGCAAAGAGC AATATATGGG ACAGATCTGA CTAGCTAGT
-2 V L S N F A F Y N C E N V L P N Q F P Y G R D A N D E N V V N R Y S
3301 ACTGTCTAAT TTGCGTTTT ACATTTGTA GAAGCTACTA CCAATCAAT TTCCAGTGG AGGATGTC AATGACGAAA ACCTGGTGA CAGAGTTTCG
-2 G V L N N R M P F C L D F E D F M S Q H S S G N M K A V I Y A Y I E
3401 GAGTGTCTAA ATACAGAAAT GCGTTTCTGT CTTGACTTTG AGGACTTCAA TAGTCAACAC AGCTCGGGCA ACATGAAGGC TGTAAATATC GCTTATATCG
-2 E T F I D C L T P E Q E Q A A M V T A A S L D K Q I I N D N Y G T K
3501 AGACTTTTCAI AGATTTGCTT ACTCGGAAAC AAGAGCAAAC GGCAATGTGG ACCGACGCT CATTAGACAA ACAGATAATC AATGACAACG TTGSAACCAA
-2 K T T Y E S K G T L L S G V R L T T F M N S V L N Y I Y T T K L A A
3601 GACTACGTAT GAATCTAAGG GAACGCTACT TTGCGGGTGG AGGCTGACGA CTTTCATGAA CAGTGTGTTG AACTATATTT ATACAAGAA ATTAGCAGCA
-2 E E K R P G D S L H N G D D Y L I G Y R S M A L P Q R C M Q N A I K
3701 GAAGAAAAGC GTCCGGTGA TAGCTTGCAC AACGAGAGAG ATGTCTTAAT AGGAGTGGG TCAATGGCTC TGCCGACAGG GTGTATGCAA AAGCCATTA
-2 K Y N V R M Q S S K C A V G A I A E F L R I D H K Q G G N G Q Y L S
3801 AATATACTGT GCGAATGCAA AGCAGCAAAI GCGCGTGGG TGCAATTTGCA GAATTCCTGC GGATAGROCA CAAACAGGGA GCGCAACGGC AGTACCTATC
-2 S R A V A T M Y H S R I E S R V S T D I R D L V Q S M E N R F A D A
3901 CAGAGCAGTG GCGCAATGG TACATTGAG AAITGAAAGT AGAGTCTGCA CTGACATAAG AGACTTAGTA CAATCAATGG AAAACAGATT CGCTGACGCA
-2 K G R G M A N D I I S G L R E Q Y Y M R Q S V L
4001 AAAGGCAGAG GAGTGGCGAA TGACATCATA TCGGATTTGC GGGAAACSTA CTACATGAGG CAGAGTGTGC TGTGACCCG ACGTAGAGGA CATATACTTG
4101 AATAAGAAAT GCGACAGAGT TGTAGSAGGA ATAGCGAGAG AAAAGGACTC GAAAATGGGT GTTTTGTATA CATCCCAACT TAGAGCGCAG AAGAAACGTG
4201 GCAATACCTT ACTTGGCGGG AGTAAATGAA TATGCAACAG AAATACACAA AGCATTAATA AATAATGTGT CAATTAAGAC TATATGTGAC AGATATAGG
4301 ACGCTACATA CGAAGCTGTC TCAATCAAGS ATCGCAAGAT GAAAACTTA CCGGAAAACA GAGATCAGTG GTTCGTTAAC GTACGACGAA TATACAAAGC
4401 ACATAAGGAT AGTCAATTTT CACAGAAATTA CCGTAAAGCA GCCTTAGTAG GATTCGCATT GGAAGTGTG GCGAGAGAAA TGCCTGATGC AACCAIACA
4501 ACAATACTGA ATACTTCACA AAGGCCACTA GACCTAATTA AGCACATATT AATAATGTGC ACAAAGACC AGCAGGGTCC ATAACAACA TC

```

Figura 10: Secuencia nucleotídica y traducida de la molécula L1 dsRNA-A. En azul se muestra la secuencia aminoacídica para los ORFs propuestos.



B

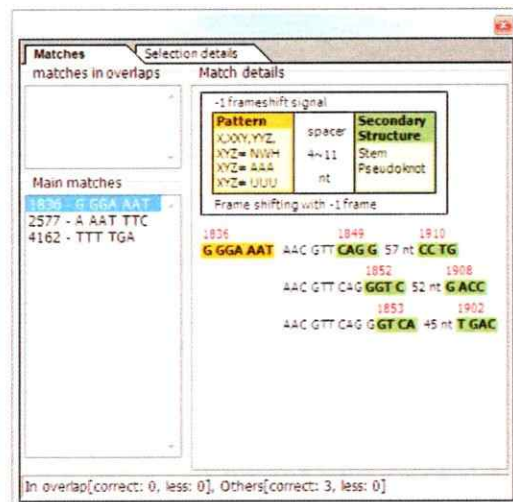


Figura 11: Sitio de corrimiento del marco de lectura en la molécula L-A 385 descrito por FS Finder. (A) Identificación de los dos ORFs descritos. (B) Estructura predicha para el corrimiento en el marco de lectura. En amarillo se muestra el sitio resbaladizo y en verde las bases nucleotídicas involucradas en la formación de la estructura secundaria pseudoknot.

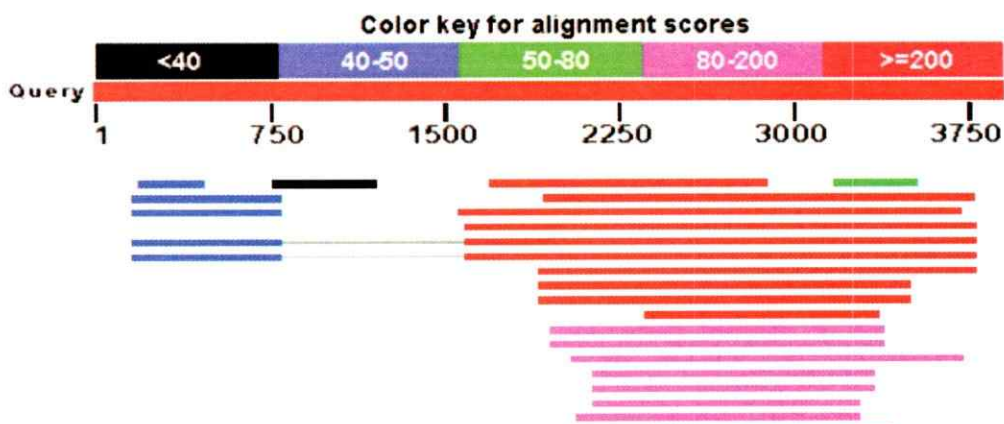
```

1 CTCAATACAC TTTCTACATA ACTCATTTGT TTGGCAGGGA TGGCACATCT CCGCTTAACG TGGACATTCA CATACCGAGT GTAGACTTGG ACCAGATGTT
101 ATTGSAACCA GCTTCCG99G CTATGAGATC AATCACC9GAC C9G9CG9CAC TCCCTTGGTG CAAATCC9GC ACCTTATTCG GCTGGATCAA AGACTACGTT
201 ACTCTTAACA GAGTAGAAGG G9CCTTTAGT GCAGCATTTG AGACTTTGAC T9CTATAGCA TTCACTCCAA TGCCATCTTA TCAAGAGAGT TTGGTGTGGG
301 ACAAGAGGAT AACGCAAGTC GTGATTGCCAA AATTTGCTCC GTGTAGAGCT AAAATACCTT CTAACCTGGG AGGTGAGGCT ATGGTGACAG ACATGGATGC
401 GCAC9ATTTT GTACT99ACG AGACTAAAAC TCCAG9GAAA GCACGTGTTT CTGGAGCCAT GATTAACAT CTASCTTTTA T99GCTACA TGGATACTT
501 TCAAAATTAG CTTCTAG9CA CGAAAATT9G AGATCAG9CT TTTTACACAG CCACGAA9AA CTTGCTATAC TTCACGATAG GACCGTA99G GCTGCATTAA
601 CCAGTGTCAAT AACT99CAAA GAAATGGTTA CTTTCATGAA CCCAAATCTC TTTGTATCTT ATGAT9TAC TCCGATGGTC AACGTAA9TA AAGTAAGCTT
701 CGAAGAGGTG TTAGAGAG9G GATACCC9TC ATCTTT9CCC GTACAC99CG T9GTGCCAAT GGTGTCA9GA TCCCTCTTCT TGTCCGCTAA CGCATCAGAC
801 GTTGCCTCAC AGTGTCACTT T9AACAG99C ATGAAGGTCA C9GTAGATGA ATATGGCACT GTGAGCC9CA AAGAT9CGCT TAGGGTGGCT CAAATGTATA
901 GAAT9TTT9G TCAC9AACTT GAAATAAGAT C9GAAAAGAC T9AG9AAATA CATGAACTGT TCGCTCC9GT G9AAGAGT9C GTCAATTACC CTTGG9CACT
1001 ACTTTACAA CAGAG99ACA CASATAGATT AAAGCTAGT TCGTCACTGC G9AGACCT9G CCGCAGT9G ACTATACCAG AC9TTTC99G TCTGACCGCT
1101 GGACAAACAA TCACCATCGA TTACACGATA CCTAGGATAG GCATGCACATA CTTCAACAAA AG99T9CTCC CCATAGCTCC GTCAATGGTA TTACCTCAA
1201 G9AAAAG9GA AATAAC9TTC AG99TCAAAG GCCATACAGT TCTCCATAAG GTCAAAAT9AA C9CTTCAAAC GTGACAAGAC CTGTGAAGGA TTTTCA99GA
1301 G9CGAAATAC AGGTGG9CAC TTTGTTGCCA GTTCTAAGGA AT9CTC9TGT C9CAATCCAA GTCAATCAGA CTTCCGAGGA GGACGTAACT GCCGAAGATG
1401 TAGATAGAA CTATAAAAC TCCAAATACA C99GAAT9A CACATCG9AC TATCTAT9C AATAC9GACA ACTATATAAC AG9AATTCAT TCATAG99CT
1501 TAACGT9CAG CCATGT9AAG GTGAG9AAG CAC999CAAC AGAT99GAGT TTGCCGACAG ACAAAAT9CC TTAACGATAT CTCTAACAA CCG999AAAC
1601 ATTGAAGCAG TACC9TTTC AGAAGCTCAG TAT9CCTTC T9GAC999GA AAGCTCAAAT AAATTC99CA TAITCGAG9C AAGTAAAAGC GTGTAT999G
1701 TAACAAT9TC T999TT99TG ATAGCTAATG GTAAG9CTAA ATTTTCTTC C9CTCAATCA AGCAATATTT GCTAACCCAA TCTAAAGCAA T9TTAATAGC
1801 CATGACCAGA CACTTCTCTA C9CTGTACGA GAGCTAC9CT TATGATGATC CTACGTCAAC G9TAAACCTA TTCAGAGACA GATTC999G TAGAAG99AA
1901 AATGAAGCGA TGACATTCAA GCAAAAT99AG GAGCT9CC9A AGTCTAAAAAT TTCAGGAG9C CACCATTTTC ACTTCACACC TATGCAAGT9 CTGAAATCGA
2001 TAG99AGAGG GAGACT99AA AG99CGATCC AG9CATCAG C9TACCAAAG GAC9CC9AAA TGAC9TTT9C CACAG99AT9 CTACTGT9GT ACACATCTAT
2101 TAGTGATCAG AT99CAGAAC TAATCAGAGG C9CG99AATA T9CAAGT9TG ATAGT9TGAG G9AATAC9TA AAAGAAAT9A AACAGCTATC AGTGGAG9CG
2201 AAGT99T9C AAAACT99T GAC99AT9AC CTAAGAACAG TAITCGAGTT AGAG9T9CTC GTAATAGAA TCGAC99T9T AGT9AGACT9 GAGAAAGAGA
2301 AG9AAACAG ACAGTCA9TA AACGTACAAA ACNTAAAGGA TTCAGAT9TG TTTAGATCTG C9T9TAAAGT ATTTGAAGAC GCCAAAG9TA TAG9CC99CG
2401 CCGAAGAGC T9CAGTT99G AGAGCTATTG G9CTAACAG9 T99CAAT9GA GT9CT999G CAGCATAAC TCACAATATC CCAGAGACAT G9AGTAC9TA
2501 ATTAG99ACC AACAGTCACT TAAAAACAAG TTTATACTA TATCCAATAT G9CCAAATGT ACAGTAGACT ACTTTTC99A CAG99AG9CG CAG9T9CAAG
2601 GTT99T9CTC AACGAAGTAT GAAT9999CA AGCAAAAGAC AATATAT99G ACAGATCTGA CTAGCTAC9T ACTGTCTAAT T9CG9TTTT ACAAIT9TGA
2701 GAACGTACTA CCAAAATCAAT T9CCAGT99G AC9AGAT9CC AATGAC9AAA AC9T99T9AA CAGAGTTT9G G9AGT9CTAA ATACAGAAAT GCCTTTCTGT
2801 CTT9ACTT9G AG9ACTTCAA TAGTCAACAC AGTTC999CA ACAT9AAG9C T9TAATATAC G9TTATAT9G AGACTTTCAT AGAIT9TCT ACTCC99AAC
2901 AAG9CAAGC G9CAAT9T9G ACC9CAG9CT CATTAGACAA ACAGATAATC AATGACAC9G TTGGAAACCA GACTAC9TAT GAATCTAAAG GAAC9TACT
3001 T9C999T9G AG99TCA9GA CTTTCATGAA CAGT9T9T9 AACTATATTT ATACAA99AA ATTAGCAGCA GAAGAAAAGC GTCC999TGA TAGCTT9CAC
3101 AAC9GAG9C AT9TCTTAAT AG9AGT9AGG TCAAT999CT T9CC9CAGAG GTGTATGCAA AAC9CCATTA AATATAAC9T G9GAATGCAA AGCAGCAAAAT
3201 G9C99T99G T9CAATT9CA GAATTC9T9C G9ATAGACCA CAAACAG9GA G9CAAC99CC AGTACCTATC CAGAGCAGT9 G9CCAAAT9G TACATTC9AG
3301 AATT9AAAGT AGAGTCT9GA CTGACATAAG AGACTT9TGA CAATCAAT9G AAAACAGATT C9CTGAC9CA AAAG9CAGAG G9AT99CGAA T9CATCATTA
3401 TCC99ATT9C G9GAACAGTA CTACAT9AGG CAGAGT9T9C TGT9T9ACAC C9ACGTAGAG GACATATACT TGATAAAGAA T9CC9CACGA GTTGTAGGAG
3501 GAATAAG9GA AGAAAAGAC T9GAAAT9G GT9TTTTGAT AACATCCCAA CTTAGAG9CG AGAAGA99T G9CAATACCG TACTT9CC9G GAGTAAATGA
3601 ATAT9CAAA CAAAATACACA AAGCATTAAA AATAAAT9TG TCAATTAAGA CTATAT9TGA CAGATTATAC GAC9CTACAT ACGAAGCT9T CTCAATCAA
3701 GAT99CAAG T9AAAATCTT WC999AAAAC AGAGATCAGT G9TT9CT9TAA C9TAC9CAGA ATATACAAAG CACATAA99G TAGTCAATT9 TCACAGAAAT
3801 AC99TAA99C AGCCTT99TA G9AAT99CAT T9GAAGT9TT G99CAGAGAA AT9CCT9AT9 CAACATAAC AACAACT9G AATACCATT9 T9GCCACT9C
3901 TC

```

Figura 12: Secuencia de la molécula L2-dsRNA obtenida por el alineamiento de los clones.

dependientes de RNA de sistemas virales, en el marco de lectura +2. Se encontró mayor identidad con el virus F de *Black raspberry* y para el virus L-A de *Saccharomyces cerevisiae* nuevamente. En el marco de lectura +3 se encontró alineamiento con la proteína de la capsida del virus LA de *S. cerevisiae* (Fig 13).



Color	Acceso	Descripción	Marco	E-value
Rojo	YP_001497151.1	RdRp [Black raspberry virus F]	+2	2e-98
Rojo	AAA50321.1	RNA polimerasa [<i>Saccharomyces cerevisiae</i> L-A]	+2	1e-97
Rojo	NP_620493.1	Proteína de unión a ssRNA [<i>Saccharomyces cerevisiae</i> L-A]	+2	7e-97
Rojo	NP_620495.1	RNA polimerasa [<i>Saccharomyces cerevisiae</i> L-A]	+2	6e-96
Rojo	NP_042581.1	RNA polimerasa [<i>Saccharomyces cerevisiae</i> L-a]	+2	1e-84
Rojo	ACA61232.1	RNA polimerasa [<i>Ribes</i> virus F]	+2	6e-73
Rosa	CAJ29959.1	RdRp [<i>Amasya cherry disease-associate</i>]	+2	1e-37
Rosa	CAJ57274.1	RdRp propuesta [<i>Cherry chlorotic rusty spof</i>]	+2	2e-36
Rosa	CAJ29958.1	RdRp propuesta [<i>Amasya cherry disease-associate</i>]	+2	2e-32
Rosa	NP_620730.2	RdRp propuesta [<i>Trichomonas vaginalis</i> virus]	+2	8e-22
Rosa	AAA62868.1	RdRp propuesta [<i>Trichomonas vaginalis</i> virus 1]	+2	8e-22
Azul	AAA50320.1	Proteína de la cápside [<i>Saccharomyces cerevisiae</i> L-A]	+3	0.75
Azul	NP_620494.1	Proteína principal de la cápside [<i>Saccharomyces cerevisiae</i> L-A]	+3	0.97

Figura 13: Alineamiento de la secuencia traducida de la molécula L2-dsRNA con la base de datos de nucleótidos traducidos (blastx) de Genebank. Esquema con las identidades donde se indica en distintos colores los puntajes obtenidos. En la tabla se observa las principales identidades obtenidas mediante este alineamiento.

3.3.3 S1-dsRNA y S2-dsRNA

Con respecto a la molécula S1 existen dos ORFs que codificarían para proteínas de 212 (Fig. 14) y 225 aminoácidos (Fig. 15), denominadas S1-A-dsRNA y S1-B-dsRNA respectivamente. En tanto que la molécula S2, codificaría para dos proteínas de 117 (Fig.16) y 216 aminoácidos (Fig.17), denominadas S2-A-dsRNA y S2-B-dsRNA.

Ninguna de estas hipotética proteínas mostró una similitud significativa con alguna proteína de la base de datos NCBI.

```

      V A T M V H S R I E S R V S T
1  GAATTCCTGC GGATAGACCA CAAACAGGGA GGCACGGCC AGTACCTATC CAGAGCAGTG GCCACAATGG TACATTGGAG AATTGAAAGT AGAGTCTCGA
-| T D I R D L V Q S M E N R F A D A K G R G M A N D I I S G L R E Q Y
101 CTGACATAAG AGACTTAGTA CAATCAATGG AAAACAGATT CGCTGACSCA AAAGSCAGAG GGATGGCGAA TGACATCATA TCCGGATTGC GGGAACAGTA
-| Y Y M R Q S V L C D T D Y E D I Y L I K N A H R V Y G G I S E E K D
201 CTACATGAGG CAGAGTGTGC TGTGTGACAC CGAGGTAGAG GACATATACT TGATAAAGAA TGCCACAGA GTTGTAGGAG GAATAAGGGA AGAAAAAGAC
-| S K M G V L I T S Q L R A Q K N Y S I P Y L P G Y N E Y A N E I H K
301 TCGAAAATGG GTGTTTGGAT AACATCCCA CTTAGAGCGC AGAAGAAGCT GTCAATACCG TACTTCCCGG GAGTAAATGA ATATGCAAAC GAAATACACA
-| K A L K I N Y S I K T I C D R L Y D A T Y D A D S I K D R K M K I L
401 AAGCATTAAA AATAAATGTG TCAATTAAGA CTATATGTGA CAGATTATAC GAGGCTACAT ACGATGCTGA CTCAATCAA GATCGCAAGA TGAAAATCTT
-| L R E N R D Q V F V N Y R R I Y K A H K G S Q L S Q N Y G K A A L Y
501 ACGGAAAAC AGAGATCAGT GGTTCGTTAA CGTACSCAGA ATATACAAAG CACATAAGGG TAGTCAATTG TCACAGAAAT ACGGTAAGC AGCCTTAGTA
-| G F A L E V L G R E M P D A T I T T I L N T S Q R P L D L I K X I L
601 GGATTCGCAT TGGAAATGTT GGGCAGAGAA ATGCGTGTAT CAACCATAAC AACATACTG AATACTTCAC AAAGGCCACT AGACCTAATT AATYACTAT
-| L
701 TATAAATGTG CACAAAGAC CAGCAGGTG

```

Figura 14: Secuencia de la molécula S1-A dsRNA de 730 pb obtenida por el alineamiento de los clones. En azul se muestra la secuencia aminoacídica para los ORFs propuestos.

```

-3      M V N Y S K Y S F E E Y L E R G Y P S S L P V H G Y V P M Y
1  ATGTACTCC GATGGTCAAC GTAAGTAAAG TAAGCTTCGA AGAAGTGTTA GAGAGGGGAT ACCCGTCATC TTTGCCCCTA CACGGGCTGG TGCCAATGGT
-3  Y S G S L F L S A N A S D V A S Q C H L E Q G M K V T V D E Y G T Y
201 GTCAGGATCC CTCTTCTTGT CCGCTAACGC ATCAGACGTT GCTTCACAGT GTCACCTTGA ACAGGGCATG AAGGTCACGG TAGATGAATA TGGCACTGIG
-3  S P E D A L R Y A G M Y R M F G H E L E I R S E K T E E I H E L F A
201 AGCCGAGAAG ATGCGCTTAS GGTGGCTCAA ATGTATAGAA TGTTTGGTCA CGAACTTGAA ATAAGATCGG AAAAGACTGA GGAATACAT GAACTGTTCC
-3  A P Y Q E C V I Y P S A L L Y N T R D T D R L K L V S S L R R P G R
301 CTCCGGTCCA AGAGTGGTCC ATTACCCCTT CGGCCTACTT TTACAACACG AGGGACACAG ATAGATTAAA GCTAGTGTCC TCACTGCGGA GACCTGGCCG
-3  R S S T I P D V S A L T A G Q T I T I D Y T I P R I G M H Y F N K R
401 CAGTACACT ATACCAGAGC TTTGGCTCT GACCGTGGG CAAACAATCA CCATCGATTA CAGATACCT AGGATAGGCA TGCCTACTTT CAACAAAAGG
-3  V L P I A P S M V L P Q R K R E I T F R V K G H T V L H K V K M N A
501 GTGCTCCCCA TAGCTCCGTC AATGGTATTA CCTCAAAGGA AAAGGGAAAT AACGTTGAGG GTCAAAAGCC ATACAGTTCT CCATAAGGTC AAAATGAACG
-3  A S N V T R P V K D F H G G E I Q V A P L L P V L R N A R Y P I Q V
601 GTTCAAAGCT GACAAGACCT GTGAAGGATT TTCACGGAGG CGAAATACAG GTGGCACCTT TGTTGCCAGT TCTAAGGAAT GCTCGTGTGC CAATCCAAGT
-3  Y N Q T S E E D Y T A E D V E
701 CAATCAGACT TGGAGGAGG ACGTAACTGC CGAAGATGTA GATAGAGAA CTATAAAAAC TCCAAATACA CAGGAACTGA CACATCGGAC TATCTAT]

```

Figura 15: Secuencia de la molécula S1-B dsRNA de 797 pb obtenida por el alineamiento de los clones. En azul se muestra la secuencia aminoacídica para los ORFs propuestos.

```

1  ACTTTTAATA TTCTTGCTCC CTTTATTGCA CGGTAGCTAC CCGGCACACA GTCTGTCTGT TCGTTTGCAA ATCCCTTCCA GAATTGGCGC CGCTTTTTC
-1                                     M V L F K F
101 AATGCCACGT TCCATTTTCA ACGACTGTAA TCAGATCGTC TATTTCGTIA TCACATCCCA AGCACCTGCC CGTCAGTGA AACATGGTCT TGTTTAAATT
-1 F H C L T A V T T F S F K R L D R F I D Q V S A I H E R G L P C P V
201 TCATTGCTTC ACAGCAGTCA CTACATTGAG CTTCAAAGCT TTAGATCGCT TTATCGATCA AGTGTCTGCT ATCCATGAAA GAGGTCTTCC TTGTCCAGTT
-1 L K P Y R F Y C C K P S Q R L K A F T R H I S V S R F Y S P T M A V
301 CTAAGCCGT ACCGTTTTGT GTTTGCAAA CCGTCGCAAC GACTTAAAGC TTTTACGCGA CACATTTCTG TGTCCCGTIT CTACTCTCCC ACTATGGCAG
-1 V V Y M T P S R F T H S R D C H P S P P S V N M A I Q A I T D L R P
401 TAGTGTAAAT GACACCCCTC AGTTTACAC ACTCCAGAGA TTGCACCTT TCGCCGCGCT CGTGGACAT GGTATACAA SCCATAACGG ACCTTCTCC
-1 P V Y Y S T N V E S A V
501 CGTTGTGGTT TCCACCAACG TAGAGTCGCG ATGTTAGTAA AAGTTGOCAC CCAGAAATAC CGASTTACTG CGGACTTTGA GCGTGTACC GTCITTTGTC
601 ACCTCTATGC CTTCCAGTTT CAATCTTTTA GCTATCAGTT TCAAAGCCTC AACCACTTTA ACAACCTCAA CTGGTACCTT CTCACGTTTG YAAT

```

Figura 16: Secuencia de la molécula S2-A dsRNA de 694 pb obtenida por el alineamiento de los clones. En azul se muestra la secuencia aminoacídica para los ORFs propuestos.

```

-1                                     Y A T M V H S R I E S R Y S T
1  GAATTCCTGC GGATAGACCA CAAACAGGGA GGCACGGCC ASTACCTATC CAGAGCAGTG GCCACAATGG TACATTGGAG AATTGAAAGT AGAGTCTGGA
-1 T D I R D L Y Q S M E N R F A D A K G R G M A N D I I S G L R E Q Y
101 CTGACATAAG AGACTTAGTA CAATCAATGG AAAACAGATT CCCTGAGGCA AAAGGCAGAG GGATGGCGAA TGACATCATA TCCGGATTGC GGGACAGTA
-1 Y Y L R Q S V L C D T D V E D I Y L I K N A H R V Y G G I S E E K D
201 CTACTTGAGG CAGAGTGTGC TGTGTGACAC CGACGTAGAG GACATATACT TGATAAAGAA TGCGCACAGA GTTGTAGGAG GAATAAGCGA AGAAAAAGAC
-1 S K M G V L I T S Q L R A Q K N G A I P Y L P G V N E Y A N E I H K
301 TCGAAAATGG GTSTTTTIGAT AACATCCCAA CTTAGAGCGC AGAAGAACGG GSCAATACCG TACTTGCCGG GAGTAAATGA ATATGCAAC GAAATACACA
-1 K A L K I N Y S I K T I C D R L Y D A T Y E A V S I K D R K M K I F
401 AAGCATTAAA AATAAATGTG TCAATTAAGA CTATATGTGA CAGATTATAC GAGGCTACAT ACGAAGCTGT CTCAATCAA GATCGCAAGA TGAATACTTT
-1 F R E N R D Q V F Y N Y R S I Y K A H K G S Q L S Q N Y G K A A L V
501 TCGGGAGAAC AGAGATCAAT GGTTCGTTAA CGTAGCGAGT ATATACAAG CACATAAGGG TAGTCAATTG TCACAGAATT ACGGTAAAGC AGCCTTAGTA
-1 G F A L E V L G R E M P D A T I T T I L N T S Q R P L D L I K Y I L
601 GGATTCGCAT TGGAAATGTT GGGCAGAGAA ATGCGTGTG CAACCATAC RACAACTG AATACTTAC AAAGGCCACT AGACCTAATT AAGTACATAT
-1 L
701 TATAAATGTG CACAAAGAC CAGCAGGGTC CATAACAAA

```

Figura 17: Secuencia de la molécula S2-B dsRNA de 739 pb obtenida por el alineamiento de los clones. En azul se muestra la secuencia aminoacídica para los ORFs propuestos.

6. DISCUSION

6.1. Clonación de las moléculas de dsRNA

6.1.1. dsRNA de mayor tamaño

Los RNA de doble hebra se han convertido en una fuente muy importante de estudios moleculares, especialmente por la determinación de secuencia de los nuevos RNA, ya que ésta no ha estado exenta de limitaciones y problemas. Variadas técnicas se han publicado para la amplificación y clonación de estos dsRNA (Attoui y col., 2000; Faridani y col., 2008; Lambden y col., 1992; Maan y col., 2007; Potgieter y col., 2002; Potgieter y col., 2009; Tzanetakis y col., 2005; Vreede y col., 1998). Una de las estrategias utilizadas corresponde a la amplificación mediante RT-PCR usando hexaoligonucleótidos de secuencia aleatoria. Esta técnica permitió obtener los primeros clones de las moléculas de mayor tamaño y a partir de estos generar partidores de secuencia conocida para avanzar por estas moléculas, confirmando así los clones anteriores, además de obtener una mayor eficiencia al eliminar la aleatoriedad del método. Además el hecho de realizar estos RT-PCR permitió identificar clones que no alineaban con los otros, y de esta manera se pudo dilucidar la presencia de dos moléculas L1, de tamaño similar y que por lo tanto no se resuelven en los geles de agarosa. Casos de co-infección de dos distintos totivirus se han descrito en *Chalara elegans* y *Sphaeropsis sapiena*, en donde los genomas lineales de los dsRNA de mayor tamaño, poseen tañamos similares (Park y col., 2005; Preisig y col., 1998).

Sin embargo fue necesario modificar una segunda técnica para poder ampliar el conocimiento hacia los extremos de cada molécula. Las técnicas de SPAT (Single-Primer Amplification Technique) y FLAC (Full-Length Amplification of cDNAs) han sido una poderosa manera de obtener una amplificación completa de estas moléculas, pero presentan la limitación de obtener resultados en dsRNA de hasta 3 kb. En el caso de

las moléculas L1-dsRNA y L2-dsRNA los oligonucleótidos utilizados como adaptadores para estas técnicas, denominados Primer 2 y PC1, hibridaban con segmentos de secuencia conocida de las moléculas, por tanto el análisis de la secuencia de los clones obtenidos por tales técnicas, resultó en encontrar clones que eran localizados en posiciones ya conocidas del alineamiento de las moléculas. Así, fue necesario diseñar un nuevo oligonucleótido adaptador, denominado NBA1 fosforilado en el extremo 3' y bloqueado en el extremo 5' con un NH₂, que además de diferir lo suficiente en secuencia con las moléculas de dsRNA, tenía una temperatura de hibridación mayor, de tal manera de aumentar la especificidad de la reacción. Para evidenciar que el adaptador era funcional se probó ligarlo en una molécula de DNA de secuencia conocida y la amplificación de esta resultó exitosa. De esta forma, mediante el reverso complementario de NBA1 y un partidor de secuencia conocida y divergente en el extremo de la molécula, se logró conocer la secuencia hacia el 3' de L1-dsRNA. Es posible que la dificultad en conocer tanto el extremo 5' de ambas L1-dsRNA, como las secuencias extremas de L2-dsRNA radique en la presencia de estructuras altamente estructuradas en estos extremos, con una difícil desnaturación mediante el protocolo utilizado.

6.1.2 dsRNA de menor tamaño

En el caso de S1 y S2 dsRNA, la técnica de RT-PCR con hexaoligonucleótidos de secuencia aleatoria solo permitió obtener clones muy pequeños, que no permitían ser amplificados nuevamente. Al ser estas moléculas menores de 1 Kb se utilizó una modificación al Método de Amplificación por Partidor Único, y se obtuvo clones que en sus extremos presentaban la secuencia correspondiente al adaptador, lo cual se confirmó al realizar RT-PCR con partidores sintetizados según la secuencia conocida, permitiendo así conocer el 100% de la secuencia de estas moléculas.

6.2. Análisis Bioinformático

El alineamiento de los clones obtenidos permitió conocer sobre el 95% de la secuencia de las moléculas de mayor tamaño, y el 100% de las de menor tamaño. Con este conocimiento, se logró determinar la presencia de dos ORF en L1-dsRNA-B que codificarían para una RNA polimerasa dependiente de RNA (3'ORF) y para una proteína de la cápside (5'ORF), relacionando estas moléculas con la presencia de un sistema viral del tipo dsRNA, que en trabajos previos había sido sugerida en base a la presencia de partículas tipo virus (Castillo y Cifuentes, 1994; Pfeiffer y col., 1996) y a la secuencia parcial de la molécula L1 dsRNA (Sanhueza, 2008).

La comparación del alineamiento de las secuencias obtenidas con la base de datos, revela porcentajes de identidad con virus de RNA de doble hebra principalmente de la familia *Totiviridae* tanto para la RdRp como para la proteína de la cápside. En base a esto se realizó la comparación de la secuencia nucleotídica traducida de la RdRp con la de todas las RdRp de esta familia que se encuentran disponibles en la base de datos. En esta comparación se encontró la presencia de los ocho motivos conservados en esta proteína descritos por J. Bruenn (Bruenn, 2003). La presencia de estos motivos permite predecir la posible funcionalidad estructural de la polimerasa.

En la familia *Totiviridae* se han descrito tres tipos de mecanismos para la traducción de la polimerasa: Un salto del ribosoma para producir la proteína de fusión, un corrimiento en el marco de lectura en -1 y +1, y un acoplamiento hipotético de término-reinicio. Por tanto se procedió al análisis de la zona cercana al ORF de la misma. En L1-dsRNA-B, los ORF están en distinto marco de lectura y se logró describir la presencia de un sitio resbaladizo y un loop que permitirían un cambio en el marco de lectura -1 del ribosoma, similar al presente en *Saccharomyces cerevisiae* L-A. Sin embargo en el caso de L1-dsRNA-A los ORF predichos están en igual marco de

lectura, y las estructuras características para un corrimiento en el marco de lectura no se encuentran aledañas al ORF de la polimerasa, por tanto, en este caso se podría postular un mecanismo de término-reinicio de la traducción, similar al encontrado en *Helminthosporium victoriae* 190S.

Puesto que la secuencia de la polimerasa es fielmente conservada en las tres moléculas descrita de mayor tamaño, lo descrito anteriormente es aplicable a ellas.

En el caso de la secuencia correspondiente a la cápside, no es de extrañar la poca similitud encontrada en la base de datos, puesto que es ésta quien dará la identidad y por tanto diferencia a cada virus descrito.

En el análisis y comparación de la secuencia traducida que codificaría a las cápsides de L1-dsRNA-A, L1-dsRNA-B, y L2-dsRNA con las secuencias descritas para otro virus, no se encontró algún sitio conservado que permitiese definir alguna estructura o secuencia características de las mismas.

En el caso de las moléculas de menor tamaño, no se encontró secuencias nucleotídicas similares en la base de datos, probablemente por su pequeño tamaño es difícil que su secuencia codifique para algún producto funcional que participe en la estructura de las VLPs. Esto apoyado en trabajos anteriores por Flores (2009) y Castillo (1995) que lograron curar la cepa UCD 67-385 de una de estas moléculas, S2-dsRNA. Las cepas curadas sólo presentaron un aumento en la concentración de L1-dsRNA y poseen VLPs morfológicamente idénticas a las de la cepa parental, por tanto indicarían que esta molécula no tendría un producto final funcional estructural en las partículas. De esta forma, S1 y S2-dsRNA serían aparentemente RNAs satélites de los dsRNAs de mayor tamaño.

7. CONCLUSIONES

1-. Las moléculas de dsRNA de mayor tamaño presentes en *X. dendrorhous*, denominadas L1, son dos moléculas, de tamaño similar, que comigran en geles de agarosa.

2-. La secuencia nucleotídica de las tres moléculas de mayor tamaño, L1-A, L1-B y L2, presenta identidad con RNA polimerasa dependiente de RNAs y con proteína de la cápside de otros virus, incluyendo micovirus. La secuencia traducida de estos RNA relaciona a estas moléculas con dsRNA genómicos del género *totivirus*. La secuencia de la RpRd presenta los ocho motivos características de la familia *Totiviridae*.

3-. Entre los ORFs descritos, L1-dsRNA-B evidencia la presencia de las estructuras características para un corrimiento en el marco de lectura hacia -1 en la traducción. En cambio L1-dsRNA-A, al tener ambos ORF en igual marco de lectura, sugiere un mecanismo de término-reinicio de la traducción. Ambos mecanismos han sido descritos para otros miembros de la familia *Totiviridae*.

4-. Las moléculas de menor tamaño, S1 y S2-dsRNA, posiblemente no codifiquen para un producto funcional, que forme parte de las VLPs, por tanto estos RNAs podrían ser moléculas satélites de los RNAs de mayor tamaño.

8. BIBLIOGRAFIA

- Attoui, H., Billoir, F., Cantaloube, J. F., Biagini, P., de Micco, P., y de Lamballerie, X. 2000. Strategies for the sequence determination of viral dsRNA genomes. *J Virol Methods*. 89:147-158.
- Baeza, M., Sanhueza, M., Flores, O., Oviedo, V., Libkind, D., y Cifuentes, V. 2009. Polymorphism of viral dsRNA in *Xanthophyllomyces dendrorhous* strains isolated from different geographic areas. *Virol J*. 6:160.
- Bruenn, J. A. 1993. A closely related group of RNA-dependent RNA polymerases from double-stranded RNA viruses. *Nucleic Acids Res*. 21:5667-5669.
- Bruenn, J. A. 2003. A structural and primary sequence comparison of the viral RNA-dependent RNA polymerases. *Nucleic Acids Res*. 31:1821-1829.
- Castillo, A., y Cifuentes, V. 1994. Presence of double-stranded RNA and virus-like particles in *Phaffia rhodozyma*. *Curr Genet*. 26:364-368.
- Castillo, A. 2005. Elementos genéticos extracromosómicos en *Phaffia rhodozyma*. Tesis de Doctorado en Ciencias con mención en Biología, Facultad de Ciencias, Universidad de Chile.
- Castro, M., Kramer, K., Valdivia, L., Ortiz, S., Benavente, J., y Castillo, A. 1999. A new double-stranded RNA mycovirus from *Botrytis cinerea*. *FEMS Microbiol Lett*. 175:95-99.
- Compel, P., Papp, I., Bibo, M., Fekete, C., y Hornok, L. 1999. Genetic interrelationships and genome organization of double-stranded RNA elements of *Fusarium poae*. *Virus Genes*. 18:49-56.
- Diamond, M. E., Dowhanick, J. J., Nemeroff, M. E., Pietras, D. F., Tu, C. L., y Bruenn, J. A. 1989. Overlapping genes in a yeast double-stranded RNA virus. *J Virol*. 63:3983-3990.
- Dinman, J. D., Icho, T., y Wickner, R. B. 1991. A -1 ribosomal frameshift in a double-stranded RNA virus of yeast forms a gag-pol fusion protein. *Proc Natl Acad Sci U S A*. 88:174-178.
- Dinman, J. D., y Wickner, R. B. 1992. Ribosomal frameshifting efficiency and gag/gag-pol ratio are critical for yeast M1 double-stranded RNA virus propagation. *J Virol*. 66:3669-3676.
- Farabaugh, P. J. 1996. Programmed translational frameshifting. *Microbiol Rev*. 60:103-134.

- Faridani, O. R., McInerney, G. M., Gradin, K., y Good, L. 2008. Specific ligation to double-stranded RNA for analysis of cellular RNA::RNA interactions. *Nucleic Acids Res.* 36:e99.
- Fauquet, C.M., Mayo, M.A., Maniloff, J., Desselberger, U., y Ball, L.A. 2005. *Virus Taxonomy. Classification and Nomenclature of Viruses.* Eighth report of the International Committee on the Taxonomy of Viruses. Elsevier. Capítulo 8, The Double Stranded RNA Viruses, pág. 571.580. San Diego, California.
- Flores, O. 2009. Relación entre la presencia de virus de dsRNA y la actividad micocida en *Xanthophyllomyces dendrorhous*. Memoria de Título, Facultad de Ciencias, Universidad de Chile.
- Ghabrial, S. A. 1998. Origin, adaptation and evolutionary pathways of fungal viruses. *Virus Genes.* 16:119-131.
- Huang, S., y Ghabrial, S. A. 1996. Organization and expression of the double-stranded RNA genome of *Helminthosporium victoriae* 190S virus, a totivirus infecting a plant pathogenic filamentous fungus. *Proc Natl Acad Sci U S A.* 93:12541-12546.
- Koltin, Y., y Day, P. R. 1976. Inheritance of killer phenotypes and double-stranded RNA in *Ustilago maydis*. *Proc Natl Acad Sci U S A.* 73:594-598.
- Lambden, P. R., Cooke, S. J., Caul, E. O., y Clarke, I. N. 1992. Cloning of noncultivable human rotavirus by single primer amplification. *J Virol.* 66:1817-1822.
- Maan, S., Rao, S., Maan, N. S., Anthony, S. J., Attoui, H., Samuel, A. R., y Mertens, P. P. 2007. Rapid cDNA synthesis and sequencing techniques for the genetic study of bluetongue and other dsRNA viruses. *J Virol Methods.* 143:132-139.
- Masison, D. C., Blanc, A., Ribas, J. C., Carroll, K., Sonenberg, N., y Wickner, R. B. 1995. Decoying the cap- mRNA degradation system by a double-stranded RNA virus and poly(A)- mRNA surveillance by a yeast antiviral system. *Mol Cell Biol.* 15:2763-2771.
- Moon, S., Byun, Y., Kim, H. J., Jeong, S., y Han, K. 2004. Predicting genes expressed via -1 and +1 frameshifts. *Nucleic Acids Res.* 32:4884-4892.
- Nuss, D. L. 2005. Hypovirulence: mycoviruses at the fungal-plant interface. *Nat Rev Microbiol.* 3:632-642.
- Park, Y., James, D., y Punja, Z. K. 2005. Co-infection by two distinct totivirus-like double-stranded RNA elements in *Chalara elegans* (Thielaviopsis basicola). *Virus Res.* 109:71-85.

- Pfeiffer, I., Kucsera, J., Varga, J., Parducz, A., y Ferenczy, L. 1996. Variability and inheritance of double-stranded RNA viruses in *Phaffia rhodozyma*. *Curr Genet.* 30:294-297.
- Potgieter, A. C., Steele, A. D., y van Dijk, A. A. 2002. Cloning of complete genome sets of six dsRNA viruses using an improved cloning method for large dsRNA genes. *J Gen Virol.* 83:2215-2223.
- Potgieter, A. C., Page, N. A., Liebenberg, J., Wright, I. M., Landt, O., y van Dijk, A. A. 2009. Improved strategies for sequence-independent amplification and sequencing of viral double-stranded RNA genomes. *J Gen Virol.* 90:1423-1432.
- Preisig, O., Wingfield, B. D., y Wingfield, M. J. 1998. Coinfection of a fungal pathogen by two distinct double-stranded RNA viruses. *Virology.* 252:399-406.
- Ribas, J. C., Fujimura, T., y Wickner, R. B. 1994. A cryptic RNA-binding domain in the Pol region of the L-A double-stranded RNA virus Gag-Pol fusion protein. *J Virol.* 68:6014-6020.
- Routhier, E., y Bruenn, J. A. 1998. Functions of conserved motifs in the RNA-dependent RNA polymerase of a yeast double-stranded RNA virus. *J Virol.* 72:4427-4429.
- Sanhueza, M. 2008. Identificación de un sistema viral del tipo dsRNA presente en *Xanthophyllomyces dendrorhous*. Tesis de Magíster en Ciencias con mención en Biología, Facultad de Ciencias, Universidad de Chile.
- Scheffter, S., Widmer, G., y Patterson, J. L. 1994. Complete sequence of *Leishmania* RNA virus 1-4 and identification of conserved sequences. *Virology.* 199:479-483.
- Scheffter, S. M., Ro, Y. T., Chung, I. K., y Patterson, J. L. 1995. The complete sequence of *Leishmania* RNA virus LRV2-1, a virus of an Old World parasite strain. *Virology.* 212:84-90.
- Schmitt, M. J., y Breinig, F. 2002. The viral killer system in yeast: from molecular biology to application. *FEMS Microbiol Rev.* 26:257-276.
- Schmitt, M. J., y Breinig, F. 2006. Yeast viral killer toxins: lethality and self-protection. *Nat. Rev. Microbiology.* 4:212-221.
- Soldevila, A. I., y Ghabrial, S. A. 2000. Expression of the Totivirus *Helminthosporium victoriae* 190S virus RNA-dependent RNA polymerase from its downstream open reading frame in dicistronic constructs. *J Virol.* 74:997-1003.

- Tai, J. H., y Ip, C. F. 1995. The cDNA sequence of *Trichomonas vaginalis* virus-T1 double-stranded RNA. *Virology*. 206:773-776.
- Tzanetakis, I. E., Keller, K. E., y Martin, R. R. 2005. The use of reverse transcriptase for efficient first- and second-strand cDNA synthesis from single- and double-stranded RNA templates. *J Virol Methods*. 124:73-77.
- Tzeng, T. H., Tu, C. L., y Bruenn, J. A. 1992. Ribosomal frameshifting requires a pseudoknot in the *Saccharomyces cerevisiae* double-stranded RNA virus. *J Virol*. 66:999-1006.
- Vreede, F. T., Cloete, M., Napier, G. B., van Dijk, A. A., y Viljoen, G. J. 1998. Sequence-independent amplification and cloning of large dsRNA virus genome segments by poly(dA)-oligonucleotide ligation. *J Virol Methods*. 72:243-247.
- Wang, A. L., Yang, H. M., Shen, K. A., y Wang, C. C. 1993. Giardiavirus double-stranded RNA genome encodes a capsid polypeptide and a gag-pol-like fusion protein by a translation frameshift. *Proc Natl Acad Sci U S A*. 90:8595-8599.
- Wickner, R. B. 1996. Double-stranded RNA viruses of *Saccharomyces cerevisiae*. *Microbiol Rev*. 60:250-265.

9. ANEXO

Tabla A1: Oligonucleótidos utilizados en reacciones de transcripción reversa y PCR

Nombre Partidor	Secuencia (5' → 3')	Orientación
3.31L	CGCCTCTGATTAGTTCTGCC	Reverso
14.4 RV	CTGACTGTCTGTTTTCTTC	Reverso
13.5	GCTCGTAAATAGAATCGACG	Directo
13.5Rv	CCGAGCTGTGTTGACTATTG	Reverso
13.6RC	CCCGAGCTGTGTTGACTATT	Directo
13.7F	GCTACATACGAAGCTGTCTC	Directo
13.7R	CTTTTGCCTCAGCGAATCTG	Reverso
13.51	CAATAGTCAACACAGCTCGG	Directo
Cap167R	GCCAGCGCCGTGTTAAATTA	Directo
Cap288Fw	CTCTTGTGGGCGAAAACCTTC	Reverso
Cap288R	GAAGTTTTCGCCACAAGAG	Directo
Cap411F	AATTCTTGCTCCACT	Reverso
Cap661Rv	CTGGTCCAAACCCTATGTCT	Reverso
Cap661F	AGACATAGGGTTTGGACCAG	Reverso
Ctg17	ATTAAGGACACGCACGTTGGGT	Directo
L6.19F	TTCGCATTGGAAGTGTGGG	Directo
L6.19R	TCCGTTGTGCAAGCTATCAC	Reverso
L7.2F	GCAGAGAAATGCCTGATGCA	Directo
L7.2R	TGCATCAGGCATTTCTCTGC	Reverso
L8.1F	ACGTTGACCATCGGAGTAAC	Reverso
L8.2F	GGTATTTTAGCTCTACACGG	Reverso
L8.2R	CCGTGTAGAGCTAAAATACC	Directo
L13.3	GGCAGAACTAATCAGAGGCG	Directo
L13.4	ATTCAGAGACAGATTCCCCG	Reverso
L13.6	ATAGTCAACACAGCTCGGG	Directo
L13.1F	GCTTAGGGTGGCTCAAATGT	Directo
L13.1R	TAGTTCTGCCATCTGATCAC	Reverso
L13.SR	TAAGCGTAGCTCTCGTACAG	Directo
L13F1	GCCATACAGTTCTCCATAAG	Directo
L13F2	CTGACACATCGGACTATCTA	Directo
L17F	GCCGACAGACAAATTCCTT	Directo
L17IF	AACTCCCATCTGTTGCCCGT	Reverso
L17R	TCAGCACTTGCATAGGTGTG	Reverso
L23F	GGTGCAAGGTTGGTCGTCAA	Directo
L23R	GCCTATACCTTTGGCGTCTT	Reverso
L269R	GCAAGTTCTTCGTGGCTGTG	Reverso
L676R	TGCACCGGAGCGAACAGTTC	Reverso

L 3'final	GATGTTTGTATGGACCCTGC	Reverso
LIV R3	CCTCTGCTGATCTGAAATGTT	Directo
LIV R4	GAACATTTAGATCAGCAGAG	Reverso
LIV R.1	GCTAATAGTGGATGAGAAG	Directo
LIV R.2	AGCCGTGGCAACTGATTAC	Directo
M19 F1	GGTAATACCATTGACGGAGC	Reverso
M19 F2	GATAGGCATGCACTACTTCA	Directo
M19 IF	TTGAGCCACCCTAAGCGCAT	Reverso
M19 IR	GGCAGAATAATCAGAGGC	Directo
M19 R1	TAAGCGTAGCTCTCGTACAG	Reverso
M19 R2	ACTTCGGCAGCTCCTCCATT	Reverso
M19F	GCTCCGTCAATGGTATTACC	Directo
M19R	CTGTACGAGAGCTACGCTTA	Directo
M20 Fw	GAGCAGTGGCCACAATGGTA	Directo
M20 Rv	TACCATTGTGGCCACTGCTC	Reverso
Mot4F	GAGGACTTCAATAGTCAACA	Directo
Mot4R	GTGTTGACTATTGAAGTCCT	Reverso
Mot5F	GGGGTGGAGGCTGACGACTT	Directo
Mot5R	AAGTCGTCAGCCTCCACCCC	Reverso
Mot6F	GCACAACGGAGACGATGTCT	Directo
Mot6R	AGACATCGTCTCCGTTGTGC	Reverso
SR	TAAGCGTAGCTCTCGTACAG	Reverso
SR2	TAGTGCATGCCTATCCTAGG	Reverso
SR3	TCCTTCACAGGTCTTGTCAC	Reverso
SR4	AAGCGCATCTTCTGGGCTCA	Reverso
S1Kan4Fw	ATGTTACTCCGATGGTCAACG	Directo
S1Kan4Rv	ATAGATAGTCCGATGTGTCAG	Reverso
S1Kan4Fw.2	GGCACTGTGAGCCCAGAAGA	Directo
S1Kan4Rv.2	CCTTCACAGGTCTTGTCACG	Reverso
S2_Fw_K1	ATTCCAAACGTGAGAGGGTA	Directo
S2_Rv_K1	ACTTTTAATATTCTTGCTCCC	Reverso
S2_Fw_K5	TTTGTTATGGACCCTGCTGGT	Directo
S2_Rv_K5	GAATTCCTGCGGATAGACCAC	Reverso
S2Kan1Fw.2	GGGTGGCAACTTTTACTACCAT	Directo
S2 Kan1 Rv.2	CCCAAGCACCTGCCCGTCAGTGT	Reverso
S2 Kan5 Fw.2	GCGCTTAGGGTGGCTCAAAT	Directo
S2 Kan5 Rv.2	CTGGCAACAAAGGTGCCCCCTG	Reverso