

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 603 553**

51 Int. Cl.:

C12N 15/86 (2006.01)

C07K 14/08 (2006.01)

C12N 7/00 (2006.01)

C07K 14/005 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **23.06.2006** **E 10182214 (6)**

97 Fecha y número de publicación de la concesión europea: **31.08.2016** **EP 2369001**

54 Título: **Virus PRRS, clones infecciosos, mutantes de los mismos, y métodos de uso**

30 Prioridad:

24.06.2005 US 694021 P

45 Fecha de publicación y mención en BOPI de la
traducción de la patente:
28.02.2017

73 Titular/es:

**REGENTS OF THE UNIVERSITY OF MINNESOTA
(100.0%)
200 Oak Street SE, Suite 600
Minneapolis, Minnesota 55455, US**

72 Inventor/es:

**FAABERG, KAY S.;
HAN, JUN;
LIU, GONGPING y
WANG, YUE**

74 Agente/Representante:

DE ELZABURU MÁRQUEZ, Alberto

ES 2 603 553 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCIÓN

Virus PRRS, clones infecciosos, mutantes de los mismos, y métodos de uso

Esta solicitud reivindica el beneficio de la Solicitud Provisional de los EE.UU. nº. de Serie 60/694.021, presentada el 24 de junio de 2005.

5 Antecedentes

El virus del síndrome reproductivo y respiratorio porcino (PRRSV) es el agente causal de una enfermedad caracterizada por trastornos respiratorios en los cerdos jóvenes e insuficiencia reproductiva en las marranas (Benfield et al., *J. Vet. Diagn. Invest.*, 4:127-133 (1992); Collins et al., *J. Vet. Diagn. Invest.*, 4:117-126 (1992); Wensvoort et al., *Vet. Q.*, 13:121-130 (1991)) y es actualmente endémico en la mayoría de los países. Este síndrome fue reconocido por primera vez como una "enfermedad misteriosa del cerdo" en los Estados Unidos en 1987 y se descubrió en Europa en 1990. Las dos cepas virales prototipo (Lelystad VR-2332) difieren en la secuencia de nucleótidos aproximadamente en un 40% y representan dos genotipos distintos, a los que se hace referencia como cepas Europea (EU o Tipo 1, Lelystad; Meulenberg et al., *Virology*, 192:62-72 (1993)) y Norteamericana (NA o Tipo 2, VR-2332; Nelsen et al., *J. Virol.*, 73:270-80 (1999)) (Fang et al., *Virus Res.*, 100:229-235 (2004); Mardassi et al., *J. Gen. Virol.*, 75:681-5 (1994); Meng et al., *Arch. Virol.*, 140:745-55 (1995); Ropp et al., *J. Virol.*, 78:3684-3703 (2004)). Se ha hecho referencia también a la enfermedad como síndrome de Wabash, enfermedad misteriosa del cerdo, síndrome reproductivo y respiratorio porcino, plaga de los cerdos, síndrome de aborto epidémico y respiratorio porcino, enfermedad de aborto azul, enfermedad azul del oído, aborto azul y aborto tardío epidémico porcino. La enfermedad se caracteriza por insuficiencia reproductiva en las marranas preñadas y problemas respiratorios en los cerdos de todas las edades. La enfermedad tiene un impacto negativo importante en la industria porcina.

PRRSV es un virus de RNA envuelto de sentido positivo perteneciente a la familia *Arteriviridae* perteneciente al orden *Nidovirales* (Cavanagh, *Arch. Virol.*, 142:629:633 (1997)). EL genoma de PRRSV varía de 15,1 a 15,5 kb de longitud (Meulenberg et al., *Virology*, 192:62-72 (1993); Nelsen et al., *J. Virol.*, 73:270-80 (1999)). El primer 75% del genoma codifica la poliproteína replicasa esencial para la replicación del virus y está constituido por dos grandes marcos de lectura abiertos (ORFs) (1a y 1b) que son procesados cotraduccionalmente en proteínas más pequeñas por proteasas viralmente codificadas (Snijder et al., *J. Gen. Virol.*, 79: 961-79 (1998)). Las proteínas estructurales están codificadas por siete ORFs aguas abajo y son traducidas por una serie anidada co-terminal 3' de mRNAs subgenómicas (sgmRNA) (Meulenberg et al., *Virology*, 192:62-72 (1993); Pattnaik et al., *Cell*, 69:1011-1020 (1992)). En la cepa VR-2332, la región codificante del genoma (15411 bases) está flanqueada por regiones 5' y 3' no traducidas de 189 y 151 nucleótidos, respectivamente.

La cepa VR-2332 de PRRSV ha sido bien caracterizada en términos de su secuencia genómica completa (Pattnaik et al., *Cell*, 69:1011-1020 (1992)), la capacidad de PRRSV para producir constitutivamente especies de RNA subgenómicas defectuosas denominadas *heteróclitos* (*en latín*: formas poco comunes) (Yuan et al., *Virology*, 275:158-169 (2000)); Yuan et al., *Virus Research*, 105:75-87 (2004)), y sus propiedades de crecimiento tanto *in vitro* como *in vivo* (Murtaugh et al., *Vet. Immunol. Immunopathol.*, 102:105-349 (2004)). Adicionalmente, un clon infeccioso de este genoma de 15,4 kb de NA PRRSV ha sido producido y examinado en cuanto a su capacidad para causar enfermedad en los cerdos (pVR-HN; Nielsen et al., *J. Virol.*, 77:3702-3711 (2003)).

PRRSV continúa causando pérdidas económicas importantes en todo el mundo. Se dispone de vacunas, pero las mismas están basadas en una sola cepa de PRRSV, y existen pruebas de que las cepas de PRRSV varían a los niveles antigénico y genético. Adicionalmente, desde que el virus fue identificado en Europa y en los Estados Unidos, han continuado surgiendo nuevos fenotipos de la enfermedad.

Sumario de la invención

Informes previos habían sugerido que las delecciones y/o mutaciones de cualquier cepa del virus PRRS eran en muchos casos extremadamente perjudiciales para el crecimiento del virus. Específicamente, laboratorios individuales habían realizado mutaciones en el extremo 3' del virus, el virus resultante era inestable y volvía rápidamente a la secuencia de tipo salvaje, o crecía muy deficientemente o no crecía en absoluto (Lee et al., *Virol.*, 331:47-62 (2005); Choi et al., *J. Virol.*, 80:723-736 (2006); Lee et al., *Virolog.*, 346:238-250 (2005)). Así, en la comparación de las secuencias de nucleótidos de la cepa europea (genotipo Tipo 1) y VR-2332 (genotipo Tipo 2), no se sabía dónde realizar mutaciones en VR-2332 nsp2 que no fueran extremadamente perjudiciales. Sin embargo, la alineación de las secuencias genómicas completas de los nuevos virus PRRS Tipo 2 con VR-2332 comenzó a proporcionar ideas en cuanto a dónde podrían producirse mutantes viables. La mutagénesis de delección ulterior demostró que la región entre los aminoácidos 324-813 de nsp2 no era necesaria para crecimiento *in vitro*.

La presente invención proporciona lo siguiente:

1. Un clon infeccioso que comprende un polinucleótido infeccioso que tiene la secuencia de nucleótidos de SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12 O SEQ ID NO: 13 y una secuencia vectorial que se replica en una célula huésped procariótica, en el que dicho polinucleótido tiene una delección en la secuencia que codifica nsp2.

2. Un polinucleótido aislado que comprende un genoma del virus PRRS que comprende una delección en la secuencia que codifica nsp2, en el que dicho polinucleótido tiene la secuencia de nucleótidos de SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12 o SEQ ID NO: 13 para su uso en la inmunización de un animal contra el virus PRRS.

5 3. Uso de un polinucleótido aislado que comprende un genoma del virus PRRS que comprende una delección en la secuencia codificadora de nsp2, en el que dicho polinucleótido tiene la secuencia de nucleótidos de SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12 o SEQ ID NO: 13 para la preparación de una composición para inmunizar un animal contra el virus PRRS.

10 4. El clon infeccioso según el punto 1 o el polinucleótido para uso según el punto 2 o el uso según el punto 3, en el que el polinucleótido se replica y produce partículas víricas infecciosas cuando se introduce en una célula permisiva.

5. El clon infeccioso según el punto 1 o el polinucleótido para uso de acuerdo con el punto 2 o el uso según el punto 3, en el que un promotor de ARN polimerasa está unido operativamente al polinucleótido.

6. El clon infeccioso según el punto 1 o el polinucleótido para uso de acuerdo con el punto 2 o el uso según el punto 3, en el que el polinucleótido comprende además un polinucleótido exógeno presente en la delección.

15 7. El clon infeccioso según el punto 6 o el polinucleótido para uso de acuerdo con el punto 6 o el uso según el punto 6, en el que el polinucleótido exógeno codifica un marcador detectable.

La presente descripción describe un polinucleótido infeccioso aislado que tiene una secuencia de nucleótidos que tiene al menos 88% de identidad con SEQ ID NO:1 y una delección de al menos 39 nucleótidos consecutivos seleccionados del nucleótido 2062 hasta el nucleótido 3864 de SEQ ID NO:1. Se describe también un polinucleótido infeccioso aislado que tiene una secuencia de nucleótidos que tiene al menos 88% de identidad con SEQ ID NO:14 y una delección de al menos 39 nucleótidos consecutivos seleccionados del nucleótido 2061 hasta el nucleótido 3545 de SEQ ID NO:14. El polinucleótido puede estar presente en un vector, en una partícula de virus aislada, presente en una célula, o una combinación de los mismos. Cuando está presente en un vector, un promotor de RNA-polimerasa puede estar enlazado operativamente al polinucleótido. El polinucleótido aislado puede ser un RNA. El polinucleótido aislado puede incluir dos o más delecciones, y cada delección puede ser independientemente de al menos 37 nucleótidos consecutivos. El polinucleótido aislado puede incluir adicionalmente un polinucleótido exógeno presente en la delección, y el polinucleótido exógeno puede codificar un polipéptido, tal como un marcador detectable.

La presente descripción también describe un polinucleótido aislado que tiene una secuencia de nucleótidos que tiene al menos 88% de identidad con SEQ ID NO:1 y al menos una delección de al menos 39 nucleótidos consecutivos seleccionados del nucleótido 2062 al nucleótido 3864 de SEQ ID NO: 1, y en donde el polinucleótido se replica y produce partículas infecciosas de virus cuando se introduce en una célula. Se describe también un polinucleótido que comprende una secuencia de nucleótidos que tiene al menos 88% de identidad con la secuencia SEQ ID NO:14 y al menos una delección de al menos 39 nucleótidos consecutivos seleccionados del nucleótido 2061 al nucleótido 3545 SEQ ID NO:14, en donde el polinucleótido se replica y produce partículas infecciosas de virus cuando se introduce en una célula. El polinucleótido aislado puede estar presente en un vector, en una partícula de virus aislada, presente en una célula, o una combinación de los mismos. Cuando está presente en un vector, puede estar enlazado operativamente al polinucleótido un promotor de RNA-polimerasa. El polinucleótido aislado puede ser un RNA. El polinucleótido aislado puede incluir dos o más delecciones, y cada delección puede ser independientemente de al menos 37 nucleótidos consecutivos. El polinucleótido aislado puede incluir adicionalmente un polinucleótido exógeno presente en la delección, y el polinucleótido exógeno puede codificar un polipéptido, tal como un marcador detectable.

La presente descripción describe además un clon infeccioso que tiene un polinucleótido con una secuencia de nucleótidos que tiene% de identidad de al menos 88 a SEQ ID NO: 1 y al menos una delección de al menos 39 nucleótidos consecutivos seleccionados del nucleótido 2062 al nucleótido 3864 de la SEQ ID NO: 1. También se describe un clon infeccioso que tiene un polinucleótido con una secuencia de nucleótidos que tiene al menos un 88% de identidad con SEQ ID NO: 14 y al menos una delección de al menos 39 nucleótidos consecutivos seleccionados del nucleótido 2061 al nucleótido 3545 de SEQ ID NO: 14. El clon infeccioso puede estar presente en una célula. Un promotor de ARN polimerasa puede estar operativamente unido al polinucleótido. El clon infeccioso puede incluir 2 o más delecciones, y en el que cada delección es independientemente al menos 37 nucleótidos consecutivos. El polinucleótido aislado puede incluir además un polinucleótido exógeno presente en la delección, y el polinucleótido exógeno puede codificar un polipéptido, tal como un marcador detectable.

También se describe un polinucleótido infeccioso aislado que comprende una secuencia de nucleótidos SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, o SEQ ID NO: 13 y un polipéptido nsp2 codificado por un polinucleótido infeccioso que comprende una secuencia de nucleótidos SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12 o SEQ ID NO: 13.

Los términos "comprende" y variaciones del mismo no tienen un significado limitante en los casos en que aparecen estos términos en la descripción y las reivindicaciones. A no ser que se especifique otra cosa, "un", "uno/a", "el/la" y "al menos uno/a" se utilizan de modo intercambiable y significan uno o más de uno.

Breve descripción de las figuras

- 5 Figura 1. A. Secuencia de nucleótidos (SEQ ID NO: 1) del polinucleótido infeccioso VR-V7 (al que se hace referencia también en esta memoria como V6G7475A). B. Secuencia de nucleótidos (SEQ ID NO: 2) del polinucleótido infeccioso VR-V5. C. Secuencia de nucleótidos (SEQ ID NO: 3) del polinucleótido infeccioso VR-V5G7475A. D. Secuencia de nucleótidos (SEQ ID NO: 4) del polinucleótido infeccioso VR-V6. E. Secuencia de nucleótidos (SEQ ID NO: 5) del polinucleótido infeccioso MN184A. F. Secuencia de nucleótidos (SEQ ID NO:6) del polinucleótido infeccioso MN184B. G. Secuencia de nucleótidos (SEQ ID NO:7) del polinucleótido infeccioso nsp2 Δ 324-434. H. Secuencia de nucleótidos (SEQ ID NO:8) del polinucleótido infeccioso nsp2 Δ 543-632. I. Secuencia de nucleótidos (SEQ ID NO:9) del polinucleótido infeccioso nsp2 Δ 633-726. J. Secuencia de nucleótidos (SEQ ID NO:10) del polinucleótido infeccioso nsp2 Δ 543-726. K. Secuencia de nucleótidos (SEQ ID NO:11) del polinucleótido infeccioso nsp2 Δ 727-813. L. Secuencia de nucleótidos (SEQ ID NO:12) del polinucleótido infeccioso nsp2 Δ 324-726. M. Secuencia de nucleótidos (SEQ ID NO:13) del polinucleótido infeccioso nsp2 Δ 324-726.

- Figura 2. Ensamblaje de los clones de longitud total de la cepa VR-2332 de PRRSV. El genoma 15.4 se amplificó en 4 secciones (I-IV) que incorporaban sitios singulares de escisión por enzimas de restricción presentes en el cDNA viral (FseI, AvrII, BsrGI) o añadidos a la secuencia de PRRSV en los extremos 5' y 3' por mutagénesis de inserción (SphI, PacI respectivamente). Un promotor de polimerasa T7 y dos residuos G sin matriz y un residuo T precedían a la secuencia viral. El vector pOK12 (24) se modificó para incluir un sitio PacI y una ribozima de la hepatitis delta aguas abajo de una cola de poli-adenosina de 50 nucleótidos.

- Figura 3. Esquema de los cambios de nucleótidos de clones infecciosos o progenie de cerdos. Se presenta El diagrama de la organización del genoma de PRRSV bajo el cual se encuentran comparaciones del genoma completo. Se representan escisiones de proteínas no estructurales supuestas en la parte superior ORF1a y 1b, representadas por flechas descendentes. Los restos de firma se identifican bajo ORF1a y 1b, con flechas ascendentes que indican su situación en el genoma de PRRSV [cisteína-proteasa α y β semejante a papaína (PCP α , PCP β); cisteína-proteasa (CP); serina/proteasa 3C (SP/3CP); polimerasa (POL); rico en cisteína/histidina (C/H); helicasa (Hel); endorribonucleasa específica de poli(U) homóloga de *Xenopus laevis* (XendoU); Ivanov et al., *Proc. Natl. Acad. Sci. USA*, 101:12694-12699 (2004); Ziebuhr et al., *J. Gen. Virol.*, 81:853-879 (2000)]. Las diferencias de nucleótidos se representan por barras verticales. 1. Cepa *wt* VR2332 (U87392) comparada con la vacuna derivada de VR-2332 (Ingelvac® MLV o RespPRRS, AF066183). 2. Cepa *wt* VR-2332 comparada con pVR-V6G7475A. 3. pVR-V5 comparado con V5-1-P3 sometido a pases *in vivo* (Sw612). 4. Cepa *wt* VR-2332 comparada con Sw612. Los cambios detallados de nucleótidos se enumeran en las Tablas 4 y 5.

- Figura 4. Seroconversión de cerdos después de la infección con PRRSV. Se infectaron cerdos en crecimiento con la cepa *wt nativa* VR-2332 ($\square$), Ingelvac® MLV (x), V5-1 P3 ($\circ$) o se dejaron sin infección ($\blacksquare$). Los días indicados, se tomaron muestras de suero y se testaron por Elisa IDEXX para indicación de la seroconversión por anticuerpos anti-PRRSV para la proteína de la nucleocápsida.

- Figura 5. A. Ensayos de calvas sobre la progenie P3 (primer linaje) de todos los clones infecciosos así como la cepa *wt* VR-2332 revelaron tamaños de calva diferentes. B. La progenie de V5-1 P3 después de crecimiento en cerdos (Sw612) producía calvas similares a la cepa *wt* VR-2332.

- Figura 6. A. Los ensayos de calvas sobre la progenie P3 (segundo linaje) de todos los clones infecciosos así como de la cepa *wt* VR-2332 exhibían tamaños de calva que eran diferentes de las preparaciones de virus del primer linaje. B. Los títulos de virus P4 indican progenie de clones infecciosos que no se replicaban como la cepa *wt* VR-2332 o el virus Sw612 a pesar de tener tamaño de calva similar.

- Figura 7. A. Las progenies P3 de la cepa *wt* VR-2332 ($\blacklozenge$), Sw612 ($\blacktriangle$), pVR-HN ($\square$), pVR-V5 (x), pVR-V5G7475A (*), pVR-V6 ($\bullet$), pVR-V6G7475A ($\circ$) se examinaron simultáneamente respecto a cinética de crecimiento en un solo paso como se reseña en el Ejemplo 1. La cepa *wt* VR-2332 y los virus Sw612 se replicaban con títulos aproximadamente 10 veces mayores en todos los momentos. pVR-V6G7475A, sin cambio alguno de aminoácido respecto al virus nativo o la vacuna, producía virus que se replicaba a un título mayor en todos los momentos que todas las progenies restantes de clones infecciosos. El título final para cada preparación de virus se enumera en la tabla que se acompaña.

- Figura 8. El análisis por transferencia Northern de diferentes pases de la progenie de pVR-V6G7475A así como Sw612 y el transcrito inicial *in vitro* revela que se producen heteróclitos tan pronto como P1 y, junto con RNA genómico, son más abundantes con los pases. Sin embargo, el RNA transcrito (Tx) no contiene especies heteróclitas fácilmente detectables.

- Figura 9. A. Representación diagramática del genoma de PRRSV. Las escisiones de proteínas no estructurales supuestas se representan encima de ORF1a y 1b, representadas por flechas descendentes. Los restos de firma se identifican bajo ORF1a y 1b, indicando su situación en el genoma de PRRSV [cisteína-proteasa α y β semejantes a

papaína (PL1); cisteína-proteasa (PL2); serina/3C-proteasa (3CL); polimerasa (RdRp); helicasa (Hel); endorribonucleasa (N) específica de poli(U) homóloga de *Xenopus laevis*; Ziebuhr et al., 2000; Ivanov et al., 2004; Gorvalenya et al., 2006]. B. Diagrama esquemático de la comparación de la proteína ORF1 (replicasa) de MN184A y MN184B y procesamiento supuesto. La degeneración observada en nsp2 se incluye en la comparación. C. Diagrama esquemático de la comparación de las proteínas ORF2-7 de MN184A y MN184B.

Figura 10. Alineación de la secuencia de aminoácidos ORF5 de PRRSV divergente. Los recuadros grises oscuros indican alta conservación de aminoácidos (> 80%; entre 16 y 19 residuos son idénticos); los recuadros grises medios (> 70%; entre 12 y 15 residuos son idénticos); los grises más claros (> 40%, entre 8 y 11 residuos son idénticos) y los recuadros sin sombreado (< 40%; menos de 8 residuos son idénticos) identifican residuos menos conservados. La región marcada con trazos indica la secuencia señal supuesta, las regiones encerradas en recuadros identifican las regiones transmembranales propuestas, se indican regiones hipervariables (HV-1 y HV-2), y se identifica la orientación propuesta de la proteína en el virión en negritas cursivas. El residuo cisteína conservado que se propone para interaccionar con la proteína M se identifica por la flecha descendente (↓). Los dos sitios de glicosilación en N supuestos conservados se identifican por asteriscos y la región hipervariable 1 contiene sitios de glicosilación en N específicos/de cepa/aislados (NxS/T). Se utilizaron para comparación las secuencias siguientes de longitud total de GenBank: VR-2332 (U87392), Ingelvac MLV (AF066183), 01NP1.2 (DQ056373), PL97-1 (AY58524), PA-8 (AF176348), SP (AF184212), BJ-4 (AF331831), HN1 (AY457635), 16244B (AF046869), HB-1 (AY150312), HB-2 (AY262352), CH-1a (AY032626), P129 (AF494042), JA142 (AY424271), SDPRRS-01-08 (AY375474), EuroPRRSV (AY366625), Lelystad (M96262), IAF-93-653 (U64931), IAF-Klop (AY184209), 98-3298 (DQ306877), 98-3403 (DQ306878), 99-3584 (DQ306879).

Figura 11. Alineación de la secuencia de aminoácidos nsp1β de PRRSV divergente. La derivación de la figura y el esquema de colores se han descrito en la leyenda de la Figura 10. Los dos residuos catalíticos supuestos completamente conservados se identifican por asteriscos y los aminoácidos encuadrados identifican la conservación de la secuencia de MN184 con aislados de Tipo 1 y EAV. El sitio de escisión propuesto se identifica por la flecha descendente (↓).

Figura 12. Alineación de la secuencia de aminoácidos nsp2 de PRRSV divergente. Los residuos catalíticos de la cisteína-proteasa supuesta completamente conservada (Cys e His) se identifican por asteriscos y los aminoácidos incluidos en los recuadros significan la conservación de la secuencia de proteasas dentro de PRRSV y EAV. Los sitios de escisión propuestos se identifican por flechas compactas (↓); los sitios de escisión posibles adicionales se indican por una flecha cortada; el péptido señal, recuadro gris compacto; las regiones transmembranales (representadas en recuadros negros cortados; y los sitios potenciales de glicosilación en N, se indican por un asterisco (*). La derivación de la figura y el esquema de colores se describieron en la leyenda de la Figura 10.

Descripción detallada de realizaciones ilustrativas

La presente invención incluye clones infecciosos del virus del síndrome reproductivo y respiratorio porcino (PRRSV) VR-2332. Como se utiliza en esta memoria, el término "clon infeccioso" es un polinucleótido que tiene dos componentes; una secuencia de vector que se replica en una célula hospedadora procarionta, y un segundo polinucleótido al que se hace referencia en esta memoria como un polinucleótido infeccioso. Cuando se transcribe *in vitro* para producir un polinucleótido de RNA y se introduce en una célula permisiva, el polinucleótido infeccioso se replica (como un RNA) y produce partículas infecciosas de virus. Así, un polinucleótido infeccioso puede estar presente en un vector como un DNA, como un RNA en una partícula de virus, o como un DNA o RNA aislado. El término "polinucleótido" se refiere a una forma polímera de nucleótidos de cualquier longitud, sea ribonucleótidos o desoxinucleótidos, e incluye DNA y RNA tanto bi- como mono-catenarios. A no ser que se indique otra cosa, un polinucleótido incluye el complemento del mismo. La secuencia de nucleótidos del complemento de un polinucleótido puede ser determinada fácilmente por una persona con experiencia en la técnica. Un polinucleótido puede incluir secuencias de nucleótidos que tienen funciones diferentes, con inclusión por ejemplo de secuencias codificantes, y secuencias no codificantes tales como secuencias reguladoras y/o regiones no traducidas. Un polinucleótido puede obtenerse directamente a partir de una fuente natural, o se puede preparar con ayuda de técnicas recombinantes, enzimáticas, o químicas. Un polinucleótido puede ser de topología lineal o circular. Un polinucleótido puede ser, por ejemplo, una porción de un vector, tal como un vector de expresión o clonación, o un fragmento.

Si se produce de forma natural, un polinucleótido puede aislarse o purificarse. Un compuesto "aislado", tal como un polinucleótido, polipéptido, o partícula de virus, es uno que está separado y apartado de su ambiente natural. Un compuesto "purificado" es uno que está exento al menos en un 60%, un 75%, o un 90% de otros componentes con los cuales se encuentra asociado naturalmente el mismo. Compuestos tales como polinucleótidos y polipéptidos que se producen fuera del organismo en el cual se encuentran naturalmente, v.g., por medios químicos o recombinantes, se considera que están aislados y purificados por definición, dado que los mismos no estaban nunca presentes en un ambiente natural.

Un ejemplo de un polinucleótido infeccioso de la presente descripción incluye el polinucleótido infeccioso VR-V7 (SEQ ID NO:1). Se hace referencia a VR-V7 también en esta memoria como V6G7475A. Otros ejemplos de polinucleótidos infecciosos de la presente descripción incluyen VR-V5 (SEQ ID NO:2), VR-V5G7475A (SEQ ID NO:3) y VR-V6 (SEQ ID NO:4). Debe indicarse que si bien las SEQ ID NOs:1, 2, 3, 4, 5, 6 y otras secuencias de

nucleótidos virales se describen en esta memoria como una secuencia de DNA, la presente descripción contempla la secuencia de RNA correspondiente, así como complementos de RNA y DNA de la misma.

Otros polinucleótidos infecciosos de la presente descripción tienen una secuencia de polinucleótidos que tiene semejanza estructural con un polinucleótido de referencia. Polinucleótidos de referencia incluyen SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5 y SEQ ID NO:6, la cepa prototipo Europea del virus PRRS, Lelystad (número de acceso a GenBank M96262; SEQ ID NO: 14), y la cepa prototipo Norteamericana del virus PRRS, VR-2332 (número de acceso a GenBank U87392; SEQ ID NO: 15). Se hace referencia a la semejanza como un "identidad porcentual" y se determina por alineación de los residuos de los dos polinucleótidos (es decir, la secuencia de nucleótidos de un polinucleótido candidato infeccioso y la secuencia de nucleótidos del polinucleótido de referencia) para optimizar el número de nucleótidos idénticos a lo largo de las longitudes de sus secuencias; se permiten lagunas en una cualquiera o ambas secuencias durante la alineación, a fin de optimizar el número de nucleótidos compartidos, aunque los nucleótidos en cada secuencia deben mantenerse sin embargo en su propio orden. En algunos aspectos, la laguna (a la que se hace referencia también como una delección) está presente en la secuencia de polinucleótidos infecciosa candidato. Un polinucleótido infeccioso candidato es el polinucleótido que tiene la secuencia de nucleótidos que se compara con el polinucleótido de referencia. Un polinucleótido infeccioso candidato puede aislarse de un animal, tal como un cerdo infectado con PRRSV, aislarse de una línea de células de cultivo, o puede producirse utilizando técnicas recombinantes, o sintetizarse química o enzimáticamente. Dos secuencias de nucleótidos pueden compararse utilizando cualquiera de los algoritmos de ordenador disponibles comercialmente utilizados como herramienta de rutina para producir alineaciones de secuencias de nucleótidos. Dos secuencias de nucleótidos se pueden comparar utilizando el programa GAP del Paquete GCG Wisconsin (Accelrys, Inc.) versión 10.3 (2001). El programa GAP utiliza el algoritmo de Needleman et al., (*J. Mol. Biol.*, 48: 443-453) 1970 a fin de encontrar la alineación de dos secuencias completas que maximiza el número de coincidencias y minimiza el número de lagunas. Se pueden utilizar los valores por defecto para todos los parámetros de búsqueda GAP, con inclusión de la Matriz de Registro = NewsgapDNA.cmp, peso de laguna 50, peso de longitud = 3, coincidencia media = 10, falta de coincidencia = 0. En la comparación de dos secuencias de nucleótidos utilizando el algoritmo de búsqueda GAP, se hace referencia a la semejanza estructural como "identidad porcentual". Un polinucleótido tiene semejanza estructural con un polinucleótido de referencia de al menos 88%, al menos 89%, al menos 90%, al menos 91%, al menos 92%, al menos 93%, al menos 94%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, o al menos 99% de identidad cuando la semejanza estructural se determina utilizando el programa GAP.

Si un polinucleótido es un polinucleótido infeccioso puede determinarse por inserción en un vector de un polinucleótido infeccioso candidato, transcripción del polinucleótido infeccioso candidato *in vitro*, transfección de una célula permisiva con las moléculas de RNA resultantes, y detección del RNA viral de la progenie, la proteína de la nucleocápsida viral de la progenie, detección de partículas infecciosas de virus, o una combinación de los mismos. El vector puede tener las características de ser de número de copias bajo y se mantiene estable después de inserción de grandes insertos (v.g., de 15 kb). Un ejemplo de un vector adecuado es pOK y pOK12 (acceso a GenBank AF223639; Vieira et al., *Gene*, 100:189-194 (1991)), y otros vectores que tienen estas características son conocidos y están disponibles. En el vector, el polinucleótido infeccioso candidato se encuentra inmediatamente aguas abajo de un promotor. Promotores útiles son aquellos que pueden ser inducidos para producir niveles altos de transcripción, tales como un promotor de RNA-polimerasa T7, por ejemplo TAATACGACTCACTATA (SEQ ID NO: 16), o los promotores de RNA-polimerasa SP6 y T3. La transcripción del polinucleótido infeccioso candidato incluye típicamente digestión con endonucleasas de restricción del vector para hacerlo lineal, y producción de transcritos de RNA por el uso de rutina y métodos de transcripción *in vitro* bien conocidos. Están disponibles comercialmente kits para transcripción *in vitro* (por ejemplo, mMessage y mMachine, disponibles de Ambion, Austin, TX).

Después de transcripción *in vitro*, el RNA se purifica utilizando métodos de rutina y se utiliza luego para transfectar una célula permisiva. Ejemplos de células permisivas incluyen, por ejemplo, BHK-21 (que permite una tanda de población de partículas de virus), CL-2621, MA-104 (ATCC CRL-2378), MARC-145 (Kim et al., *Arch. Virol.*, 133:477-483 (1993)), líneas de células clonadas a partir de estas líneas de células, o macrófagos alveolares primarios de porcino. Métodos para transfección eficiente de células incluyen el uso de bromuro de 1,2-dimiristiloxipropil-3-dimetilhidroxi-etil-amonio y colesterol (DMRIE-C) y otros productos disponibles comercialmente, p.ej. DMRIE-C. Métodos para transfección eficiente de macrófagos alveolares primarios de porcino se conocen en la técnica (Groot Bramel-Verheige et al., *Viol.*, 278:380-389 (2000)). Generalmente, pueden utilizarse para transfección 2 a 3 microgramos de RNA, pero pueden utilizarse cantidades menores y mayores. Después de un periodo de tiempo adecuado, la presencia de RNA viral de la progenie puede detectarse por ejemplo, mediante la reacción en cadena de la polimerasa con transcriptasa inversa (RT-PCR). Análogamente, la proteína de la nucleocápsida viral de la progenie puede ser detectada, por ejemplo, por un anticuerpo específico de la nucleocápsida. Adicionalmente, puede detectarse si las partículas de virus producidas por células transfectadas con un nucleótido candidato infeccioso llegarán a infectar otra célula por exposición de células permisivas no infectadas al sobrenadante de las células infectadas. Opcionalmente, puede observarse efecto citopático (CPE). Un polinucleótido infeccioso candidato se considera que es un polinucleótido infeccioso cuando el mismo produce RNA viral de la progenie, proteínas virales de la progenie (nucleocápsida, membrana, GP5, y otras), e infecta otras células permisivas.

Un polinucleótido infeccioso puede incluir una delección de nucleótidos que codifican proteína 2 no estructural (nsp2), uno de varios polipéptidos (12 predichos) presentes en la poliproteína codificada por ORF1. En un virus PRRS, y polinucleótidos infecciosos del mismo, los nucleótidos que codifican el primer aminoácido de nsp2 pueden

determinarse por identificación del sitio de escisión de la proteasa 1 beta semejante a papaína, que se predice está situada después del aminoácido glicina ORF1 en la posición 383 de VR-2332.

Con respecto a la identificación de los nucleótidos que codifican el último aminoácido de nsp2, no se ha determinado empíricamente el sitio de escisión C-terminal exacto de nsp2 de la poliproteína codificada por ORF1a, por lo que los nucleótidos que corresponden al extremo 3' de la región codificante se desconocen. Sin embargo, se han propuesto dos predicciones de sitios de restricción C-terminal, una Gly|Gly (donde la línea vertical entre los dos residuos glicina indica la localización de la escisión) en el aminoácido 980 en VR-2332, y la otra en el aminoácido 1197 en VR-2332. En la alineación de todas las secuencias PRRSV disponibles, existen varios dobletes Gly|Gly totalmente conservados dentro de esta porción, que pueden ser también el sitio de escisión en el terminal C de nsp2 de la poliproteína (aminoácidos 646, 980, 1116, 1196, 1197, en VR-2332). Las localizaciones de los dobletes Gly|Gly en los otros virus y polinucleótidos infecciosos pueden identificarse por comparación de las secuencias de nsp2 y los dobletes Gly|Gly descritos en la Figura 12. Los estudios actuales sugieren que pueden existir al menos tres sitios de escisión en nsp2, correspondientes a los aminoácidos 280, 1116, 1196 o 1197.

El polipéptido nsp2 incluye un dominio de cisteína-proteasa semejante a quimotripsina altamente conservado (identificado como CP en la Figura 3 y PL2 en la Figura 9) presente en el término N, y 3-4 dominios transmembranales predichos cerca del término C de nsp2 (donde el número de dominios transmembranales varía dependiendo de la localización del sitio de restricción C-terminal). Típicamente, la delección de los nucleótidos que codifican los aminoácidos del dominio PL2 o la totalidad de los dominios transmembranales predichos da como resultado un polinucleótido que puede replicarse en células permisivas pero no producirá partículas infecciosas de virus. Así, un clon infeccioso descrito aquí no incluye típicamente la delección del dominio PL2 entero o la totalidad de los dominios transmembranales predichos.

Los nucleótidos que codifican el dominio de cisteína- proteasa semejante a quimotripsina son los nucleótidos 1474 a 1776 de VR-V7 (SEQ ID NO:1), los nucleótidos 1474 a 1776 de VR-2332 (número de acceso a GenBank U87392), y los nucleótidos 1482 a 1784 de Lelystad (número de acceso a GenBank M96262). La localización de un dominio de cisteína-proteasa semejante a quimotripsina en la secuencia de nucleótidos de otros virus PRRS puede identificarse por alineación de la secuencia de aminoácidos del polipéptido nsp2 codificado por un virus PRRS con la alineación de la secuencia de aminoácidos descrita en la Figura 12, y determinación de cuáles son los nucleótidos que codifican dichos aminoácidos que se alinean con el dominio de cisteína-proteasa semejante a quimotripsina. Alternativamente, las secuencias de aminoácidos de los polipéptidos nsp2 de otros virus PRRS pueden identificarse por alineación de la secuencia de aminoácidos del polipéptido nsp2 codificado por un virus PRRS con la secuencia de aminoácidos de los polipéptidos nsp2 producidos por otros arterivirus, tales como el virus de la arteritis de los equinos (EAV) y el virus aumentador de la lactato-deshidrogenasa (LDV).

Los nucleótidos que codifican los dominios transmembranales predichos de VR-V7 (SEQ ID NO:1), VR-2332 (número de acceso a GenBank U87392) y Lelystad (número de acceso a GenBank M96262) se muestran en la Tabla 1.

Tabla 1. Nucleótidos nsp2 que codifican dominios transmembranales predichos

	VR-V7	VR-2332	Lelystad
Dominio transmembranal I	881 a 901	881 a 901	761 a 781
Dominio transmembranal II	913 a 934	913 a 934	793 a 814
Dominio transmembranal III	963 a 980	963 a 980	843 a 860
Dominio transmembranal IV	985 a 1003	985 a 1003	865 a 883

La localización de los dominios transmembranales en la secuencia de nucleótidos de otros virus PRRS puede identificarse por alineación de la secuencia de aminoácidos del polipéptido nsp2 codificado por un virus PRRS con la alineación de la secuencia de aminoácidos descrita en la Figura 12, y determinación de qué nucleótidos codifican aquellos aminoácidos que están alineados con los dominios transmembranales. Alternativamente, la localización de los dominios transmembranales puede identificarse con un algoritmo de computadora, tal como el algoritmo

PredictProtein como ha sido descrito por Rost et al. (*Nucleic Acids Res.*, 32 (publicación del servidor de la Web): W321-326 (2004), o el algoritmo TMHMM descrito por Krogh et al. (*J. Mol. Biol.*, 305:567-580 (2001)) y disponible en internet.

- La delección presente en los polinucleótidos infecciosos aquí descritos puede hallarse típicamente entre los nucleótidos que codifican el dominio de cisteína-proteasa semejante a quimotripsina y los nucleótidos que codifican los dominios transmembranales, y no da como resultado un cambio de marco en el marco de lectura de ORF1. Como se ha expuesto anteriormente, la delección no incluye típicamente todos los nucleótidos que codifican el dominio de cisteína-proteasa semejante a quimotripsina, todos los nucleótidos que codifican los dominios transmembranales, o la combinación de los mismos. En algunos aspectos, por ejemplo cuando el polinucleótido infeccioso tiene semejanza estructural con SEQ ID NO:1, el límite 5' de una delección se encuentra en el nucleótido 2305, el nucleótido 2205, el nucleótido 2105, o el nucleótido 2062, y el límite 3' de una delección se encuentra en el nucleótido 3774, el nucleótido 3804, el nucleótido 3834, o el nucleótido 3864. En otros aspectos, por ejemplo cuando el polinucleótido infeccioso tiene semejanza estructural con SEQ ID NO:14, el límite 5' de una delección se encuentra en el nucleótido 2304, el nucleótido 2204, el nucleótido 2104, o el nucleótido 2061, y el límite 3' de una delección se encuentra en el nucleótido 3455, el nucleótido 3495, el nucleótido 3525, o el nucleótido 3545. La delección puede ser de al menos 39 nucleótidos, 48 nucleótidos, o 57 nucleótidos. En algunos aspectos, la delección puede ser al menos 267 nucleótidos, al menos 276 nucleótidos, o al menos 285 nucleótidos. En algunos aspectos, la delección no es > 489 nucleótidos, no > 459, no > 429, o no > 402 nucleótidos. Un polinucleótido infeccioso puede tener más de una delección en la región nsp2.
- Ejemplos de polinucleótidos infecciosos derivados de VR-V7 y que contienen una delección se describen en la Tabla 2.

Tabla 2. Polinucleótidos infecciosos derivados de VR-V7 (SEQ ID NO:1).

Polinucleótido*	Nucleótidos delecionados de SEQ ID NO:1	Aminoácidos delecionados de ORF1	Títulos virales (PFU/ml)	Sumario de fenotipo**
nsp2 Δ0180-323	1876-2304	563-705	-	no viable
nsp2 Δ242-323	2056-2304	623-705	-	no viable
nsp2 Δ324-434	2305-2637	706-816	+ (~10 ⁵)	tamaño de calva pequeño
nsp2 Δ324-523	2305-2904	706-905	+ (~10 ⁵ -10 ⁶)	intermedio
nsp2 Δ543-632	2962-3231	925-1014	+ (~10 ⁵)	tamaño de calva pequeño
nsp2 Δ633-726	3232-3513	1015-1108	+ (~10 ⁵)	tamaño de calva pequeño
nsp2 Δ543-726	2962-3513	925-1108	+ (~10 ⁵)	tamaño de calva pequeño
nsp2 Δ727-813	3514-3774	1109-1195	+ (~10 ⁵)	tamaño de calva pequeño
nsp2 Δ324-726	2305-3513	706 -1108	+ (~10 ¹⁻²)	ND
nsp2 Δ324-813	2305-3774	706-1195	-	no viable
nsp2 Δ727-845	3514-3870	1109-1227	-	no viable
nsp2 Δ324-845	2305-3870	706-1227	-	no viable

* La delección se refiere a los aminoácidos de nsp2 que se delecionan, v.g., en el virus nsp2 Δ 180-323, se han delecionado los aminoácidos 180-323 de nsp2.

** El tamaño de la calva se refiere a las calvas producidas por VR-2332 de tipo salvaje.

Un polinucleótido infeccioso que contiene una delección puede incluir un polinucleótido exógeno insertado en el lugar de la delección. Un polipéptido "exógeno" hace referencia a una secuencia nucleotídica extraña, es decir, una secuencia de nucleótido que no está presente normalmente en un virus PRRS o un clon infeccioso del mismo. El polinucleótido exógeno puede incluir, y preferiblemente incluye de hecho un polipéptido. Polinucleótidos exógenos adecuados incluyen aquéllos que codifican un marcador detectable, v.g., una molécula que es detectada fácilmente por diversos métodos. Ejemplos incluyen polipéptidos fluorescentes (v.g., proteínas fluorescentes verdes, amarillas, azules, o rojas), luciferasa, cloranfenicol-acetil-transferasa, y otras moléculas (tales como c-myc, flag, 6xhis, el péptido HisGln (HQ) de fijación de metales, y el epítipo V5) detectables por su fluorescencia, actividad enzimática o propiedades inmunológicas, y son típicamente útiles cuando se detectan en una célula, por ejemplo, una célula cultivada, o una muestra de tejido que se ha extraído de un animal. Otros polinucleótidos exógenos que pueden utilizarse son aquéllos que codifican polipéptidos expresados por otras entidades, tales como células y patógenos. La expresión de un polinucleótido exógeno da como resultado un polinucleótido infeccioso que expresa antígenos extraños. Ejemplos de secuencias de nucleótidos exógenos incluyen aquéllas que codifican proteínas expresadas por patógenos, p.ej. patógenos de porcino, tales como el circovirus Tipo 2 porcino, *Mycoplasma hyopneumoniae* (v.g., las proteínas P46 y P65 de *M. hyopneumoniae*, *Lawsonia intracellularis* (v.g. las proteínas de la membrana exterior de *L. intracellularis*), el ORF5 de diferentes cepas de PRRSV, y *Streptococcus suis* (v.g., la proteína de 38 kDa de *S. suis*). El polipéptido nsp2 tiene epítipes de células B y se espera sea inmunógeno. La inclusión de epítipes extraños en un polipéptido nsp2 se espera que dé como resultado una respuesta inmune a los epítipes extraños. Ejemplos adicionales de polinucleótidos exógenos incluyen aquéllos que codifican modificadores de la respuesta biológica, tales como, por ejemplo, IFN- α , IFN- γ , IL-12, IL-2, TNF- α , e IL-6.

El polinucleótido exógeno se inserta en la región de delección de tal modo que el mismo se encuentra en marco con el marco de lectura abierto que codifica nsp1 α y nsp1 β , y puede insertarse en tándem más de un polinucleótido exógeno, por ejemplo, pueden estar presentes secuencias de nucleótidos que codifican tres epítipes c-myc. El tamaño total del polinucleótido infeccioso que contiene un polinucleótido exógeno insertado en el sitio de la delección es típicamente no > 16.000 bases, no > 15.800 bases, no > 15.600 bases, no > 15.400 bases, o no > 15.200 bases (con inclusión de la cola poliA). Puede estar presente una inserción en un polinucleótido infeccioso que tiene la delección nsp2 Δ 324-434, nsp2 Δ 324-523, nsp2 Δ 543-632, nsp2 Δ 633-726, nsp2 Δ 543-726, nsp2 Δ 727-813, o nsp2 Δ 324-726. Ejemplos de clones infecciosos que contienen un polinucleótido exógeno en la localización de una delección incluyen un polinucleótido infeccioso que tiene la delección nsp2 Δ 324-434 que contiene una región codificante que codifica una proteína fluorescente verde de 238 aminoácidos, un polinucleótido infeccioso que tiene la delección nsp2 Δ 543-632 que contiene una región codificante que codifica una proteína fluorescente verde de 238 aminoácidos, un polinucleótido infeccioso que tiene la delección nsp2 Δ 324-434 que contiene una región codificante que codifica un epítipo c-myc de 10 aminoácidos (EQKLISEEDL, SEQ ID NO: 17), un polinucleótido infeccioso que tiene la delección nsp2 Δ 324-434 que contiene una región codificante que codifica un epítipo c-myc de 10 aminoácidos, y un polinucleótido infeccioso que tiene las delecciones nsp2 Δ 324-726 o nsp2 Δ 543-726, cada una de las cuales contiene una región codificante que codifica la repetición en tándem del epítipo c-myc de 10 aminoácidos.

Un polinucleótido infeccioso está presente típicamente en un vector, y se hace referencia a la combinación de polinucleótido infeccioso y vector con un clon infeccioso, que se produce por genética inversa. Un vector es un polinucleótido replicante, tal como un plásmido, fago, o cósmido, al cual puede estar unido otro polinucleótido a fin de producir la replicación del polinucleótido unido. La construcción de vectores que contienen un polinucleótido de la invención emplea métodos estándar de DNA recombinante conocidos en la técnica (véase, v.g., Sambrook et al, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbour Laboratory Press (1989)). Un vector puede proporcionar clonación adicional (amplificación del polinucleótido), es decir, un vector de clonación, o expresión del polipéptido codificado por la región codificante, es decir, un vector de expresión, o la combinación de los mismos. El término vector incluye vectores de plásmidos, vectores virales, vectores de cósmidos, o vectores de cromosomas artificiales. Típicamente, un vector es capaz de replicación en un hospedador bacteriano, por ejemplo *E. coli*. El vector puede ser un plásmido.

La selección de un vector depende de una diversidad de características deseadas en el constructo resultante, tales como un marcador de selección o una tasa de replicación del vector. Un vector adecuado para uso como parte de un clon infeccioso puede ser a la vez un vector de clonación y un vector de expresión. Los vectores útiles tienen un número de copias bajo en una célula hospedadora. Células hospedadoras adecuadas para clonación o expresión de los vectores de esta invención son células procariotas o eucariotas. La célula hospedadora puede secretar cantidades mínimas de enzimas proteolíticas. Procariotas adecuados incluyen eubacterias, tales como organismos gram-negativos, por ejemplo, *E. coli* o *S. typhimurium*. Células hospedadoras ilustrativas útiles para producir, manipular, y mantener un clon infeccioso son DH5 α , DH-1 (ATCC 33849), y AG-1.

Un vector incluye secuencias reguladoras enlazadas operativamente al polinucleótido infeccioso. El término "enlazado operativamente" hace referencia a una yuxtaposición de componentes tal que los mismos se encuentran en una relación que les permite funcionar de su manera deseada. Una secuencia reguladora está "enlazada operativamente" a un polinucleótido infeccioso cuando aquélla está unida de tal modo que la expresión de la región codificante se realiza en condiciones compatibles con la secuencia reguladora. Típicamente, un promotor es uno que proporciona fijación de especificidad alta de una RNA-polimerasa, y promotores de este tipo incluyen T7, SP6, y T3. Típicamente, el promotor está situado inmediatamente aguas arriba del primer nucleótido del polinucleótido

infeccioso. Puede estar insertado un GGT entre el promotor y el primer nucleótido del polinucleótido infeccioso. Opcionalmente, el vector contiene también una ribozima del virus de la hepatitis delta aguas debajo de la región poliA.

Opcionalmente, el vector incluye una o más secuencias marcadoras de selección, que codifican típicamente una molécula que desactiva o detecta o es detectada de algún otro modo por un compuesto en el medio de crecimiento. Por ejemplo, la inclusión de una secuencia marcadora de selección puede hacer la célula transformada resistente a un antibiótico, o puede conferir metabolismo específico del compuesto a la célula transformada. Ejemplos de una secuencia marcadora de selección son secuencias que confieren resistencia a kanamicina, ampicilina, cloranfenicol, tetraciclina, y neomicina.

Cuando se produce una delección de nucleótidos que codifican un polipéptido nsp2 en un clon infeccioso, pueden utilizarse métodos estándar de DNA recombinante conocidos en la técnica (véase, v.g., Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press (1989)). Como reconocerán las personas expertas, es una práctica estándar durante la construcción de un clon infeccioso (y cuando se realizan delecciones de construcción en un clon infeccioso) verificar por análisis de la secuencia de nucleótidos la presencia de secuencias de nucleótidos esperadas, tales como delecciones u otras alteraciones y la ausencia de otras mutaciones. Análogamente, cuando se testa un polinucleótido infeccioso candidato para determinar si el mismo es infeccioso, es práctica estándar verificar por análisis de la secuencia de nucleótidos la ausencia de virus contaminante de tipo salvaje.

La presente descripción también incluye polinucleótidos infecciosos aislados descritos en SEQ ID NO:5 y SEQ ID NO:6 así como polinucleótidos infecciosos que tienen semejanza estructural con SEQ ID NO:5 o SEQ ID NO:6. Métodos para determinación de la semejanza estructural se describen en esta memoria. Preferiblemente, un polinucleótido infeccioso de este aspecto puede tener semejanza estructural con SEQ ID NO:5 o SEQ ID NO:6 de al menos 88%, al menos 89%, al menos 90%, al menos 91%, al menos 92%, al menos 93%, al menos 94%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, o al menos 99%. Se considera que un polipéptido que tiene semejanza estructural con SEQ ID NO:5 o SEQ ID NO:6 es un polinucleótido infeccioso si, cuando está presente en una partícula viral y se expone a células permisivas, el polinucleótido se replica en la células permisivas y produce partículas infecciosas de virus.

La presente descripción también incluye partículas de virus aisladas. Como se utiliza en esta memoria, los términos "partícula de virus" y "partícula viral" se utilizan intercambiamente y hacen referencia a un polinucleótido de la presente invención rodeado por una envoltura. Una partícula de virus de la presente descripción puede, cuando se añade a una célula permisiva cultivada, replicarse para dar como resultado la producción de más partículas virales.

Una partícula de virus puede hacerse crecer por pases *in vivo* o en cultivo de células. El sometimiento a pases *in vivo* incluye inocular un cerdo (Faaberg et al., Patente U.S. 7.041.443). El sometimiento a pases en cultivo de células incluye exponer células cultivadas a la partícula de virus e incubar las células en condiciones adecuadas para que el virus se reproduzca y produzca más partículas virales. Las células cultivadas no pueden ser una línea de células inmortalizada o transformada (es decir, las células no son capaces de dividirse indefinidamente). Se pueden utilizar macrófagos alveolares primarios de porcino para el sometimiento a pases en cultivo de células (Faaberg et al., Patente U.S. 7.041.443).

Un virus de la presente descripción puede estar desactivado, es decir, puede estar incapacitado para reproducirse *in vivo* y/o en cultivo de células. Métodos de desactivación se conocen en la técnica e incluyen, por ejemplo, tratamiento de una partícula de virus de la invención con un agente químico desactivador estándar tal como un reactivo aldehídico que incluye formalina, acetaldehído y análogos; alcoholes reactivos ácidos que incluyen cresol, fenol y análogos; ácidos tales como ácido benzoico, ácido benceno-sulfónico y análogos; lactonas tales como beta-propiolactona y caprolactona; y lactamas, carbodiimidas y compuestos carbonil-diheteroaromáticos activados(as) tales como carbonil-diimidazol. Puede utilizarse también irradiación tal como por ejemplo irradiación ultravioleta y gamma para desactivar el virus.

También se incluyen en la presente descripción partículas de virus atenuadas (es decir, virus que tienen una capacidad reducida para causar los síntomas de la enfermedad misteriosa porcina en los cerdos), y métodos de producción de una partícula de virus atenuada. Métodos de producción de un virus atenuado se conocen en la técnica. Típicamente, un virus de la presente descripción se somete a pases, es decir se utiliza para infectar una célula en cultivo, se deja reproducir, y se cosecha finalmente. Este proceso se repite hasta que la virulencia del virus en los cerdos se reduce. Por ejemplo, el virus puede someterse a pases diez veces en cultivo de células, después de lo cual se mide la virulencia del virus. Si la virulencia no ha disminuido, el virus que no se inyectó en el animal se somete a pases diez veces más en cultivo de células. Este proceso se repite hasta que se reduce la virulencia. En general, la virulencia se mide por inoculación de cerdos con virus, y evaluación de la presencia de síntomas clínicos y/o DL₅₀ (véase, por ejemplo, Halbur et al., *J. Vet. Diagn. Invest.*, 8:11-20 (1996), Halbur et al., *Vet. Pathol.*, 32:200-204 (1995), y Park et al., *Am. J. Vet. Res.*, 57:320-323 (1996)). La virulencia puede reducirse hasta que el virus atenuado no causa la muerte de los animales, y tampoco causa síntomas clínicos de la enfermedad.

Típicamente, un cultivo de células útil para producir un virus atenuado de la presente descripción incluye células de origen mamífero no porcino. Ejemplos de cultivos de células de mamífero no porcino incluyen, por ejemplo, la línea de células MA-104 (ATCC CRL-2378), la línea de células MARC-145 (Kim et al., *Arch. Virol.*, 133:477-483 (1993)), y la línea de células CL-2621 (Baustita et al., *J. Vet. Diagn. Invest.*, 5:163-165 (1993)). Se puede utilizar un cultivo de células mixto para producir una partícula de virus atenuada de la presente descripción. En un cultivo de células mixto existen al menos dos tipos de células presentes. Un cultivo de células mixto puede incluir una línea de células inmortalizada o transformada y un cultivo de células primario. Un cultivo de células mixto es particularmente útil cuando un virus se reproduce lentamente, o no se reproduce en absoluto, en una línea de células inmortalizada o transformada. Ejemplos de una línea de células inmortalizada o transformada para uso en un cultivo de células mixto incluyen, por ejemplo, la línea de células MARC-145 (Kim et al., *Arch. Virol.*, 133:477-483 (1993)), y la línea de células MA-104 (ATCC CRL-2378). Los cultivos de células primarias para uso en un cultivo de células mixto pueden ser de origen porcino. Un ejemplo preferido de un cultivo de células primario para uso en un cultivo de células mixto son macrófagos primarios alveolares de porcino.

Adicionalmente la presente descripción incluye los polipéptidos codificados por regiones codificantes de nsp2 presentes en los polinucleótidos descritos en la Tabla 2, con inclusión de aquéllos que son viables. Se incluyen también en la presente descripción anticuerpos, con inclusión de anticuerpos monoclonales y policlonales, que se fijan específicamente a un polipéptido codificado por las regiones codificantes de nsp2 presentes en los polinucleótidos descritos en la Tabla 2. El término "anticuerpo", a no ser que se especifique lo contrario, incluye fragmentos de anticuerpos enteros que retienen su actividad de fijación para un antígeno diana. Tales fragmentos incluyen fragmentos Fv, F(ab') y F(ab')₂, así como anticuerpos monocatenarios (scFv). Como se utiliza en esta memoria, un anticuerpo que puede "fijarse específicamente" a un polipéptido es un anticuerpo que interacciona únicamente con el epítipo del antígeno que indujo la síntesis del anticuerpo, o interacciona con un epítipo estructuralmente afín. Un anticuerpo que "se fija específicamente" a un epítipo interaccionará, en las condiciones apropiadas, con el epítipo incluso en presencia de una diversidad de dianas de fijación posibles. Como se utiliza en esta memoria, el término "complejo polipéptido:anticuerpo" hace referencia al complejo que resulta cuando un anticuerpo se fija específicamente a un polipéptido, o una subunidad o análogo del mismo. En algunos aspectos, un anticuerpo de la presente descripción incluye aquéllos que no se fijan específicamente a un polipéptido nsp2 de longitud total codificado por VR-2332 (v.g., número de acceso a GenBank U87392, aminoácidos 384-1363 del ORF1 (véase también Allende et al. *J. Gen. Virol.*, 80:307-315 (1999) o aminoácidos 384-1580 del ORF1 (véase también Ziebuhr et al., *J. Gen. Virol.*, 81:853-879 (2000)). Tales anticuerpos pueden identificarse utilizando métodos de rutina conocidos en la técnica.

Los anticuerpos de la presente descripción se pueden preparar utilizando el polipéptido intacto. Opcionalmente, un polipéptido nsp2 descrito en esta memoria puede estar unido o conjugado covalentemente a un polipéptido portador para mejorar las propiedades inmunológicas del polipéptido. Polipéptidos portadores útiles se conocen en la técnica.

La preparación de anticuerpos policlonales es bien conocida. Los anticuerpos policlonales pueden obtenerse por inmunización con un inmunógeno de una diversidad de animales de sangre caliente tales como caballos, vacas, cabras, ovejas, perros, pollos, conejos, ratones, hámster, cobayos y ratas así como animales transgénicos tales como ovejas, vacas, cabras o cerdos transgénicos. Los anticuerpos resultantes pueden aislarse de otras proteínas utilizando una columna de afinidad que tenga un resto de fijación Fc, tal como proteína A, o análogos.

Los anticuerpos monoclonales pueden obtenerse por diversos métodos familiares para los expertos en la técnica. Resumidamente, células de bazo de un animal inmunizado con un antígeno deseado se inmortalizan, comúnmente por fusión con una célula de mieloma, (véase, por ejemplo, *Antibodies: A Laboratory Manual*, Harlow et al., eds., Cold Spring Harbor Laboratory Press: Cold Spring Harbor, Nueva York, (1988)). Los anticuerpos monoclonales pueden aislarse y purificarse a partir de cultivos de hibridoma por métodos bien conocidos en la técnica.

El anticuerpo puede producirse recombinantemente, por ejemplo, por presentación de fago o por métodos combinatorios. La presentación de fago y los métodos combinatorios pueden utilizarse para aislar anticuerpos recombinantes que se fijan a un polipéptido descrito en esta memoria, o una subunidad biológicamente activa o análogo de la misma (véase, por ejemplo, Ladner et al., Patente EE.UU. nº. 5.223.409). Tales métodos pueden utilizarse para generar anticuerpos monoclonales humanos.

La presente descripción también proporciona composiciones que incluyen un polinucleótido infeccioso, polinucleótido PRRS, partícula de virus, o anticuerpo de la presente descripción. Tales composiciones incluyen típicamente un portador farmacéuticamente aceptable. Como se utiliza en esta memoria, "portador farmacéuticamente aceptable" incluye solución salina, disolventes, medios de dispersión, recubrimientos, agentes antibacterianos y antifúngicos, agentes isotónicos y retardantes de la absorción, y análogos, compatibles con administración farmacéutica. Compuestos activos adicionales pueden incorporarse también en las composiciones.

Una composición se puede preparar por métodos bien conocidos en la técnica de la farmacia. En general, una composición puede formularse de modo que sea compatible con su ruta de administración proyectada. Ejemplos de rutas de administración incluyen perfusión y ruta parenteral, v.g. intravenosa, intradérmica, subcutánea, oral (v.g. inhalación), transdérmica (tópica), y transmucosal. Las soluciones o suspensiones pueden incluir los siguientes componentes: un diluyente estéril tal como agua para administración, solución salina, aceites fijos, polietilén-glicoles,

glicerina, propilen-glicol u otros disolventes sintéticos; agentes antibacterianos tales como alcohol bencílico o metil-parabenos; antioxidantes tales como ácido ascórbico o bisulfito de sodio; agentes quelantes tales como ácido etilendiaminatetraacético; tampones tales como acetatos, citratos o fosfatos; electrolitos, tales como ion sodio, ion cloruro, ion potasio, ion calcio, e ion magnesio, y agentes para el ajuste de la tonicidad tales como cloruro de sodio o dextrosa. El pH puede ajustarse con ácidos o bases, tales como ácido clorhídrico o hidróxido de sodio. Una composición puede confinarse en ampollas, jeringuillas desechables o viales de dosis múltiples hechos de vidrio o plástico.

Las composiciones pueden incluir soluciones acuosas estériles (en el caso de ser solubles en agua) o dispersiones y polvos estériles para la preparación extemporánea de soluciones o dispersiones estériles. Para administración intravenosa, portadores adecuados incluyen solución salina fisiológica, agua bacteriostática, Cremophor EL™ (BASF, Parsippany, N.J.) o solución salina tamponada con fosfato (PBS). Una composición es típicamente estéril y, cuando es adecuada para uso inyectable, debe ser fluida en tal grado que exista susceptibilidad fácil de administración con jeringuilla. La composición debería ser estable en las condiciones de fabricación y almacenamiento y estar conservada contra la acción contaminante de microorganismos tales como bacterias y hongos. El portador puede ser un disolvente o medio de dispersión que contenga, por ejemplo, agua, etanol, poliol (por ejemplo, glicerol, propilenglicol, y polietilenglicol líquido, y análogos), y mixturas adecuadas de los mismos. La prevención de la acción de los microorganismos puede conseguirse por diversos agentes antibacterianos y antifúngicos, por ejemplo, parabenos, clorobutanol, fenol, ácido ascórbico, timerosal, y análogos. En muchos casos, se pueden incluir agentes isotónicos, por ejemplo, azúcares, polialcoholes tales como manitol, sorbitol, cloruro de sodio en la composición. La absorción prolongada de las composiciones inyectables puede producirse por inclusión en la composición de un agente que retarda la absorción, por ejemplo, monoestearato de aluminio y gelatina.

Pueden prepararse soluciones estériles por incorporación del compuesto activo (es decir, un polinucleótido infeccioso o virus PRRS de la presente invención) en la cantidad requerida en un disolvente apropiado con uno o una combinación de ingredientes enumerados anteriormente, en caso requerido, seguido por esterilización mediante filtración. Generalmente, las dispersiones se preparan por incorporación del compuesto activo en un vehículo estéril, que contiene un medio de dispersión básico y los otros ingredientes requeridos de los arriba enumerados. En el caso de polvos estériles para la preparación de soluciones estériles inyectables, los métodos de preparación pueden ser secado a vacío y liofilización, que produce un polvo del ingrediente activo más cualquier ingrediente adicional deseado a partir de una solución del mismo o de los mismos previamente esterilizada por filtración.

Las composiciones orales incluyen generalmente un diluyente inerte o un portador comestible. Para el propósito de administración terapéutica oral, el compuesto activo puede incorporarse con excipientes y utilizarse en la forma de tabletas, trociscos, o cápsulas, v.g., cápsulas de gelatina. Estas composiciones pueden estar también conformadas en un polvo o suspendidas en una solución acuosa de tal modo que estos polvos y/o soluciones puedan añadirse al pienso animal o al agua de bebida de los animales. Estas composiciones pueden estar convenientemente edulcoradas o saborizadas por diversos agentes conocidos para promover la absorción de la vacuna oralmente por el cerdo.

Los compuestos activos se pueden administrar también por cualquier método adecuado para administración de agentes polinucleótidos, v.g., utilizando cañones de genes, bioinyectores, y parches dérmicos así como métodos sin aguja tales como la tecnología de vacunas de DNA de micropartículas descrita por Johnston et al (Patente EE.UU. n°. 6.194.389). Adicionalmente, es posible el suministro intranasal, como se describe, por ejemplo, en Hamajima et al., *Clin. Immunol. Immunopathol.*, 88:205-210 (1998). Pueden utilizarse también liposomas y microencapsulación.

Los compuestos activos se pueden preparar con portadores que protegerán el compuesto contra la eliminación rápida del cuerpo, tales como una formulación de liberación controlada, con inclusión de implantes. Pueden utilizarse polímeros biodegradables biocompatibles, tales como etileno-acetato de vinilo, polianhídridos, ácido poliglicólico, colágeno, poliortoésteres, y poli(ácido láctico). Tales formulaciones se pueden preparar utilizando técnicas estándar. Los materiales pueden obtenerse también comercialmente de, por ejemplo, Alza Corporation y Nova Pharmaceuticals, Inc. Pueden utilizarse también suspensiones de liposomas como portadores farmacéuticamente aceptables. Éstas se pueden preparar de acuerdo con métodos conocidos por los expertos en la técnica.

La toxicidad y eficacia terapéutica de tales compuestos activos pueden determinarse por procedimientos farmacéuticos estándar en cultivos de células o animales experimentales, v.g., para determinación de la DL₅₀ (la dosis letal para el 50% de la población) y la DE₅₀ (la dosis terapéuticamente eficaz en el 50% de la población). La ratio de dosis entre los efectos tóxico y terapéutico es el índice terapéutico, y puede expresarse como la ratio DL₅₀/DE₅₀. Se pueden usar compuestos que exhiban índices terapéuticos altos.

Los datos obtenidos a partir de ensayos de cultivo de células y estudios en animales pueden utilizarse en la formulación de una gama de dosificaciones para uso en el campo. La dosificación de tales compuestos puede estar basada preferiblemente en una gama de concentraciones circulantes que incluyen la DE₅₀ con toxicidad escasa o nula. La dosificación puede variar dentro de este intervalo dependiendo de la forma de dosificación empleada y la ruta de administración utilizada.

Las composiciones se pueden administrar una o dos veces al día hasta una o más veces por semana, con inclusión de una sola vez en días alternos. El profesional experto apreciará que ciertos factores pueden influir en la dosificación y la temporización requeridas para tratar eficazmente a un individuo, incluyendo, pero sin carácter limitante, la gravedad de la enfermedad o trastorno, tratamientos previos, el estado general de salud y/o la edad del individuo, y otras enfermedades presentes. Además, el tratamiento de un individuo con una cantidad eficaz de un polipéptido puede incluir un solo tratamiento o una serie de tratamientos.

Se describen también en esta memoria métodos para utilización de las composiciones aquí descritas. En un aspecto, se describen métodos de tratamiento de uno o más síntomas de la enfermedad misteriosa del cerdo en un animal que puede estar causada por infección por un virus PRRS. El método incluye administrar una cantidad eficaz de una composición a un animal que padece o se encuentra en riesgo de padecer la enfermedad misteriosa del cerdo, o síntomas de la enfermedad misteriosa del cerdo.

El tratamiento de la enfermedad misteriosa del cerdo, o los síntomas de la enfermedad misteriosa del cerdo, puede ser profiláctico o, alternativamente, puede iniciarse después de la aparición de la enfermedad o los síntomas de la misma. Como se utiliza en esta memoria, el término "síntoma" se refiere a la evidencia objetiva en un individuo de la enfermedad misteriosa del cerdo. Los síntomas asociados con la enfermedad misteriosa del cerdo y las evaluaciones de tales síntomas son rutinarios y se conocen en la técnica. Ejemplos de síntomas incluyen aborto, anorexia, fiebre, estado letárgico, neumonía, coloración roja/azul de las orejas, respiración laboriosa (disnea), y ritmo respiratorio incrementado (taquipnea). Se hace referencia en esta memoria a un tratamiento que es profiláctico, por ejemplo, iniciado antes que un individuo manifieste síntomas de una afección causada por un virus PRRS, como tratamiento de un individuo que se encuentra "en riesgo" de desarrollar la enfermedad o síntomas de la misma. Típicamente, un animal "en riesgo" es un animal que se encuentra en un área en la que se han diagnosticado animales que presentan la enfermedad o síntomas de la misma y/o es probable que se vea expuesto a un virus PRRS. De acuerdo con ello, la administración de una composición puede realizarse antes, durante, o después de la aparición de las condiciones descritas en esta memoria. El tratamiento iniciado después de la aparición de una condición puede dar como resultado disminución de la gravedad de los síntomas de una de las condiciones, o la eliminación completa de los síntomas.

En algunos aspectos, los métodos incluyen típicamente administrar a un animal una composición que incluye una cantidad eficaz de una partícula de virus de la presente descripción. Una "cantidad eficaz" es una cantidad efectiva para prevenir la manifestación de síntomas de la enfermedad misteriosa del cerdo, reducir la gravedad de los síntomas de la enfermedad, y/o eliminar por completo los síntomas. Típicamente, la cantidad eficaz es una cantidad que da como resultado una respuesta inmune humoral y/o celular que protege al animal durante la exposición futura a un virus PRRS. La partícula de virus utilizada en la composición puede contener un polinucleótido infeccioso que tiene una delección como se describe en esta memoria. Opcionalmente, el polinucleótido infeccioso incluye también un polinucleótido exógeno presente en la localización de la delección. Una ventaja de la utilización de una partícula de virus que tiene una delección (o un polinucleótido exógeno presente en la localización de la delección) es que la misma puede distinguirse fácilmente de otros virus PRRS, con inclusión del virus PRRS de tipo salvaje presentes en el campo. La partícula de virus puede identificarse por aislamiento del virus de un animal seguido, por ejemplo, por secuenciación, digestión con enzimas de restricción, o amplificación de nucleótidos específicos basada en PCR. A una partícula de virus "marcada" de este tipo se hace referencia a menudo en la técnica como vacuna marcadora.

En otros aspectos de la presente descripción, los clones infecciosos y/o polinucleótidos infecciosos descritos en esta memoria pueden utilizarse para investigar inserciones de genes viables, investigar RNA o proteínas alternativos expresados distintos del virus de longitud total, investigar recombinación viral, e investigar propiedades inmunógenas de nsp2 de longitud total con relación a nsp2 truncado.

Ejemplos

Ejemplo 1

Se desarrollaron clones de cDNA de longitud total del virus del síndrome reproductivo y respiratorio porcino de Norteamérica (PRRSV), cepa prototipo VR-2332, poseyendo cada versión progresiva menos cambios de nucleótidos que las versiones previas cuando se comparaban con la cepa *wt* VR-2332. Se recuperó el virus de la progenie de cada clon infeccioso y se analizó para verificación de la secuencia de nucleótidos, la tasa de crecimiento *in vitro* y el tamaño de las calvas. La progenie de un clon infeccioso confirmó una replicación robusta *in vivo*, observada por la aparición de anticuerpos alfa-PRRSV a la misma tasa que el virus *wt*. El análisis por transferencia Northern *in vivo* de la progenie reveló también que especies de RNA subgenómicas defectuosas, denominadas heteróclitos (formas no comunes), estaban presentes junto con genomas de longitud total. El análisis concurrente por transferencia Northern de una serie de pases de cultivos de células MA-104 infectadas reveló que el virus recombinante adquiría sólo gradualmente un perfil simultáneo de RNA de longitud total y heteróclito similares a la especie de RNA observada en la infección *in vivo*.

Materiales y métodos

- Células y cepas virales. Células MA-104 o sus células descendientes MARC-145 (ATCC CRL-11171), una línea de células epiteliales de riñón del mono verde africano que soporta replicación de PRRSV (Meng et al., *J. Vet. Diagn. Invest.*, 8:374-81 (1996)), se mantuvieron en medio esencial mínimo de Eagle (EMEM) (JRH Biosciences 56416), complementado con 1 mg/ml de NaHCO_3 y 10% de suero bovino fetal (FBS), a 37°C con 5% de CO_2 . Las células cultivadas se transfectaron con RNA o se infectaron con virus cuando el crecimiento monocapa había alcanzado 70-80% de confluencia. Las cepas prototipo PRRSV de Norteamérica VR-2332 e Ingelvac® MLV han sido descritas previamente (Yuan et al., *Virus Res.*, 79:189-200 (2001)). La cepa VR-2332 crece hasta títulos equivalentes en ambas líneas de células.
- 10 Purificación del RNA viral. Se purificó RNA viral (vRNA) como se describe. (Chen et al., *J. Gen. Virol.*, 75:925-930 (1994); Yuan et al., *Virus Res.*, 79:189-200 (2001)). Resumidamente, el sobrenadante de células MARC-145 infectadas con VR-2332 se cosechó el día 4 después de la infección (p.i.). Después de eliminación de los residuos celulares por centrifugación a 12.000 rpm, los sobrenadantes se estratificaron sobre una capa amortiguadora de 2 ml de sacarosa 0,5 M y se centrifugaron a 76.000 x g durante 4 horas. Los viriones reducidos a un sedimento se resuspendieron en 0,5 ml de LES (LiCl 0,1 M/EDTA 5 mM/1,0% SDS) y se digirieron ulteriormente por adición de 100 µg de proteinasa K a 56°C para eliminar totalmente la proteína. Después de 10 minutos de incubación, se extrajo el vRNA varias veces con fenol ácido y fenol/cloroformo y se precipitó luego en etanol de 70% v/v. El vRNA reducido a un sedimento se resuspendió inmediatamente en 50 µl de agua o tampón TE exento de RNasa (Tris-HCl 10 mM, EDTA 1 mM, pH 8,0) y se guardó a -80°C.
- 20 Construcción de cDNA viral de longitud total. La síntesis del cDNA se realizó con el Kit RT-PCR Enhanced Avian HS (Sigma, HSRT-100). Se utilizaron ocho cebadores PCR (Tabla 3) para amplificar cuatro fragmentos de cDNA solapantes que cubrían el genoma completo de VR-2332 (Figura 2). Las condiciones de los ciclos eran 94°C durante 2 minutos, seguido por 35 ciclos de 94°C durante 15 segundos, 68°C durante 4-5 segundos, seguidos por 68°C durante 5 minutos. Cada fragmento PCR se purificó con el Kit Gel Extraction QIAEX II (Qiagen) y se clonó en el vector pCR® 2.1-TOPO® con el Kit de Clonación TOPO TA® (Invitrogen K450001). Los plásmidos que representaban cada fragmento se sometieron a análisis de la secuencia de nucleótidos. Los fragmentos con las mutaciones de nucleótidos mínimas comparadas con la secuencia parental VR-2332 (número de sumisión a GenBank U87392) se utilizaron para ensamblar el cDNA de longitud total, como se muestra en la Figura 2. En cada región solapante, se utilizó un solo sitio de enzima de restricción para unir los fragmentos flanqueantes. Cuatro fragmentos digeridos, que representaban la secuencia genómica de longitud total, se ensamblaron con precisión escalonadamente en un vector de plásmido modificado de bajo número de copias (pOK12HDV-PacI). El vector se modificó para incluir la ribozima HDV por inserción de un fragmento de 244 pb *SmaI* a *SacII* que contenía la ribozima antígeno de HDV y una secuencia terminadora de RNA-polimerasa T7 del vector de transcripción 2.0 (Johnson et al., *J. Virol.*, 71:3323-3327 (1997); Pattnaik et al., *Cell*, 69:1011-1020 (1992)) en los sitios correspondientes en pOK12 (Vieira et al., *Gene*, 100:189-194 (1991)). El sitio de la enzima de restricción *NcoI* en este fragmento de 244 pb se reemplazó con un sitio singular *PacI* por mutación de oligonucleótidos con juegos de cebadores 5' pOK12HDV-2157/3' pOK12HDV-257 y 5' pOK12HDV-257/poliA-modificado (Tabla 3), seguido por fusión PCR. En los clones cDNA de longitud total, la secuencia genómica viral estaba precedida por el promotor de RNA-polimerasa T7, 1 ó 2 residuos G y un residuo T, y seguida por una cola de ácido poliadenílico de 50 nucleótidos. Los clones ensamblados se propagaron en la cepa DH5α de *Escherichia coli* y se presentaron luego para la confirmación de la secuencia nucleotídica del genoma total.

Tabla 3. Cebadores oligonucleotídicos utilizados en este estudio. Los cebadores directos se indican con una barra oblicua (/) después del designador, y los cebadores inversos van precedidos por una barra oblicua. Los sitios de las enzimas de restricción insertadas se muestran en cursiva subrayada

Cebador	Posición en el genoma*	Secuencia
Clonación:		
ConductorT7-VR largo	1-31	5'-ACAT <u>G</u> CATGCTTAATACGACTCACTATAGTATGACGTATAGGTGTTGGCTCTATGCCTTGG (SEQ ID NO:18)
/3'-4300	4617-4635	5'-CTGGGCGACACAGTCCTA (SEQ ID NO:19)
5'4056.5'-4056- <i>AscII</i>	4055-4080	5'-CTTCTC <u>G</u> GCGCGCCGAATGGGAGT (SEQ ID NO:20)
/3'-7579	7578-7603	5'-TCATCATACCTAG <u>G</u> GCCTGCTCCACG (SEQ ID NO:21)
5'-7579/	7578-7603	5'-CGTGGAGCAGGC <u>C</u> TAGGTATGATGA (SEQ ID NO:22)
/P32	13293-13310	5'-TGCAGGCGAACGCCTGAG (SEQ ID NO:23)

Cebador	Posición en el genoma*	Secuencia
Clonación:		
VR1509/	11938-11958	5'-GTGAGGACTGGGAGGATTACA (SEQ ID NO:24)
/extremo 3'-FL	15405-15411	5'-CTCTTTAATTAAGTAG(T) ₃₀ AATTTCG (SEQ ID NO:25)
Mutagénesis:		
5'-pOK12HDV-257/(SphI, PacI)	pOK12HDV-PacI 257-282	5'-GATGCATGCCATTAATTAAGGGTCGGC (SEQ ID NO:26)
'3'/3'-pOK12HDV-257(SphI, PacI)	pOK12 HDV-PacI 257-282	5'-GCCGACCCTTAATTAATGGCATGCATC (SEQ ID NO:27)
T7leader-VR-2G/	1-5	5'-ACATGCATGCTTAATACGACTCACTATAGGTATGAC (SEQ ID NO:28)
7475G2A/	7453-7477	5'-5Phos/CTGTGTGGACATGTCACCATTGAAA (SEQ ID NO:29)
13860C2T/	13843-13867	5'-5Phos/GTGTATCGTGCCGTTCTGTTTTGCT (SEQ ID NO:30)
14979A2G/	14958-14982	5'-5Phos/CAGATGCTGGGTAAGATCATCGCTC (SEQ ID NO:31)
Análisis por transferencia Northern:		
/3'-UTR	15298-15336	5'-GCACAATGTCAATCAGTGCCATTACCCACACATTCTTCC (SEQ ID NO:32)
/1a-p222	221-261	5'-TAGACTTGCCCTCCGCCATAAACACCTGGCATTGGGGGT (SEQ ID NO:33)
* La posición en el genoma está basada en la Presentación a GenBank U87392		

Modificación y análisis de la secuencia de los clones de cDNA de longitud total. Se utilizó el Kit de Mutagénesis Dirigida Multisite QuikChange® (Stratagene) para modificar todos los clones de cDNA desde pVR-V4 a pVR-V6G7475A. Las inserciones completas de plásmidos de cDNA genómico se presentaron luego al Centro de Análisis Genético Avanzado de la Universidad de Minnesota (AGAC) para análisis de la secuencia de nucleótidos con cebadores de secuenciación apropiados (Tabla 3). Las diferencias de secuencia entre pVR-V4 hasta pVR-V6G7475A, así como las de VR-2332 parental, su cepa de vacuna atenuada correspondiente, Ingelvac MLV, y pVR-HN, el primer clon infeccioso de VR-2332, se enumeran en la Tabla 4. (Nelsen et al., *J. Virol.*, 73:270-80 (1999); Yuan et al., *Virus Res.*, 79:189-200 (2001); Nielsen et al., *J. Virol.*, 77:3702-3711 (2003)).

Tabla 4. Diferencias de nucleótidos entre las cepas de PRRSV y los clones infecciosos VR-2332. Únicamente se muestran las posiciones en las cuales se observaron diferencias de nucleótidos. Los nucleótidos que se representan en la cepa VR-2332 se muestran en recuadros sin sombrear. Los recuadros con sombreado claro representan diferencias de nucleótidos que son exclusivas para el clon infeccioso, los recuadros con sombreado medio hacen resaltar aquellos nucleótidos que se observan también en Ingelvac® MLV, y los recuadros con sombreado negro indican nucleótidos exclusivos de porcino. Las regiones que no se secuenciaron se indican por una barra oblicua

Base*	Region	VR-2332	V4	V5	V5-1.P3	V5-2.P3	V5-Swine 612	V5G7475A	V5G7475A -P3	V6	V6G7475A	V6G7475A -P3	VR-HN	MLV
-4	5'UTR					T								
-3			G	G		T		G		G			G	
-2			G	G		T		G		G			G	
-1									T/Ø					
48	NSP1a	A	A	A	A	A	A	A	A	A	A	R(G/A)	A	A
102		A	A	A	A	A	G	A	A	A	A	A	A	A
258		C	C	C	C	C	C	C	C	C	C	C	A	A
309		A	G	G	G	G	G	G	G	G	G	G	A	C
415		T	T	T	-	Y(C/T)	T	T	-	T	T	-	T	A
642		T	C	C	-	C	C	C	-	C	C	-	T	T
784		G	G	G	-	G	G	G	-	G	G	-	G	T
827	NSP1b	C	C	C	-	C	T	C	-	C	C	-	C	C
1074		C	C	C	-	C	C	C	-	C	C	-	T	C
1107		A	G	G	-	G	G	G	-	G	G	-	A	A
1122		A	A	A	-	R(G/A)	A	A	-	A	A	-	A	A
1181		C	C	C	-	C	C	C	-	C	C	-	C	A
1294		A	A	A	-	R(G/A)	A	A	-	A	A	-	A	A
1379		C	C	C	-	C	T	C	-	C	C	-	C	C
1595	NSP2	C	A	C	-	C	C	C	-	C	C	-	C	C
2192		C	C	C	-	C	C	C	-	C	C	-	C	C
3040		G	G	G	-	G	G	G	-	G	G	-	G	C
3457		G	G	G	-	G	G	G	-	G	G	-	G	C
3657		C	C	C	-	Y(C/T)	C	C	-	C	C	-	C	C
4407		T	C	C	-	C	C	C	-	C	C	-	T	T
4593		A	G	G	-	G	G	G	-	G	G	-	A	A
4881	NSP3	T	G	G	-	G	G	G	-	G	G	-	G	A
4865		T	T	T	-	Y(C/T)	T	T	-	T	T	-	T	T
4866		A	G	G	-	G	G	G	-	G	G	-	A	A
5097		G	A	A	-	A	A	A	-	A	A	-	A	A
5247		T	C	C	-	C	C	C	-	C	C	-	T	T
5519		C	C	C	-	C	C	C	-	C	C	-	T	C
5610		T	T	T	-	T	T	T	-	T	T	-	A	T
6345	NSP5	A	A	A	-	A	A	A	-	A	A	-	A	A

	C	G	G	G	-	-	G	T	G	-	-	G	T	G	-	-	C	A	G
6674	G	T	T	T	-	-	T	T	T	-	-	T	T	T	-	-	A	C	T
6853	A	A	A	A	-	-	A	A	A	-	-	A	A	A	-	-	R(G/A)	C	A
6966	C	C	C	C	-	-	C	C	C	-	-	C	C	C	-	-	Y(C/T)	C	A
7188	C	C	C	C	-	-	C	C	C	-	-	C	C	C	-	-	Y(C/T)	C	C
7189	C	C	C	C	-	-	C	C	C	-	-	C	C	C	-	-	M(A/C)	C	C
7213	G	A	A	A	-	-	A	A	A	-	-	A	A	A	-	-	A	G	C
7329	A	G	G	G	-	-	G	G	G	-	-	G	G	G	-	-	A	A	G
7475	T	C	C	C	-	-	C	C	C	-	-	C	C	C	-	-	C	A	A
7554	T	C	C	C	-	-	C	C	C	-	-	C	C	C	-	-	C	C	T
9220	T	A	A	A	-	-	T	T	T	-	-	T	T	T	-	-	T	T	T
9649	T	A	A	A	-	-	T	T	T	-	-	T	T	T	-	-	G	G	G
9918	T	T	T	T	-	-	T	T	T	-	-	T	T	T	-	-	T	T	T
9958	G	T	T	T	-	-	T	T	T	-	-	T	T	T	-	-	T	T	T
10040	A	G	A	A	-	-	A	A	A	-	-	A	A	A	-	-	A	A	A
10533	T	T	T	T	-	-	T	T	T	-	-	T	T	T	-	-	T	T	T
10643	T	T	T	T	-	-	T	T	T	-	-	T	T	T	-	-	T	T	T
10697	T	T	T	T	-	-	T	T	T	-	-	T	T	T	-	-	T	T	T
10739	C	T	C	C	-	-	C	C	C	-	-	C	C	C	-	-	C	C	C
10781	G	A	G	G	-	-	G	G	G	-	-	G	G	G	-	-	A	A	A
10803	T	C	C	C	-	-	T	T	T	-	-	T	T	T	-	-	C	C	C
10895	C	C	C	C	-	-	C	C	C	-	-	C	C	C	-	-	C	C	C
11055	T	T	T	T	-	-	T	T	T	-	-	T	T	T	-	-	T	T	T
11081	G	A	A	A	-	-	G	G	G	-	-	G	G	G	-	-	A	A	A
11221	G	T	T	T	-	-	T	T	T	-	-	T	T	T	-	-	T	T	T
11229	G	G	G	G	-	-	G	G	G	-	-	G	G	G	-	-	A	A	A
11259	C	G	G	G	-	-	C	C	C	-	-	C	C	C	-	-	G	C	C
11327	C	T	C	C	-	-	C	C	C	-	-	C	C	C	-	-	C	C	C
11329	G	C	C	C	-	-	C	C	C	-	-	C	C	C	-	-	C	C	C
11501	G	G	G	G	-	-	G	G	G	-	-	G	G	G	-	-	C	C	C
11666	C	C	C	C	-	-	C	C	C	-	-	C	C	C	-	-	A	A	A
11744	G	C	C	C	-	-	G	G	G	-	-	G	G	G	-	-	G	G	G
11760	G	G	G	G	-	-	R(A/G)	G	G	-	-	G	G	G	-	-	G	G	G
11882	A	G	A	A	-	-	A	A	A	-	-	A	A	A	-	-	A	A	A
12076	A	G	A	A	-	-	A	A	A	-	-	A	A	A	-	-	A	A	A
12102	G	G	G	G	-	-	G	G	G	-	-	G	G	G	-	-	G	G	G
12153	A	G	A	A	-	-	A	A	A	-	-	A	A	A	-	-	A	A	A
12432	A	G	A	A	-	-	A	A	A	-	-	A	A	A	-	-	A	A	A
12501	T	A	T	T	-	-	T	T	T	-	-	T	T	T	-	-	T	T	T
12600	G	G	G	G	-	-	G	G	G	-	-	G	G	G	-	-	G	G	G
12943	C	C	C	C	-	-	C	C	C	-	-	C	C	C	-	-	C	C	C
12950	T	T	T	T	-	-	T	T	T	-	-	T	T	T	-	-	T	T	T
13011	G	G	G	G	-	-	G	G	G	-	-	G	G	G	-	-	T	T	T
No ORF	T	T	T	T	-	-	T	T	T	-	-	T	T	T	-	-	C	C	T

[illegible]

* Las bases negativas hacen referencia a aquellos nucleótidos presentes en el RNA después de transcripción y derivados del promotor de RNA-polimerasa inmediatamente aguas arriba del polinucleótido infeccioso. Estos nucleótidos derivados del promotor ya no están presentes típicamente en un polinucleótido infeccioso después que el mismo se ha sometido a pases 9 veces.

- Transcripción *in vitro*. El clon de cDNA de longitud total se linealizó por escisión con *PacI*, que corta aguas abajo de la cola poliA. Se produjeron transcritos de RNA protegidos terminalmente [análogo de la caperuza m⁷G(5')ppp(5')G] utilizando el Kit mMESSAGE MACHINE™ (Ambion) y una ratio optimizada 2:1 de análogo metilado de la caperuza a GTP. Se generaron aproximadamente 50 a 60 µg de RNA a partir de 2 µg de matriz de DNA en 20 µl de mixtura de reacción. El aumento de la ratio de análogo de la caperuza a GTP reducía sustancialmente el rendimiento de RNA. Se purificó subsiguientemente el RNA con fenol ácido-cloroformo seguido por precipitación con isopropanol y se resuspendió en tampón TE exento de nucleasas (pH 8,0). El RNA se evaluó respecto a calidad por comparación de tamaños con RNA viral de VR-2332 de tipo salvaje en un gel de agarosa desnaturalizante con 1% de glixal, y se cuantificó por espectrofotometría a DO₂₆₀.
- 10 Transfección de las células MARC-145. Se generó un procedimiento de transfección modificado basado en el método descrito por Nielsen et al., (*J. Virol.*, 77:3702-3711 (2003)). Para la transfección, se sembraron las células MARC-145 en placas de 6 pocillos (2-3 x 10⁵ células/pocillo) en 3 ml de medio completo [EMEM complementado con 10% de suero bovino fetal (FBS)] y se incubaron luego a 37°C, 5% CO₂ durante 20-24 horas hasta aproximadamente 80% de confluencia (Collins et al., *J. Vet. Diagn. Invest.*, 4:117-126 (1992)). Se combinaron 4 µg de RNA transcrito *in vitro* diluido en 500 µl de Medio de Suero Reducido Opti-MEM® I (Invitrogen) y 2 µl de bromuro de 1,2-dimiristiloxipropil-3-dimetil-hidroxi-etil-amonio y colesterol (DMRIE-C; Invitrogen) diluidos en 1 ml de medio Opti-MEM®, y se agitaron con brevedad enérgicamente. Las células MARC-145 se lavaron una sola vez con 2 ml de medio Opti-MEM® y se cubrieron inmediatamente después con la solución del complejo lípido:RNA. Se utilizó DMRIE-C sin RNA (2 µl) como control negativo y se empleó como control positivo DMRIE-C con 10-100 ng de RNA viral purificado de la cepa (tipo salvaje) *wt* VR-2332. Después de 4 horas de exposición a los complejos lípido:RNA, se lavaron las monocapas y se añadió medio completo nuevo (EMEM con 10% FBS). Los sobrenadantes de las células transfectadas se monitorizaron diariamente respecto a la aparición de efecto citopático (CPE) y se sometieron a pases sobre MARC-145 nuevas 72-96 horas después de la transfección.
- 15 20 Detección del RNA viral de la progenie. Para testar el RNA viral de la progenie, se cosecharon sobrenadantes de cultivo celular de células MARC-145 transfectadas e infectadas. Se aisló el RNA con el Kit de RNA viral QiaAmp (Qiagen). Se realizó la RT-PCR con pares de cebadores seleccionados, específicos para los nucleótidos de la cepa VR-2332 que eran indicativos de residuos mutados de clones infecciosos (Tabla 3). La confirmación de la progenie de clones infecciosos se obtuvo por verificación de la secuencia nucleotídica de nucleótidos específicos de clones presentes en los productos RT-PCR.
- 25 30 Detección de la proteína de la nucleocápsida viral de la progenie. Se utilizaron ensayos de inmunofluorescencia indirecta (IFA) para detectar la expresión de proteínas virales en células MARC-145 transfectadas con RNA transcrito *in vitro*, o infectadas con el virus de la progenie, dispuestas en cubreobjetos. Las células infectadas se fijaron en paraformaldehído al 3,7% con solución salina tamponada con fosfato (PBS), de pH 7,5, a la temperatura ambiente durante 10 minutos. Las células fijadas se lavaron con PBS, se incubaron a 37°C durante 45 minutos en el anticuerpo monoclonal SDOW17 específico de la proteína de la nucleocápsida del PRRSV (Magar et al., *Can. J. Vet. Res.*, 59:232-234 (1995)) y se incubaron posteriormente con inmunoglobulina G (IgG) anti-ratón de cabra conjugada con isotiocianato de fluoresceína a 37°C durante 45 minutos más (dilución 1:100) (Sigma). Se lavaron los cubreobjetos con PBS, se montaron en portaobjetos utilizando aceite de soporte de gel, y se observaron bajo un microscopio de fluorescencia.
- 35 40 45 Ensayo de calvas virales. Monocapas de células MARC-145 en placas de 6 pocillos se infectaron con sobrenadante de células (en diluciones al décuplo) procedente de células MARC-145 transfectadas o infectadas por incubación a la temperatura ambiente durante 1 hora. Las monocapas infectadas se lavaron subsiguientemente una sola vez con EMEM nuevo/10% FBS, se cubrieron inmediatamente con Agarosa SeaPlaque estéril al 1% (BioWhittaker Molecular Applications, Rockland, Maine) en 1X MEM (Sigma M4144)/10% FBS/2% (p/v) NaHCO₃/1X glutamina/1X aminoácidos no esenciales/10 mM HEPES/2% (v/v) gentamicina, y se incubaron a 37°C/5% CO₂, invertidas, durante 5 días. Después de eliminación cuidadosa de la agarosa, las células se tiñeron con 5% de violeta cristal en etanol al 20% durante 10-30 minutos para visualización del tamaño de las calvas.
- 50 55 Curva de crecimiento viral. Monocapas MARC-145 en matraces T-75 se inocularon con PRRSV parental o recombinante diluido en EMEM exento de suero a una multiplicidad de infección (MOI) de 0,001. Después de 1 hora de fijación a la temperatura ambiente con mezcladura cuidadosa, se eliminaron los inóculos y se lavaron 3 veces las monocapas con EMEM exento de suero. Después del lavado, se añadieron 4 ml de medio completo y los matraces se incubaron subsiguientemente durante hasta 5 días a 37°C, 5% CO₂. Se cosecharon inmediatamente partes alícuotas (0,5 ml) después de la adición de medio (momento puntual de 0 horas) y al cabo de 24, 48, 72, 96 y 120 horas, y se guardaron a -80°C. Se utilizaron diluciones seriadas de las muestras para infectar células MARC-145 nuevas y las células se procesaron luego como se ha descrito arriba. Después de eliminar la agarosa, se visualizaron y contaron las calvas. Los resultados de la curva de crecimiento se expresaron como PFU/ml.
- 60 Inoculación *in vivo* del virus de la progenie. Diez cerdos de 4 semanas de edad de raza y sexo mixtos procedentes de una piara PRRSV-seronegativa se dividieron en 3 grupos, constituidos cada uno por dos animales. El primer grupo recibió 10^{3,5} dosis infecciosas de cultivo de tejido al 50% (TCID₅₀) de virus clonado (pVL-V5, tercer pase en células MARC-145) por ml, el segundo grupo recibió 10^{5,4} TCID₅₀ por ml de la cepa del virus parental VR-2332 (cuarto pase en células MARC-145), y el tercer grupo se inoculó falsamente con EMEM. Todos los animales

recibieron 2 ml de inóculo por inyección intramuscular. Los animales se mantuvieron en salas separadas a lo largo del experimento y se observaron diariamente respecto a síntomas clínicos. Todos los cerdos se sacrificaron por eutanasia el día 28 después de la infección. Para recuperar el virus, las muestras individuales de suero se diluyeron 5 veces con EMEM incompleto y se pusieron sobre monocapas nuevas de MARC-145 durante 1 a 2 horas a la temperatura ambiente con agitación cuidadosa. Se retiraron luego los inóculos y se añadió EMEM completo. Las células infectadas se incubaron a 37°C, 5% de CO₂ y se observaron diariamente. Una vez que fue evidente el CPE, los sobrenadantes de células infectados se congelaron a -80°C hasta su caracterización posterior.

Análisis por Transferencia Northern. Transcritos de pVR-V6G7475A se transfectaron en células MA104 y se sometieron luego a pases sobre células nuevas durante varios pases. Para el análisis por transferencia Northern subsiguiente, los sobrenadantes del pase 1 (P1), P3, P6, P8 y P10 se diluyeron en una ratio 1:50 y se utilizaron luego para infectar células (1 ml/matraz T75) el mismo día. Al mismo tiempo, se diluyó al décuplo suero de porcino infectado y se utilizó luego (1 ml) para infectar un matraz T75 separado. El efecto citopático se observó el día 3 p.i. para todos los matraces. El RNA intracelular se extrajo utilizando un Kit RNeasy Midi (Qiagen) y se sometió a electroforesis (15 µg/muestra) en un gel desnaturizante de glioxal como se ha descrito previamente (Nelsen et al., *J. Virol.*, 73:270-80 (1999)). Como control se corrió RNA transcrito de pVR-V6G7475A (100 ng). Después de transferencia del RNA a Membrana de Transferencia de Nailon Magnagraph de 0,45 micrómetros (Osmonics), se sondó la membrana con oligonucleótido marcado/1a-p222, marcado en un extremo con γ -³²P-ATP (Amersham) utilizando polinucleótido-quinasa (Promega) como se ha descrito previamente (Nelsen et al., *J. Virol.*, 73:270-80 (1999)).

Análisis de la secuencia de ácido nucleico del virus de la progenie. Se realizó una amplificación rápida de 5' y 3' de los extremos del cDNA (RACE) con el Kit de Amplificación de cDNA SMART™ RACE (BD Bioscience) o el Juego 5' ó 3'- Full Race Core (TaKaRa Bio Inc.) en RNA viral aislado con el Mini Kit de RNA Viral QIAmp® (Qiagen). La secuencia de nucleótidos remanente se determinó a partir de los productos RT-PCR de pares de cebadores desarrollados para cubrir el genoma entero de la cepa VR-2332 (Tabla 3), como se ha descrito previamente (Yuan et al., *Virus Res.*, 79:189-200 (2001)). Los productos se presentaron para determinación de la secuencia de ácido nucleico en el Centro de Análisis Genético Avanzado en la Universidad de Minnesota. La secuencia viral completa con al menos recubrimiento triple se ensambló inicialmente con el paquete SeqMan del software de análisis de secuencias Lasergene® (DNASTAR, Inc.), y se analizó ulteriormente utilizando el software GCG Wisconsin Package versión 10.3 (Accelrys Inc.). Se utilizaron la cepa VR-2332 (Acceso a GenBank U87392), la cepa Ingelvac® MLV (Acceso a GenBank AF066183) y el clon de cDNA pVR-HN (Acceso a GenBank AY150564; Nielsen et al., *J. Virol.*, 77:3702-3711 (2003)) en todas las comparaciones de nucleótidos para cepas de virus recombinantes.

Resultados

Modificación del vector pOK12. pOK12 (Acceso a GenBank AF223639; Vieira et al., *Gene*, 100:189-194 (1991)), un vector de clonación de bajo número de copias, se modificó por digestión con SmaI (sitio de enzima en el par de bases 263 en pOK12) y Sall (sitio en pb 307) e inserción del fragmento de 244 pb SmaI-Sall del vector 2.0 (7) que contenía la ribozima del virus de la hepatitis delta (HDV). El vector (pOK12HDV) se modificó ulteriormente por mutagénesis de un sitio KpnI existente (sitio pOK12 HDV en el pn 273) para insertar un sitio de enzima de restricción PacI mediante el uso del par de cebadores 5'-pOK12 HDV-257SphIPacI/3'-pOK12HDV-257SphIPacI. La ribozima de HDV Se añadió para proporcionar una escisión eficaz exactamente en el extremo 3' del tramo políA. Los estudios realizados revelaron que la modificación no era necesaria para obtener virus de progenie infeccioso.

Construcción de clones de cDNA de longitud total. La estrategia de clonación se representa en la Figura 2. Cuatro fragmentos de genoma solapantes se amplificaron a partir de RNA viral purificado de VR-2332 por RT-PCR utilizando los pares de cebadores indicados (Figura 2, Tabla 3). Cada fragmento se clonó individualmente en el vector pCR® 2.1-TOPO® para generar el clon intermedio pCR-SphI-FseI (segmento I), pCR-FseI-AvrII (segmento II), pCR-AvrII-BsrGI (segmento III), y pCR-BsrGI-PacI (segmento IV). Los clones de cDNA se digirieron luego con dos enzimas de restricción singulares como se indica por el nombre del clon. Los cuatro fragmentos se purificaron en gel y se ligaron escalonadamente al vector pOK12HDV-PacI para generar un clon de cDNA de longitud total de PRRSV (pVR-V4). En el clon de cDNA de longitud total, la secuencia genómica viral estaba dirigida por el promotor de RNA-polimerasa T7 y seguida por una cola de ácido poliadenílico de 50 nucleótidos. Los transcritos de RNA del clon pVR-V4 no exhibían la infectividad típica del PRRSV cuando se transfectaron en células permisivas, aunque pudo detectarse RNA viral a lo largo de varios pases. Cuando se comparó con la cepa VR-2332, se detectaron un total de 45 mutaciones de nucleótidos (Tabla 4) que conducían a 21 cambios de aminoácidos (Tabla 5), aunque varias mutaciones eran iguales que las identificadas previamente en Ingelvac® MLV (Yuan et al., *Virus Res.*, 61:87-98 (1999)).

Tabla 5. Diferencias de aminoácidos entre las cepas PRRSV y los clones infecciosos de VR-2332. Se muestran únicamente las posiciones en las cuales se observaron diferencias de nucleótidos con la posición del aminoácido correspondiente dentro de la región genómica identificada. Los aminoácidos que se representan en la cepa VR-2332 se muestran en recuadros sin sombreado y las identidades de los aminoácidos de los clones infecciosos con VR-2332 se representan con recuadros en blanco. El texto en cada recuadro individual representa cambios silenciosos o cambios de aminoácidos debido a las diferencias de nucleótidos que se muestran en la Tabla 2. Los recuadros con sombreado claro representan diferencias de nucleótidos que son exclusivas para el clon infeccioso, los recuadros con sombreado mediano hacen resaltar aquellos nucleótidos que se aprecian también en Ingelvac® MLV, y los recuadros con sombreado negro indican nucleótidos exclusivos de porcino. Los aminoácidos separados por barras oblicuas indican números de aminoácidos ORF2a/ORF2b. Las regiones que no se secuenciaron se indican por una barra oblicua.

Posición NT	Posición AA	Región	VR-2332	V4	V5	V5-1-P3	V5-2-P3	V5-Sw612	V5G7475A	V6	V6G7475A	VR-HN	MLV
258	23	NSP1α	V									Silencioso	
309	40		Q	Silencioso	Silencioso	Silencioso	Silencioso	Silencioso	Silencioso	Silencioso	Silencioso		
642	151		P	Silencioso	Silencioso	-	Silencioso	Silencioso	Silencioso	Silencioso	Silencioso		
784	199	NSP1β	V		-								
827	213		A		-			V					
1074	295		Y		-							Silencioso	
1107	306		L	Silencioso	-		Silencioso	Silencioso	Silencioso	Silencioso	Silencioso		
1181	331		S		-								F
1379	397		A		-			V					
1595	469	NSP2	A	D	-								
2192	668		S		-								N
3040	951		D		-								N
3457	1090		D		-								N
4407	1406		P	Silencioso	-		Silencioso	Silencioso	Silencioso	Silencioso	Silencioso		
4593	1468		Q	Silencioso	-		Silencioso	Silencioso	Silencioso	Silencioso	Silencioso		
4681	1498	NSP3	S	A	A	-	A	A	A	A	A	A	A
4866	1559		V	Silencioso	-		Silencioso	Silencioso	Silencioso	Silencioso	Silencioso		
5097	1636		R	Silencioso	-		Silencioso	Silencioso	Silencioso	Silencioso	Silencioso	Silencioso	Silencioso
5247	1686		V	Silencioso	-		Silencioso	Silencioso	Silencioso	Silencioso	Silencioso		
5519	1777		T		-							1	
5610	1807		L		-							Silencioso	
6345	2052	NSP5	P		-								Silencioso

6853	2222	D			-	Silencioso G	Silencioso G	Silencioso				N	
6966	2259	D			-	Silencioso	Silencioso	Silencioso				Silencioso	
7329	2380	K	Silencioso G	Silencioso	Silencioso G	Silencioso	Silencioso	Silencioso	Silencioso				
7475	2429	E											
7554	2455	V	Silencioso	Silencioso	Silencioso	Silencioso	Silencioso	Silencioso	Silencioso			Silencioso	
9220	3011	L	P		-								
9649	3154	G	E		-								
9918	3244	L			-								Silencioso
9958	3257	G	E		-	G/E					E	E	
10040	3284	V	Silencioso		-							H	
10533	3449	Y			-								
10643	3485	V			-							Silencioso	
10697	3503	A	Silencioso		-						Silencioso	Silencioso	
10739	3517	H	Silencioso		-								
10781	3531	T	Silencioso			-						Silencioso	Silencioso
10803	3539	C	R		-							Silencioso	Silencioso
10895	3569	D			-							Silencioso	Silencioso
11055	3623	S	T		-						T	T	
11081	3631	P	Silencioso		-						Silencioso	Silencioso	
11221	3678	G	E		-	E					E	E	
11229	3681	V			-							L	
11259	3691	R	G		-								
11327	3738	H	Silencioso		-								
11329	3739	G	A	A	-	A	A	A	A	A	A	A	A
11501	3771	E	Silencioso		-								
11666	3826	P	Silencioso		-						Silencioso	Silencioso	
11744	3852	W	C		-								
11882	3898	K	Silencioso		-								
12076	2	K	E		-							F/Y	
12102	10/9	LD			-								
12153	27	P/I	P/V		-								
12432	120	E	Silencioso		-								
12501	143	D	E		-								
12600	176	G			-							Silencioso	E
12943	83	G			-							Silencioso	Silencioso
12950	85	D			-								

12973	93		M	R		-													
13011	106		G			-											S		S
13825	13		R			-													Q
13860	25		F		L	-	L	L	L										
14238	151	ORF5	R			-													G
14336	183		G		Silencioso	-	Silencioso	Silencioso	Silencioso	Silencioso	Silencioso	Silencioso	Silencioso	Silencioso	Silencioso	Silencioso			
14404	10		H		Silencioso	-	Silencioso	Silencioso	Silencioso	Silencioso									
14420	16		Q			-													F
14686	104	ORF6	L			-	Silencioso												
14735	121		R	G	G	-	G	G	G	G	G	G	G	G	G	G	G	G	G
14737	121		R	G	G	-	G	G	G	G	G	G	G	G	G	G	G	G	G
14979	31	ORF7	A	TT	T	-	T	T	T	T									

Dado que muchas mutaciones en pVR-V4 ocurrían en la región crítica que codifica helicasa, polimerasa y otros restos de Nidovirus supuestos (Figura 3, Tabla 4), se generaron clones adicionales del segmento genómico III (pCR-AvrII-BsrGI) y se secuenciaron en su totalidad. Después de reemplazamiento del segmento III de pVR-V4 con el fragmento más exacto de la secuencia obtenido, se determinó de nuevo la secuencia de nucleótidos del clon de longitud total entero del genoma (pVR-V5). Excepto en la región reemplazada y en cuatro mutaciones espontáneas (nucleótidos 1595, 13860, 14336, y 14404), estos dos clones genómicos eran idénticos (Tabla 4). El análisis de la secuencia de pVR-V5 demostró que este clon albergaba un total de 23 mutaciones comparado con la cepa VR-2332. De estos 23 cambios, únicamente 8 mutaciones de nucleótidos codificaban un cambio en aminoácido y 5 de las mutaciones de residuos de aminoácido eran idénticas a Ingelvac® MLV y por tanto no se predijo que afectarían desfavorablemente a la replicación *in vitro* (Tabla 4).

El clon pVR-V6 se derivó de mutagénesis orientada del segmento genómico IV para reparar los nucleótidos 13860 y 14979 utilizando los cebadores 13860C2T/ y 14979A2G/, respectivamente. La mutación de estos dos nucleótidos corregiría el residuo de aminoácido 25 de GP5 (L→F) y el residuo 31 de la proteína de la nucleocápsida (T→A). El análisis de la secuencia del clon pVR-V6 confirmó que los nucleótidos se habían corregido de nuevo a nucleótidos de tipo salvaje (*wt*) VR-2332 y no habían dado como resultado ningún otro cambio de nucleótido en ninguna otra parte del genoma cuando se compararon con pVR-V5 (Tablas 4 y 5). Finalmente, la mutagénesis orientada en el segmento III del genoma utilizando el oligómero 7475 G2A se completó tanto en pVR-V5 como en pVR-V6 a fin de corregir una alteración de *wt* VR-2332 en el nucleótido 7475. El cambio de G→A en el nucleótido 7475 dio como resultado una glicina (G) en el aminoácido 2429 de ORF1 en los dos clones recombinantes al ácido glutámico (E) observado en la cepa viral VR-2332 parental. Los dos clones finales, pVR-V5G7475A y pVR-V6G7475A se secuenciaron de nuevo en su totalidad y se encontró que tenían sólo (nucleótido 7475) alterado respecto a los plásmidos recombinantes originales pVR-V5 y pVR-V6, respectivamente (Tabla 5). pVR-V6G7475A contiene así 11 cambios de nucleótidos y ningún cambio de aminoácido con respecto a la cepa VR-2332, además de los observados también en Ingelvac® MLV.

Como puede verse esquemáticamente en la Figura 3 para el constructo final (pVR-V6G7475A), y se detalla en las Tablas 4 y 5, todos los clones de longitud total poseen todavía cambios de nucleótidos dispersos en todo el genoma, fundamentalmente en las regiones de ORF1 definidas deficientemente. Sin embargo, el gran grupo de cambios de nucleótidos en ORF1b que evitaban supuestamente que pVR-V4 completase la replicación viral se repararon en versiones posteriores de los clones del genoma de longitud total. Sólo quedaba una mutación de nucleótido (nt 11329 que codificaba la mutación G3739A) en ORF1b de pVR-V5 y clones posteriores, y esta mutación no evita que Ingelvac® MLV infecte y se replique eficientemente en células cultivadas. Las Tablas 4 y 5 muestran también la información de residuos para el clon infeccioso previamente publicado, pVR-HN (Nielsen et al., *J. Virol.*, 77:3702-3711 (2003)), que se ha demostrado se replica en animales. Existe un aumento sustancial en el número de residuos en pVR-HN (15 nucleótidos) que exhiben directamente la secuencia de Ingelvac® MLV en el constructo final, pVR-V6G7475A (7 nucleótidos).

Caracterización del virus recombinante. Se produjeron transcritos de RNA de longitud total de cada clon de cDNA. La transfección de células MARC-145 con los transcritos de cDNA o RNA viral de *wt* VR-2332 (vRNA) dio como resultado CPE, caracterizado por aglomeración de células seguida por lisis, a las 48 a 72 horas después de la transfección. Los CPE causados por los transcritos recombinantes estaban retardados y eran algo distintos comparados con el inducido por el vRNA de VR-2332 *wt*, en el cual el CPE se presenta como agregación enérgica, desprendimiento, y disgregación. A las 96 horas de la transfección, la mayoría de las células transfectadas con vRNA de VR-2332 habían sufrido lisis y se habían desprendido de la placa, mientras que se apreciaba un CPE menos severo en las células transfectadas con los transcritos de RNA clonado derivados *in vitro*.

El virus (P0) se cosechó de las células transfectadas y se utilizó una parte alícuota (10 µl diluidos hasta 1 ml en medio de cultivo) para infectar células MARC-145 para amplificación del virus de la progenie. Después de la detección del CPE, el virus (P1) se cosechó de nuevo y se utilizó una parte alícuota para reinfección de células MARC-145. El virus recombinante en el sobrenadante de las células (P2) se utilizó para purificación del RNA viral, que se empleó luego para obtener fragmentos RT-PCR con los pares de cebadores 5'-6800/3'-ORF1b (nt 6796-7614) y P51/05P4 (nt 13757-14341). Los fragmentos PCR obtenidos se presentaron para análisis de la secuencia de nucleótidos a fin de confirmar que la infectividad observada era debida a transcritos de RNA de longitud total transfectados del constructo infeccioso y no un resultado de contaminación debida al virus *wt*. Se observaron mutaciones de nucleótidos en los residuos 7329, 7475, 7554, y diferencias en el nucleótido 13860 en el virus de la progenie de pVR-V5, y 7329, 7554, y 13860 se detectaron en el virus de pVR-V5G7475A. Análogamente, se detectaron mutaciones en los residuos 7329, 7475 y 7554 en la progenie pVR-V6 y se detectaron mutaciones en 7329 y 7554 en el virus resultante de pVR-V6G7475A (Tabla 4 y 5). No se observaron mutaciones correspondientes en el virus P2 procedente de las transfecciones de vRNA *wt*.

Análisis de inmunofluorescencia de los virus recombinantes. Se utilizaron ensayos de inmunofluorescencia directa para detectar la expresión de la proteína de la nucleocápsida de PRRSV en células MARC-145 infectadas. Todas las células infectadas por transcritos de virus recombinante (P2 y siguientes) así como vRNA eran positivas por este método. Era claramente evidente una acumulación nucleolar masiva de la proteína de la nucleocápsida, como había sido consignado previamente por Rowland et al. (*Virus Res.*, 64:1-12 (1999)).

Infección *in vivo* con virus recombinante derivado de pVR-V5. Los virus recombinantes recuperados de P3 de células MARC-145 transfectadas con transcritos de RNA del clon de cDNA pVR-V5 se inocularon en cerdos jóvenes en paralelo con VR-2332 *wt*, virus de la vacuna Ingelvac® MLV y solución salina (control negativo). Se recogieron muestras de sangre los días 0, 3, 5, 7, 14, 21 y 28 p.i. y se analizaron respecto a seroconversión por HerdChek PRRS 2XR ELISA (IDEXX) y respecto a recuperación de virus. El día 28, todos los animales infectados habían experimentado seroconversión con aproximadamente la misma cinética, revelando que los virus recombinantes pVR-V5 se replicaban satisfactoriamente *in vivo* (Figura 4). Los síntomas clínicos estaban ausentes de todos los animales durante el curso del experimento, pero esto no era una sorpresa, dado que la cepa VR-2332 *wt* no produce en muchos casos enfermedad manifiesta en los cerdos jóvenes y da como resultado ganglios linfáticos engrosados sólo transitoriamente, típicamente hacia el día 14 p.i.

Se utilizó una muestra de suero de un animal infectado con progenie de pVR-V5 (Sw612), tomada a los 14 días p.i., para infectar monocapas recientes de MARC-145 para recuperación de virus recombinante sometido a pases *in vivo*. Como se ha descrito anteriormente, el virus derivado de transfección *in vitro* de transcritos de RNA del clon pVR-V5 causaba sólo CPE mínimo (evidenciado por agregación de células infectadas) mientras que el virus recuperado del suero del día 14 del animal de test causaba CPE típico (agregación de células, desprendimiento, y disgregación) a las 96 horas de la infección. Esto sugería que se había producido un cambio en el genotipo o fenotipo viral mientras pVR-V5 se replicaba *in vivo*.

Con objeto de dilucidar la razón del cambio aparente de fenotipo, se completó el análisis de la secuencia genómica total en virus recombinado de un cerdo (Sw612) y se sometió luego a pases una sola vez en células MARC-145 para amplificar la progenie de Sw612 (Figura 3, Tablas 4 y 5). Cuando se comparó con el virus utilizado para infectar cerdos, pVR-V5, se retenían en Sw612 17 cambios de nucleótidos específicos del clon de cDNA infeccioso, algunos de los cuales se observan también en Ingelvac® MLV (7/17 nucleótidos). Los dos residuos G no virales seguidos por un residuo T presente en el extremo 5' del transcrito del clon pVR-V5 original no se apreciaban en el virus derivado de la infección *in vivo*. Se observaba degeneración en las posiciones de los nucleótidos 9958 (R), 14336 (Y) y 15411 (Y). El nucleótido semejante a VR-2332 *wt* (G) en la posición 9958 exhibía degeneración con un nucleótido semejante a Ingelvac® MLV (A). Este cambio da como resultado una mutación de un residuo glicina a un residuo ácido glutámico, respectivamente (Tabla 2). En la posición 14336, se detectó degeneración como una base específica de clon infeccioso (C) y una base específica de VR-2332 *wt* (T), lo que reflejaba una mutación silenciosa.

Había otra mutación (nt 7475) en la cual un residuo G había cambiado al residuo A *wt*. Sin embargo, había otras cinco diferencias de nucleótidos (nt 102, 827, 1379, 14686 y 15411) no apreciadas en ninguno de los otros virus de este estudio. El nucleótido 102 está localizado en la secuencia conductora, que se cree no se traduce. No obstante, si se tradujera la secuencia conductora, el ORF codificado (nucleótidos 1-100 de VR-2332) se extendería en un solo residuo de aminoácido (W). Las mutaciones en los residuos 827 y 1379 conducían a mutaciones en OFR1a, dando como resultado en ambos casos un cambio de aminoácido de la alanina codificada por VR-2332 *wt* por una valina Sw612. El residuo guanina en el nt 7475 de pVR-V5 había mutado a adenina *wt*. Esto daba como resultado una mutación de un aminoácido no conservadora G3294A, que está basada en producto de escisión nsp7 de proteasa predicho de ORF1a, y esta región genómica no tiene función alguna definida hasta la fecha. El nucleótido 14686, localizado en ORF6, exhibía un cambio desde una guanina VR-2332 *wt* a una alanina en Sw612, que codifica todavía el aminoácido glicina. El otro cambio de nucleótido singular ocurría exactamente en el extremo 3' de la secuencia viral (nt 15411), antes del comienzo de la cola poliA. En este caso, un residuo timina previamente conservado revelaba degeneración con un residuo citosina. Estos cambios genéticos, aunque informativos, no revelaban inmediatamente la o las causas del cambio en el fenotipo de crecimiento observado. Sin embargo, ello revelaba de hecho la naturaleza errante de la replicación de PRRSV *in vivo* y sugiere que una secuencia genómica viral moderadamente diferente de VR-2332 *wt* era capaz de replicarse eficazmente (Figura 3).

Comparación del tamaño de la calva viral. Se completaron determinaciones del tamaño de las calvas de los virus recombinantes así como de VR-2332 *wt* en paralelo en células MARC-145 a las 120 horas p.i. (Fig. 5A). La cepa VR-2332 formaba calvas que tenían un tamaño medio de 3 mm, mientras que la progenie del pase 3 del clon de cDNA de pVR-HN formaba calvas ligeramente menores (2,5 mm como promedio). En contraste, se obtuvieron únicamente calvas minúsculas de los virus recombinantes derivados de pVR-V5 y pVR-V6, y éstas eran sólo fácilmente apreciables por examen al microscopio (Fig. 5A). Los virus recombinantes recuperados de los clones pVR-V5G7475A y pVR-V6G7475A formaban, por término medio, calvas de 1,5 mm y 2 mm respectivamente. Sin embargo, en otro ensayo, las calvas producidas por la progenie viral (Sw612) recuperadas de la infección *in vivo* de virus recombinante derivado de VR-FLV5 eran mucho mayores, aproximadamente iguales en tamaño y en número a las derivadas de VR-2332 *wt* (Fig. 5B).

Quedaban sólo volúmenes mínimos de los sobrenadantes de células que contenían cada virus recombinante. Por tanto, a fin de examinar detalladamente el papel del cambio de nucleótido en la determinación del tamaño de las calvas, se transfectaron transcritos nuevos de RNA producidos a partir de pVR-V5, pVR-V6, pVR-V5G7475A y pVR-V6G7475A en células MARC-145 (denominado segundo linaje). Los virus de la progenie del pase 3 de cada clon infeccioso a los 5 días post-infección se analizaron de nuevo respecto a tamaño de las calvas en comparación con VR-2332 *wt*, VR-HN y los virus Sw612. En contraste con el ensayo de calvas previo, los tamaños de todas las calvas aparecían similares, yendo únicamente los virus recombinantes obtenidos de pVR-V5, pVR-V6, pVR-V5G7475A ligeramente menores que los virus *wt* VR-2332, Sw612 y pVR-V6G7475A obtenidos *in vivo* (Fig. 6A). Sin embargo,

los virus recombinantes no mimetizaban todavía directamente la infección viral auténtica, como se muestra por los títulos aproximadamente 10 veces más bajos cuando se comparaban con *wt* VR-2332 o con virus pVR-V5 recombinante que se había sometido a pases en cerdos (Sw612) (Fig. 6B).

Análisis de la Secuencia de Nucleótidos de las Preparaciones de Virus del Primer y Segundo linajes. El análisis limitado de la secuencia de nucleótidos (debido a limitación del stock de virus) de virus del pase 3 derivado de pVR-V5 inoculados en cerdos (V5-1-P3) y el análisis completo de la secuencia de nucleótidos del virus derivado del pase 3 de pVR-V5 (V5-2-P3) se completaron a fin de revelar la razón genética de las discrepancias en el tamaño de las calvas. Tales análisis revelaron que los dos virus V5 preparados independientemente diferían en secuencia en el extremo 5' (Tabla 4). El virus que había producido calvas minúsculas (V5-1-P3) no tenía nucleótido extraño alguno en el extremo 5', como se muestra en la secuencia de nucleótidos de la cepa *wt* VR-2332, mientras que el que producía calvas de mayor tamaño (V5-2-P3) poseía 4 residuos timidina sin matriz en el término 5' (Tabla 4). La secuencia de nucleótidos viral V5-1-P3 restante que pudieron obtener los autores de la invención coincidía exactamente con la del virus V5-2-P3, así como la del clon parental. No obstante, el análisis de la secuencia completa del virus V5-2-P3 revelaba que el virus exhibía degeneración de nucleótidos en varios sitios genómicos. Se obtuvieron resultados similares cuando se analizaron regiones limitadas de los virus del segundo linaje VR-FLV5G7475A-P3 y VR-FLV6G7475A-P3. Estas dos últimas infecciones de la progenie de clones exhibían términos 5' diferentes al mismo tiempo que exhibían degeneración en su secuencia.

Curvas de Crecimiento Viral. Se completaron simultáneamente determinaciones de la curva de crecimiento viral en un solo paso utilizando células MARC-145 y virus del pase 3 (segundo linaje) (Fig. 7). Los virus recombinantes recuperados de pVR-V5, pVR-V5G7475A, pVR-V6, pVR-V6G7475A y pVR-HN exhibían tasas de crecimiento viral en un solo paso similares, pero sus picos de replicación eran todos ellos significativamente más bajos que la cepa *wt* VR-2332 y Sw612, la progenie *in vivo* de pVR-V5. Asimismo, las tasas de replicación de las preparaciones de virus recombinantes derivadas de pVR-V5, pVR-V6 y pVR-HN se reducían un tanto en comparación con el virus derivado de pVR-V5G7475A y pVR-V6G7475A. Los dos últimos clones infecciosos codifican diferencias tan pequeñas como de 13 y 11 nucleótidos, respectivamente, dando como resultado 2 y 0 cambios de aminoácido con respecto a la secuencia de *wt* VR-2332 además de los cambios observados en Ingelvac® MLV. Estos datos revelan por tanto que los virus con un cambio tan pequeño como de 11 nucleótidos respecto a *wt* VR-2332 y su progenie atenuada Ingelvac® MLV están algo deteriorados en replicación. Correspondientemente, los títulos resultantes de los virus *wt* VR-2332 y SW612 eran aproximadamente 6-15 veces mayores que el de los virus recombinantes que no se habían sometido a pases en cerdos (Fig. 7).

Análisis Northern de vRNA. Se ha demostrado que las especies sgRNA deficientes de PRRSV, identificadas previamente como RNAs subgenómicos heteróclitos (latín: formas no comunes), son un constituyente de la infección de PRRSV y no pueden separarse de los genomas virales de longitud total por métodos estándar tales como pase de células cultivadas a multiplicidades de infección bajas o centrifugación en gradiente de sacarosa (Yuan et al., *Virology*, 275:158-169; 30 (2000); Yuan et al., *Virus Res.*, 105:75-87 (2004)). Para explorar si se producen o no heteróclitos de PRRSV durante la transcripción *in vitro* de los clones del genoma de cDNA de longitud total o aparecen después de transformación/infección subsiguiente, se completó el análisis por transferencia Northern. El transcrito de RNA de longitud total y los pases 1, 3, 6, 8 y 10 del virus producido a partir de las células MA-104 transfectadas se utilizaron para inocular matraces T-75 nuevos de células M-104 con 100 µl de sobrenadante diluido 1:100, así como suero Sw612 diluido 1:10 (2 ml total/matraz). Después de 4 días, se cosechó RNA intracelular de PRRSV y se separaron 15 µg de cada preparación por electroforesis a través de un gel de agarosa desnaturante, y se transfirieron a una membrana de nailon. Después de reticulación del RNA, la membrana se hibridó con una sonda radiomarcada con ³²P complementaria al extremo 5' de ORF1a que selecciona genomas de VR-2332 de longitud total así como heteróclitos (1a-222; 29). Como se muestra en la Figura 8, el transcrito de RNA es en su mayor parte una sola banda, que migra como vRNA de longitud total, mientras que las especies de RNA de PRRSV del pase 1 y posteriores migran a la vez como especies de longitud total y especies de tamaño subgenómico identificadas previamente como heteróclitos. Adicionalmente, la fuerza de hibridación aumenta a lo largo de los pases. Dado que el virus se cosechó de un volumen igual de sobrenadante de células infectadas en el mismo momento, esta solución sugiere que el vRNA se hace más eficiente con las replications a lo largo del tiempo. Por último, cuando se compara el virus generado por Sw612 con el virus generado por cultivo de células, el patrón de bandas de RNA es indistinguible, lo que sugiere fuertemente que las especies de RNA defectivas se forman fácilmente y se replican tanto *in vitro* como *in vivo* y son por tanto una parte natural de la infección de PRRSV.

Discusión

En teoría, un clon de cDNA infeccioso de un virus debería ser idéntico a la secuencia parental a fin de generar un sistema genético inverso que mimetice la infección del tipo salvaje. Se ejerció un esfuerzo considerable para reproducir un genoma de VR-2332 de la cepa PRRSV de fidelidad total, pero debido a mutaciones espontáneas impredecibles en varios sitios, no se ha tenido éxito todavía en la derivación de un clon infeccioso que no presente diferencia alguna respecto a la cepa VR-2332 *wt* secuenciada en el laboratorio de los inventores. Las DNA-polimerasas de alta fidelidad, utilizadas en este estudio, están disponibles para reducir las mutaciones artificiales, pero dicha mutación no puede evitarse durante la transcripción inversa (Malet et al., *J. Virol. Methods*, 109:161-70 (2003)). Adicionalmente, el hecho de que PRRSV exhibe evolución viral y variación de cepas asombrosa (Chang et al., *J. Virol.*, 76:4750-6 (2002); Murtaugh et al., *Adv. Exp. Med. Biol.*, 440:787-94 (1998); Yoon et al., *Adv. Exp. Med.*

Biol., 494:25-30 (2001)), se recombina fácilmente a alta frecuencia dando como resultado recombinantes intergénicos entre cepas (Yuan et al., *Virus Res.*, 61:87-98 (1999)), sufre recombinación intragénica para formar RNAs subgenómicos de PRRSV y heteróclitos (Nelsen et al., *J. Virol.*, 73:270-8025 (1999); Yuan et al., *Virology*, 275:158-169 (2000); Yuan et al., *Virus Research*, 105:75- 87 (2004)) y exhibe a menudo degeneración de nucleótidos en sitios nucleotídicos impredecibles en aislados de campo hacen que esta meta inicial consuma mucho tiempo y conduzca a una ganancia insignificante. Un constructo infeccioso de DNA que posee tan pocas como 11 mutaciones de nucleótidos, en comparación con la cepa VR-2332, fuera de los dominios que se sabe están implicados en la replicación viral (extremos 5' y 3', ORF1b) se consideró suficiente para la producción de virus *wt* y los objetivos aguas abajo del uso de clones infecciosos para interrogantes sobre patogénesis y estudios estructura:función. pVR-HN es más similar a Ingelvac® MLV en la región del virus que codifica el resto helicasa (nsp10). Adicionalmente, la comparación patogénica de estos dos clones infecciosos puede arrojar luz acerca de las diferencias entre la cepa parental, VR-2332, y su progenie de la cepa de vacuna, Ingelvac® MLV.

Puede deducirse información valiosa acerca de la construcción y evaluación de los clones infecciosos para la cepa VR-2332 de PRRSV. En primer lugar, la cepa VR-2332 de PRRSV no puede tolerar todas las mutaciones para supervivencia. Mutaciones particulares de nucleótidos o aminoácidos pueden favorecer o dificultar la replicación viral, y el reto consiste en averiguar cuáles son letales para la supervivencia. En el clon pVR-V4, que no producía viriones infecciosos, había un total de 42 diferencias de nucleótidos con respecto a la cepa parental *wt* VR-2332. En estos 42 cambios de nucleótidos, varios nucleótidos dan como resultado mutaciones silenciosas (20 residuos) o existen en otras cepas de PRRSV conocidas (9 mutaciones de residuos de aminoácidos mimetizan directamente Ingelvac® MLV), lo que conduce a la predicción de que estos cambios pueden ser no letales para la replicación del virus. Se predijo así que once cambios de nucleótidos que conducían a 12 cambios de aminoácidos y dos mutaciones de nucleótidos 3' UTR, no observadas cada una en Ingelvac® MLV, son letales para la cepa VR-2332 de PRRSV. En pVR-V5 y constructos posteriores, se corrigieron 19 cambios, que incluían varias mutaciones silenciosas y 9 cambios aberrantes de aminoácidos no observados en el genoma de Ingelvac® MLV y otros 8 cambios observados en la cepa de la vacuna. Esto condujo a la primera evidencia de que los constructos eran infecciosos, aunque en pVR-V5 estaban todavía presentes dos mutaciones de aminoácidos, una de las cuales se alteró por mutagénesis orientada para producir pVR-V6. El cambio de aminoácido restante se reparó en pVR-V5G7475A y pVR-V6G7475A, aunque estos clones albergan todavía mutaciones silenciosas que no se encuentran en la cepa VR-2332 y la cepa de vacuna derivada.

De este estudio se obtuvieron varias observaciones singulares. Ante todo, cada linaje de virus producido puede dar como resultado una secuencia terminal 5' singular que no se detectaba en la cepa *wt* VR-2332. Los autores de la invención tampoco pueden correlacionar todavía el tamaño de las calvas con la secuencia de nucleótidos. En segundo lugar, se observaron cambios de nucleótidos singulares después de replicación en cerdos, lo cual puede reflejar la naturaleza inherente de la polimerasa de PRRSV. Todos los cambios de nucleótidos eran de naturaleza transitoria y no exhibían sesgo (5 A/G y 4 C/T). Aunque la reversión G A en el nucleótido 7475 se observó después del pase *in vivo*, no se pudo correlacionar este sitio con el tamaño de calva incrementado subsiguiente debido a que habían ocurrido otros cambios sin matriz. Adicionalmente, los análisis de la secuencia del genoma total del pase 3 de un virus derivado de V5 que producía calvas de mayor tamaño (V5-2-P3) revelaron una secuencia diferente del terminal 5' del virus V5 portador de calvas minúsculas utilizado para infectar cerdos (V5-1-P3). Sin embargo, puede llegarse a la conclusión de que las mutaciones no eran letales para la replicación del virus, dado que este virus, después de pases en cerdos, producía calvas de tamaño *wt* en células MARC-145 y crecía prácticamente al mismo ritmo que el virus parental (Figs. 5A, 6 y 7).

Presenta un interés considerable el hecho de que el análisis de la secuencia de el tercer pase *in vitro* de V5, V5G7475A y V6G7475A parecía sugerir que el complejo de replicasa de PRRSV permite que ocurran transiciones frecuentes, y transversiones infrecuentes, al tiempo que sufre replicación viral. Esto puede reflejar una replicasa viral que ha evolucionado de tal modo que la misma puede generar nuevas formas genéticas de un genoma de PRRSV y evaluar luego su competencia entre otras variantes, dando como resultado un virus óptimamente "adaptado". Estas observaciones se han anotado también durante el sometimiento a pases secuenciales de PRRSV *in vivo* (Chang et al., *J. Virol.*, 76:4750-63 (2002)). Se están haciendo actualmente esfuerzos de secuenciación para examinar los genomas de longitud total de pases ulteriores, cuando se detecta una replicación más robusta. Por último, está claro ahora que la replicasa de VR-2332 de la cepa PRRSV sintetiza fácilmente heteróclitos al mismo tiempo que está produciendo vRNA de longitud total. Esta cepa prototipo, aislada y caracterizada en 1992, puede ser exclusiva en la adquisición gradual de adaptación de la replicación, dado que otros investigadores que producen clones infecciosos de cepas más recientes no han observado el mismo efecto (Truong et al., *Virology*, 3325:308-319 (2004)). El papel de la formación de heteróclitos y la aparición concomitante de replicación viral vigorosa sugieren que existe un papel ventajoso de los heteróclitos en la evolución del PRRSV.

Ejemplo 2

Muchos aislados virulentos de un PRRSV aparentemente nuevo se identificaron recientemente en el estado de Minnesota, EE.UU. El análisis de la secuencia de nucleótidos de ORF5 y la comparación con la base de datos de PRRSV del Laboratorio de Diagnóstico Veterinario de la Universidad de Minnesota (> 5000 aislados) reveló que los aislados eran del linaje Tipo 2, pero eran significativamente diferentes de los aislados previos. Adicionalmente, los mismos estaban muy estrechamente relacionados con los aislados observados previamente en Canadá a principio

de los años 1990 (Mardassi et al., *J. Gen. Virol.*, 75:681-685 (1994)) y en el estado de Minnesota en 1998. El análisis del polimorfismo de longitud de fragmentos de restricción (RFLP) de ORF5 demostró también que los mismos pertenecían al mismo grupo de virus que estos casos anteriores, conocidos como aislados 1-8-4 (Wesley et al., *J. Vet. Diagn. Invest.*, 10:140-144 (1998)) y se designaron por ello aislados MN184. Debido a la notable disimilitud con todos excepto uno de los aislados de PRRSV MN aislados previamente, dos de estos nuevos aislados se amplificaron una sola vez en macrófagos alveolares de porcino (PAM), la célula hospedadora, y el análisis del genoma de longitud total se completó en los virus, designados como MN184A y MN184B. Estos dos aislados se recogieron en momentos diferentes de dos explotaciones ganaderas separadas.

Materiales y Métodos

- 10 Para secuenciar los aislados MN184, se extrajo RNA viral (vRNA) del sobrenadante de células infectadas PRRSV con el Mini Kit de RNA Viral QIAmp® (Qiagen, Valencia, CA) y se realizó una RT-PCR (kit RT-PCR Qiagen® OneStep). Los cebadores (disponibles a petición) se diseñaron basándose en las secuencias publicadas de diferentes cepas de PRRSV depositadas en GenBank así como la secuencia MN184 de nueva generación. La secuencia de nucleótidos 5' de los dos aislados PRRSV se derivó utilizando el Kit Core RACE 5'-Full (TaKaRa Bio, Madison, WI). La 3' RACE se realizó con el Kit de Amplificación de cDNA SMART™ RACE (Clontech, Mountain View, CA). Los productos de la RT-PCR se purificaron en gel (QIAquick® Qiagen), se clonaron en el vector pGEM-T (Promega, Madison, WI) y se seleccionaron de 3 a 5 clones para cada producto RT-PCR para secuenciación. La determinación de la secuencia de nucleótidos se completó en ambas direcciones con los cebadores PCR específicos o los cebadores de los promotores SP6 y P7 codificados por el vector. Los productos se sometieron al
- 20 Centro de Análisis Genético Avanzado de la Universidad de Minnesota para determinación de la secuencia con un analizador automático de fragmentos de DNA ABI377. Se obtuvo una secuencia de calidad que representaba al menos un recubrimiento triple del genoma. Los datos de secuencia se ensamblaron y analizaron utilizando el programa de análisis de secuencias GeneTool (BioTools Inc., Edmonton, Alberta, CA) y Lasergene (DNASTAR, Madison, Wis.).
- 25 Se generaron alineaciones múltiples de secuencias con CLUSTALX (Thompson et al., *Nucleic Acids Res.*, 24:4876-4882 (1997)) o el Paquete Wisconsin Version 10.3 (Accelrys Inc., San Diego, CA). Las secuencias de PRRSV de longitud total se alinearon utilizando ClustalX (Versión 1.83.1; matriz de pesos IUB DNA, penalidad por laguna 15.00, penalidad por longitud de laguna 6.66). La alineación resultante se analizó ulteriormente analizando el Programa de Distancias del Paquete Wisconsin Version 10.3 (método de distancias Jukes-Cantor, coincidencias parciales debidas a los símbolos degenerados considerados). Para la Figura 10, las secuencias se alinearon con el programa Pileup del paquete Wisconsin (Matriz de Registro Blosum62, Peso de Laguna = 8, Peso de Longitud = 2, Extremos Pesados). La alineación se registró para redundancia y se coloreó para identidad porcentual utilizando Jalview (Clamp et al., *Bioinformatics*, 12: 426-427 (2004)), y se transfirió luego a Adobe® Photoshop® CS, versión 8.0, para transformación en escala de grises. Para la Figura 11, las secuencias se alinearon con el programa Pileup del Paquete Wisconsin (Matriz de Registro Blosum 62, Peso de Laguna = 8, Peso de Longitud = 2, Extremos Pesados). Para la Figura 12, se predijo un péptido señal utilizando el servidor SignalP (Bendtsen et al., *J. Mol. Biol.*, 340:783-795 (2004)). Las regiones transmembranales se derivaron por PHDhtm (Rost et al., *Protein Sci.*, 5:1704-1718 (1996) y los sitios de N-glicosilación potencial se identificaron por PROSITE (Bairoch et al., *Nucleic Acids Res.*, 25:217-221 (1997)) utilizando el servidor PredictProtein (Rost et al., *Nucleic Acids Res.*, 32:W321-W326 (2003)). Las secuencias se alinearon con el programa Pileup del Paquete Wisconsin (Matriz de Registro Blosum 62, Peso de Laguna = 8, Peso de Longitud = 2, Extremos Pesados).
- 40

Resultados

- La alineación genómica demostró que estos dos PRRSV eran muy distintos (> 14,5% de disimilitud de nucleótidos de otros genomas secuenciados de longitud total de Norteamérica Tipo 2, pero la comparación con las secuencias de longitud total Tipo 1 (Europeas) confirmó que los aislados eran exclusivamente de origen de genotipo Tipo 2 dado que eran sólo aproximadamente 59% similares a nivel de nucleótidos a ambas cepas EuroPRRSV y Lelystad. Llamativamente, estos aislados de MN184 Tipo 2 representaban los genomas de PRRSV más cortos detectados hasta la fecha (15019 nucleótidos, sin incluir la cola poliA). Por otra parte, no se discernía área específica alguna que sugiriera que estos aislados se derivaban de recombinación viral entre las cepas Tipo 1 y Tipo 2.
- 50 El análisis de la secuencia de longitud total reveló que los dos aislados MN184 eran en realidad genéticamente distintos. Los mismos compartían 98,0% de semejanza de nucleótidos o 2% de diferencia. Este porcentaje de disimilitud era inesperado debido a su aparición simultánea brusca en Minnesota, sin ningún aislado reciente claramente afín observado en la base de datos de los inventores de PRRSV en dicho momento. La Tabla 6 presenta la comparación detallada de nucleótidos y aminoácidos entre los dos aislados, y la Figura 9 representa las
- 55 diferencias de aminoácidos observadas entre estas dos cepas. Ambos aislados exhibían degeneración de nucleótidos en varias regiones del genoma, predominantemente en la región nsp2 predicha de ORF1 (Tabla 6). El hecho de que se observaba degeneración de nucleótidos en estos aislados sugería que PRRSV puede estar constituido por varias especies individuales, a las que se hace referencia a menudo como un enjambre de secuencias virales afines pero distintas en los animales infectados.

Tabla 6. Análisis detallado de las regiones genómicas individuales de PRRSV y proteínas traducidas, y número de bases degeneradas detectadas en cada región. La degeneración se define como más de un nucleótido detectado para una base particular en archivos huella separados de tres o más archivos traza.

Región	Bases	Longitud en nucleótidos	Similitud de nucleótidos, %	Identidad de nucleótidos, %	Número de Bases Degeneradas (184A/184B)	Longitud en amino-ácidos	Similitud de Aminoácidos, %	Identidad de Aminoácidos, %
5' UTR	1-190	190	99.5	98.9	1/0	-	-	-
ORF1A	191-7309	7119	98.5	96.7	16/109	2372	96.8	96.5
NSP1a	191-688	498	98.8	98.5	1/0	166	97.6	97.6
NSP1b	689-1339	651	98.3	97.5	2/3	217	97.2	95.9
NSP2	1340-3886	2547	98.0	94.6	10/76	849	94.2	94.2
NSP3	3887-5224	1338	98.7	98.7	0/0	446	99.3	98.9
NSP4	5225-5836	612	98.5	96.4	0/13	204	97.1	97.1
NSP5	5837-6346	510	99.2	95.3	3/17	170	97.1	97.1
NSP6	6347-6394	48	100.0	100.0	0/0	16	100	100
NSP7	6395-7171	777	99.3	99.3	0/0	259	99.6	99.2
NSP8	7172-7309	138	99.3	99.3	0/0	46	97.6	97.6

ORF1B	7306-11679	4374	99.2	98.9	5/4	1457	99.5	99.2
NSP9	7288-9225	1938	98.9	98.8	1/1	646	99.4	98.9
NSP10	9226-10548	1323	99.3	98.9	3/3	441	99.8	99.3
NSP11	10549-11217	669	99.3	99.3	0/0	223	99.5	99.5
NSP12	11218-11679	462	99.6	99.4	1/0	153	99.3	99.3
ORF2a/GP2	11681-12451	771	99.0	98.3	1/0	222	98.0	97.3
ORF2b/E	11686-11907	222	99.6	99.6	0/0	73	100	100
ORF3/GP3	12304-13068	765	98.6	98.6	0/0	254	97.6	97.6
ORF4/GP4	12849-13385	537	98.5	98.5	0/0	178	98.9	98.9
ORF5/GP5	13396-13998	603	97.8	97.7	1/0	200	96.5	96.5
ORF6/M	13983-14507	525	99.6	97.4	0/0	174	100	100
ORF7/N	14497-14868	372	98.9	98.9	0/0	123	97.6	97.6
3' UTR	14869-15019	151	100	98.0	1/1	-	-	-

Con objeto de identificar más estrechamente las regiones individuales de estos aislados MN184 que exhibían la disimilitud máxima con respecto a otras cepas PRRSV y asignar la región o regiones que dieran cuenta de la diferencia en la longitud del genoma viral del Tipo 2, se compararon estos dos aislados con la secuencia de la cepa VR-2332 prototipo de Tipo 2. Las diferencias entre los dos aislados pudieron discernirse de nuevo, exhibiendo el aislado MN184B una semejanza ligeramente mayor con la cepa VR-2332 que el aislado MN184A. Las comparaciones de nucleótidos y aminoácidos con VR-2332 demostraron que regiones individuales del aislado MN184 variaban desde 81,5-94,7% y 78,4-100%, respectivamente, pero las regiones correspondientes a ORF5 (86,4-86,7% y 87,0-87,5%, respectivamente) predecían nsp1β (83,8-84,0% y 84,8-85,4%, respectivamente), y nsp2 (81,5-85,5% y 78,4-79,5%, respectivamente) eran las más variables. Era sumamente interesante que únicamente la región genómica nsp2 predicha mostraba una diferencia en longitud de nucleótidos y que ambos aislados MN184 exhibían la misma delección en nsp2, detallada más adelante. La comparación revelaba también que las UTR's 5' y 3' eran las regiones más conservadas del genoma (94,7% y 94,0%, respectivamente), indicando una conservación de secuencia en regiones importantes para la replicación y transcripción viral.

ORF5 codifica una proteína estructural de PRRSV heterogénea (GP5) y se utiliza a menudo para identificación diagnóstica de PRRSV (Kapur et al., *J. Gen. Virol.*, 77:1271-1276 (1996)). GP5 es una proteína con tres dominios transmembranales predichos con un endodominio y un ectodominio. El ectodominio de 30 aminoácidos se compone de un dominio corto altamente conservado que contiene usualmente al menos dos sitios de glicosilación en N limitados por dos regiones hipervariables. Se ha demostrado que el dominio altamente conservado de esta región de 30 aminoácidos codifica el epítipo de fijación viral en las cepas Tipo 2 (Plagemann, *Virology*, 290:11-20 (2001); Ostrowski et al., *J. Virol.*, 76:4241-4250 (2002); Plagemann et al., *Arch. Virol.*, 147:2327-2347 (2002)). GP5 del mismo conjunto de genomas de longitud total, así como los aislados originales RFLP184 identificados en Canadá (IAF-93-653, IAF-Klop) y en 1998-1999 en Minnesota (98-3298, 98-3403, 99-3584) se alinearon (Fig. 10). La alineación de PRRSV GP5 reveló identidades de aminoácidos comprendidas entre 82,5% y 87,7% entre los nuevos aislados MN184 y otras cepas Tipo 2 distintas de RFLP184. Es interesante que las diferencias de aminoácidos entre los nuevos aislados MN184 y los aislados anteriores RFLP184 eran muy grandes (5,7%-12,2%) y por tanto no se detectó ningún origen claro del nuevo virus RFLP184. La alineación limitada demuestra que la mayor parte de las diferencias de aminoácidos observadas se encontraban en las regiones hipervariables (Fig. 10). Los dos sitios de glicosilación en N conservados se mantenían en los aislados MN184, excepto en lo que respecta a la degeneración de nucleótidos detectada que codificaba el aminoácido 44 en el aislado MN184B.

Nsp1β codifica una cisteína-proteasa semejante a papaína (den Boon et al., *J. Virol.*, 69:4500-4505 (1995)). Se completó una alineación de aminoácidos de los aislados MN184 con un conjunto redundante de secuencias nsp1β Tipo 2 disponibles así como las cepas Tipo 1 EuroPRRSV y Lelystad (Fig. 11). La proteína nsp1β posee cierto número de aminoácidos completamente conservados, y los residuos catalíticos propuestos se mantenían en todos los genomas secuenciados (den Boon et al., *J. Virol.*, 69:4500-4505 (1995)). La alineación, ordenada por similitud de aminoácidos, indica que los aislados MN184 son más similares a las cepas Tipo 1 que las otras secuencias Tipo 2 de longitud total secuenciadas. En particular, 5 aminoácidos (encuadrados en Fig. 11) mimetizan directamente las cepas Tipo 1. Sin embargo, los aminoácidos que se conservaban en las otras secuencias Tipo 2 no redundantes estaban también mayoritariamente conservados en los aislados MN184, pero aminoácidos dispersos y la similitud de aminoácidos (84,8-85,4%) revelaban una proteína Tipo 2 más divergente que se había evidenciado hasta la fecha. Así, la alineación define adicionalmente residuos mantenidos de nsp1β que pueden ser críticos para el ciclo de replicación de PRRSV.

Una alineación de aminoácidos de secuencias no redundantes de nsp2, ordenada por identidad de parejas, se muestra en la Figura 12. Un dominio de cisteína-proteasa afín a quimotripsina altamente conservado (PL2) está presente en el término N, predicho previamente por alineación con el virus de la arteritis equina (EAV) nsp2 (Snijder et al., *J. Gen. Virol.*, 79:961-979 (1998); Ziebuhr et al., *J. Gen. Virol.*, 81:853-879 (2000)). Existen 3-4 dominios transmembranales predichos cerca del término C de esta proteína (McGuffin et al., *Bioinformatics*, 16:404-405 (2000)), pero el sitio de escisión exacto en el terminal C no ha sido determinado empíricamente. Se han propuesto dos predicciones del sitio de escisión C-terminal (una G|G en el aminoácido 980 de nsp2 VR-2332 (Allende et al., *J. Gen. Virol.*, 80:307-315 (1999)) y la otra en el aminoácido 1197 (Ziebuhr et al., *J. Gen. Virol.*, 81:853-879 (2000)), pero existen varios dobles G|G completamente conservados en esta proteína (aminoácidos 646, 980, 1116, 1196, y 1197 de nsp2 de VR-2332; flechas descendentes en Fig. 12). Un trabajo anterior había indicado también que la proteína nsp2 predicha es rica en prolina y contiene epítopos múltiples de células B potenciales (Oleksiewicz et al., *J. Virol.*, 75:3277-3290 (2001); Fang et al., *Virus Res.*, 100:229-235 (2004); Ropp et al., *J. Virol.*, 78:3684-3703 (2004)). La región intermedia grande de PRRSV nsp2 (aminoácidos 148-880 de nsp2 de VR-2332) no tiene función asignada alguna, pero es de longitud muy variable. Adicionalmente, la diferencia de longitud entre las cepas Tipo 1 y Tipo 2 de PRRSV secuenciadas ha sido mapeada a esta región variable intermedia de nsp2 (Fig. 12). Hasta ahora, se ha demostrado que los genomas Tipo 1 secuenciados son 313-364 bases más cortos que la mayoría de los PRRSV Tipo 2 (Meulenberg et al., *Virology*, 192:62-72 (1993); Fang et al., *Virus Res.*, 100:229-235 (2004); Ropp et al., *J. Virol.*, 78:3684-3703 (2004)). Sin embargo, la alineación de secuencias múltiples estableció que el genoma de MN184 contiene la nsp2 más corta predicha hasta la fecha (2547 pb), 393 pb más corta que la cepa de VR-2332 Tipo 2 prototipo. Adicionalmente, la misma contenía 3 delecciones discontinuas en la proteína traducida con tamaños de delección consistentes en 111, 1 y 19 aminoácidos, respectivamente, correspondientes a las posiciones de los aminoácidos en la cepa VR-2332 de PRRSV nsp2 de 324-434, 486 y 505-523, respectivamente (Fig. 12). Las tres

deleciones daban como resultado la pérdida de varios residuos prolina y predecían epítopes de células B. Además de estas deleciones, se observaban también alteraciones significativas en la secuencia de aminoácidos de nsp2 de otras cepas Tipo 2, correspondientes en algunos casos al aminoácido Tipo 1 observado en la misma posición relativa (Fig. 12). La comparación de la proteína predicha nsp2 de los dos genotipos PRRSV demostró que la identidad de aminoácidos dentro de los virus Tipo 2 variaba desde 66% a 99% y desde 88 a 90% en los virus Tipo 1, pero difería notablemente entre genotipos (semejanza < 45). En particular, los aislados de MN184 exhibían 66-80% de identidad de aminoácidos con todas las proteínas nsp2 Tipo 2 predichas y únicamente 43-45% de identidad con las cepas Tipo 1. Cuando se examinó la alineación de secuencias múltiples en Fig. 12, se observó también que todos los casos de inserción o delección en ambos genotipos ocurrían en esta región hipervariable intermedia. Sobre este punto, Shen et al. (*Arch. Virol.*, 145:871-883 (2000)) informaron por primera vez que la cepa SP Tipo 2 de PRRSV de Norteamérica tiene una inserción única de 36 aminoácidos con relación a la posición entre los aminoácidos 813 y 814 de PRRSV VR-2332 nsp2. Otro investigador encontró una delección única de 12 aminoácidos en la posición 466-477 en el aislado de PRRSV HB2(sh)/20002 nsp2 (Gao et al., *Arch. Virol.*, 149:1341-1351 (2004)). Existía una delección de 17 aminoácidos en aislados de PRRSV afines a los europeos de nueva identificación cuando se comparaban con la cepa LV (Fang et al., *Virus Res.*, 100; 229-235 (2004); Ropp et al., *J. Virol.*, 78:3684-3703 (2004)). Los casos de mutación no ocurrían consistentemente a lo largo del mismo tramo de aminoácidos, aunque las deleciones observadas entre los aislados MN184 y otros virus Tipo 2 abarcan la mayor parte de la delección más larga detectada entre el Tipo 1 y otros PRRSV Tipo 2. Todos estos datos sugerían que el ORF de nsp2 contiene un resto proteasa conservado y regiones predichas de alcance transmembranal que pueden ser necesarias para la replicación de PRRSV, pero esto es sumamente susceptible de mutación en la gran sección intermedia.

La aparición repentina de aislados de campo de PRRSV en Minnesota que reflejan el patrón 184 RFLP es todavía un misterio, pero las consecuencias de este evento están comprobándose todavía actualmente. El Laboratorio de Diagnóstico Veterinario de Minnesota realiza actualmente la secuenciación rutinaria en aislados similares de 184 RFLP de aproximadamente una cuarta parte del número total de solicitudes de secuencia de ORF5. Adicionalmente, el patrón de 184 RFLP ha sido detectado ahora no sólo en Minnesota, sino también en Iowa, Wisconsin, Dakota del Sur, Kansas, Missouri, Illinois, Nebraska, Kentucky, Oklahoma y Wyoming. Los autores de la presente invención eligieron derivar las secuencias de longitud total de dos aislados debido a la necesidad de entender si esto podría ser más que un solo tipo de virus y al hecho de que la piara de cerdos diagnosticada con el aislado MN184A se presentaba con un caso más leve de PRRS que la piara infectada con el aislado ML184B, tal como fue comunicado por el patólogo encargado del caso. Las cepas no han sido inoculadas en animales naïf para comprobar las presentaciones de los casos, pero es interesante observar que el aislado ML184B tenía muchas más degeneraciones de nucleótidos detectadas cuando se analizó el genoma, y esto podría reflejar la gravedad de la enfermedad informada.

Este análisis del genoma aumentó la comprensión por los autores de la invención de la inmensa variación de secuencias de nucleótidos y aminoácidos que existe en el campo. Factores que impulsan esta variación pueden estar relacionados con el modo en que se están tratando actualmente los cerdos, el transporte interestatal e internacional de los cerdos y el semen de los cerdos macho, la intermezcladura de aislados de PRRSV diferentes en las piaras y la naturaleza del propio virus. La generación de secuencias del genoma total permite también a los inventores monitorizar dónde se tolera la variación del genoma y qué regiones están más conservadas. Como resultado de este estudio, así como de una publicación previa (Ropp et al., *J. Virol.*, 78:3684-3703 (2004)), está emergiendo un cuadro que indica que nsp2, nsp1β y ORF5 son proteínas extraordinariamente versátiles.

Este estudio ha proporcionado también una evidencia clara de que el tamaño de nsp2 no puede utilizarse ya para diferenciar entre los dos genotipos de PRRSV. El nuevo descubrimiento de que nsp2 evolucionó para exhibir un genoma Tipo 2 con tres deleciones discontinuas, que conducen al genoma más corto hasta la fecha (15019 kb), sugiere que PRRSV puede estar evolucionando para eliminar regiones genómicas prescindibles y hacer más compacto el genoma. Finalmente, aunque la significación de variaciones genéticas en PRRSV puede sólo conjeturarse actualmente, el cambio evolutivo observado en ORF5, nsp1β y nsp2 debería estar relacionado razonablemente con la adaptación biológica de PRRSV durante la presión de selección.

La descripción y los ejemplos detalladas que anteceden se han dado únicamente para claridad de comprensión. No deben deducirse de ellos limitaciones innecesarias en ningún caso.

A no ser que se indique otra cosa, todos los números que expresan cantidades de componentes, pesos moleculares, etcétera, utilizados en la memoria descriptiva y las reivindicaciones deben entenderse como modificados en todos los casos por el término "aproximadamente". De acuerdo con ello, a no ser que otra cosa indique lo contrario, los parámetros numéricos expuestos en la memoria descriptiva y las reivindicaciones son aproximaciones que pueden variar dependiendo de las propiedades deseadas que pretendan obtenerse por la presente invención. Como mínimo, y no como un intento de limitar la doctrina de equivalencias dentro del alcance de las reivindicaciones, cada parámetro numérico debería interpretarse teniendo presente al menos el número de dígitos significativos consignados y aplicando técnicas ordinarias de redondeo.

A pesar de que los intervalos y parámetros numéricos que exponen el alcance general de la invención son aproximaciones, los valores numéricos expuestos en los ejemplos específicos se han expresado con la mayor

precisión posible. Todos los valores numéricos, sin embargo, contienen inherentemente un intervalo resultante necesariamente de la desviación estándar encontrada en sus medidas de test respectivas.

Todos los encabezamientos se incluyen para conveniencia del lector y no deberían utilizarse para limitar el significado del texto que sigue al encabezamiento, a no ser que se especifique en tal sentido.

- 5 La descripción también comprende los siguientes puntos:
 1. Un polinucleótido infeccioso aislado que comprende una secuencia de nucleótidos que tiene al menos un 88% de identidad con SEQ ID NO: 1 y una delección de al menos 39 nucleótidos consecutivos que corresponden del nucleótido 2062 al nucleótido 3864 de SEQ ID NO: 1.
 2. Un polinucleótido aislado que comprende una secuencia de nucleótidos que tiene al menos un 88% de identidad con SEQ ID NO: 1 y al menos una delección de al menos 39 nucleótidos consecutivos que corresponden del nucleótido 2062 al nucleótido 3864 de SEQ ID NO: 1 y en el que el polinucleótido se replica y produce partículas virales infecciosas cuando se introduce en una célula.
 3. Un clon infeccioso que comprende un polinucleótido que tiene una secuencia de nucleótidos que tiene al menos un 88% de identidad con SEQ ID NO: 1 y al menos una delección de al menos 39 nucleótidos consecutivos que corresponden del nucleótido 2062 al nucleótido 3864 de SEQ ID NO: 1.
 4. Un polinucleótido infeccioso aislado que comprende una secuencia de nucleótidos que tiene al menos un 88% de identidad con SEQ ID NO: 14 y una delección de al menos 39 nucleótidos consecutivos que corresponden del nucleótido 2061 al nucleótido 3545 de SEQ ID NO: 14.
 5. Un polinucleótido aislado que comprende una secuencia de nucleótidos que tiene al menos un 88% de identidad con SEQ ID NO: 14 y al menos una delección de al menos 39 nucleótidos consecutivos que corresponden del nucleótido 2061 al nucleótido 3545 de SEQ ID NO: 14 y en el que el polinucleótido se replica y produce partículas virales infecciosas cuando se introduce en una célula.
 6. Un clon infeccioso que comprende un polinucleótido que tiene una secuencia de nucleótidos que tiene al menos un 88% de identidad con SEQ ID NO: 14 y al menos una delección de al menos 39 nucleótidos consecutivos que corresponden del nucleótido 2061 al nucleótido 3545 de SEQ ID NO: 14.
 7. El polinucleótido aislado de los puntos 1, 2, 4 y 5, en el que el polinucleótido está presente en un vector.
 8. El polinucleótido aislado de los puntos 1, 2, 4, o 5 o el clon infeccioso de los puntos 3 ó 6, en el que el polinucleótido comprende 2 o más delecciones, y en el que cada delección es independientemente al menos 37 nucleótidos consecutivos.
 9. El polinucleótido aislado de los puntos 1, 2, 4 ó 5, en el que el polinucleótido está presente en una partícula de virus aislada.
 10. El polinucleótido aislado de los puntos 1, 2, 4, o 5, en el que el polinucleótido es un polinucleótido de ARN.
 11. El polinucleótido aislado de los puntos 1, 2, 4 ó 5 o el clon infeccioso de los puntos 3 ó 6, en el que el polinucleótido está presente en una célula.
 12. El polinucleótido aislado de los puntos 1, 2, 4, o 5 o el clon infeccioso de los puntos 3 ó 6, en el que un promotor de ARN polimerasa está unido operativamente al polinucleótido.
 13. El polinucleótido aislado de los puntos 1, 2, 4, o 5 o el clon infeccioso de los puntos 3 ó 6, en el que el polinucleótido comprende además un polinucleótido exógeno presente en la delección.
 14. El polinucleótido aislado o el clon infeccioso del punto 13, en el que el polinucleótido exógeno codifica un marcador detectable.
 15. Un polinucleótido infeccioso aislado que comprende una secuencia de nucleótidos SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12 o SEQ ID NO: 13.
 16. Un polipéptido nsp2 codificado por un polinucleótido infeccioso que comprende una secuencia de nucleótidos SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, O SEQ ID NO: 13.

REIVINDICACIONES

1. Un clon infeccioso que comprende un polinucleótido infeccioso que tiene la secuencia de nucleótidos de SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12 O SEQ ID NO: 13 y una secuencia vectorial que se replica en una célula huésped procariótica, en el que dicho polinucleótido tiene una delección en la secuencia que codifica nsp2.
2. Un polinucleótido aislado que comprende un genoma del virus PRRS que comprende una delección en la secuencia que codifica nsp2, en el que dicho polinucleótido tiene la secuencia de nucleótidos de SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10 , SEQ ID NO: 11, SEQ ID NO: 12 o SEQ ID NO: 13 para su uso en la inmunización de un animal contra el virus PRRS.
3. Uso de un polinucleótido aislado que comprende un genoma del virus PRRS que comprende una delección en la secuencia que codifica nsp2, en el que dicho polinucleótido tiene la secuencia de nucleótidos de SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO : 10, SEQ ID NO: 11, SEQ ID NO: 12 o SEQ ID NO: 13 para la preparación de una composición para inmunizar un animal contra el virus PRRS.
4. El clon infeccioso según la reivindicación 1 o el polinucleótido para uso según la reivindicación 2 o el uso según la reivindicación 3, en el que el polinucleótido se replica y produce partículas virales infecciosas cuando se introduce en una célula permisiva.
5. El clon infeccioso según la reivindicación 1 o el polinucleótido para uso según la reivindicación 2 o el uso según la reivindicación 3, en el que un promotor de ARN polimerasa está unido operativamente al polinucleótido.
6. El clon infeccioso según la reivindicación 1 o el polinucleótido para uso según la reivindicación 2 o el uso según la reivindicación 3, en el que el polinucleótido comprende además un polinucleótido exógeno presente en la delección.
7. El clon infeccioso según la reivindicación 6 o el polinucleótido para uso según la reivindicación 6 o el uso según la reivindicación 6, en el que el polinucleótido exógeno codifica un marcador detectable.

Fig. 1A-3

ACTAAGAGAGGTGGCCTGTGCTGCGGACCCGATCACCTCTGTGCTAACACCATTTATAGTTTGGTGATCTATGCACA
 GCATATGGTGCTTAGTTACTTCAAAAGTGGTCACCCCATGGCCTTCTGTTCTTACAAGACCAGCTAAAGTTTGGAGACA
 TGCTCAAGGTTCAACCCCTGATCGTCTATTCCGACGACCTCGTGCTGTATGCCGAGTCTCCACCATGCCAACTATCAC
 TGGTGGGTTGAACATCTGAATTTGATGCTGGGGTTTCAGACGGACCCAAAGAAGACAGCAATAACAGACTCGCCATCATT
 TCTAGGCTGTAGAATAATAATGGGCGCCAGCTAGTCCCCAACCGTGACAGGATCCTCGCGGCCCTCGCCTATCACATGA
 AGGCGAGTAATGTTTCTGAATACTATGCCCTCAGCGGCTGCAATACTCATGGACAGCTGTGCTTGTGGAGTATGATCCT
 GAATGTTTGAAGAACTGTAGTTGGAATAGCGCAGTGCGCCCGCAAGGACGGCTACAGCTTTCCCGGCACGCCGTTCTT
 CATGTCCATGTGGGAAAACTCAGGTCCAATTATGAGGGGAAGAAGTCGAGAGTGTGCGGGTACTGCGGGGCCCGGCC
 CGTACGCTACTCGCTGTGGCCTCGACGTCTGCATTACCAACACCCACTTCCACCAGCATTGTCCAGTCACAATCTGGTGT
 GGCCATCCAGCGGGTTCTGGTCTTGTAGTGAGTGCAAAATCCCCTGTAGGGAAAGGCACAAGCCCTTTAGACGAGGTGCT
 GGAACAGTCCCGTATAAGCCCCACGGACCGTTATCATGCATGTGGAGCAGGGTCTCACCCCTTGTATCCAGGTAGAT
 ACCAACTCGCCGCGGATAGTCTCTGTGAGGCGTGGAATTAGGGGAAATGAAGTTGGACTACCAGACGGTGATTATGCT
 AGCACGCTTCTGCTCCCTACCTGCAAAAGAGATCAACATGCTCGCTGTGCTTCCAATGATTGCGCAGGTTTATCATCT
 CGGCCACCCCGGTGCTGGGAAAACATACTGGCTCCTTCAACAGGTCCAGGATGGTGATGTTATTTACACACCACTCACC
 AGACCATGCTTGACATGATTAGGCTTTGGGGACGTGCCGTTCAACGTCCCGGCAGGCACAACGCTGCAATTTCCCGTC
 CCTCCCGCACCCGTCGTTGGGTTTCGCATCTAGCCGGCGGTGGTGTCTGGCAAGAATTCCTTCTAGATGAAGCAGC
 GTATTCGACATCTCCGCTTCCCAAGGTTCGACACAACTTGGGATTTTATTTCTCACCTGATTAAACACAGTTTGCATAAACT
 CAGTGGGTTTGTATCTCATTTGCTATGTTTGTGACATCATGCCCTCAAACTCAACTGAAGACCATCTGGAGGTTTGGACAG
 AATATCTGTGATGCCATTACAGCAGATTACAGGGACAACTCATGTCCATGGTCAACACAACCCGTGTGACCTACGTGGA
 AAAACCTGTGAGGTATGGGCGGTCCTCACCCCTTACCAGGGACCGAGAGGACGACGCCATCACTATTGACTCCAGTC
 AAGGCCCCACATTCGATGTGGTTACATTGCATTTGCCCACTAAAGATTCACTCAACAGGCAAGAGCCCTTGTGCTATC
 ACCAGGGCAAGACACGCTATCTTGTGTATGACCCACACAGGCAGTGCAGGGCTTGTGTTGATCTTCTGCAAAAGGCAC
 GCCCGTCAACCTCGCAGTGCACTGCGACGGGCAGCTGATCGTGCTGGATAGAAATAACAAAGATGCACGGTTGCTCAGG
 CTCTAGGCAACGGGATAAATTTAGGGCCACAGACAAGCGTGTGTAGATTCTCTCCGCGCCATTTGTGCTGATCTAGAA
 GGGTCGAGCTCTCCGCTCCCAAGGTTCGACACAACTTGGGATTTTATTTCTCACCTGATTAAACACAGTTTGCATAAACT
 CCCAGTAGAATTTGCACCTCACTGGCCCGTGGTGTCAACCCAGAACAATGAAAAGTGGCCGGATCGGCTGGTTGCCAGCC
 TTCGCCCTATCCATAAATACAGCCGCGCTGCATCGGTGCCGCTATATGTTGGGCCCTTCGCTGTTTCTAGGCACTCCT
 GGGGTGCTGTACATATCTCAAAAATTTGTTAAGGGCGGGCTCAAGTGCTTCCGAGACGGTTTTTCAGCACCGGCCG
 AATTGAGGTAGACTGCCGGGAATATCTTGATGATCGGGAGCGAGAAGTTGCTGCGTCCCTCCACACGCTTTCATTGGCG
 ACGTCAAAGGCACTACCGTTGGAGGATGTATCATGTCACTCCAGATACCTCCCGCGCTCCTTCCCAAGGAATCAGTT
 GCGGTAGTCGGGTTTTCAAGCCCCGGAAGCCGCGAAAGCATTGTGCACACTGACAGATGTGTACCTCCAGATCTTGA
 AGCCTATCTCCACCCGAGACCCAGTCCAAGTGCTGGAATATGATGTTGGACTTCAAAGAAGTTCGACTAATGGTCTGGA
 AAGACAAAACAGCCTATTTCCAACTTGAAGGTGCTATTTACCTGGTATCAGCTTGCCAGCTATGCCCTCGTACATCCGT
 GTTCCCGTCAACTCTACGGTGTACTTGGACCCCTGCATGGGCCCGCCCTTTGCAACAGGAGAGTCGTGCGGTCCACCCA
 CTGGGGGGCTGACCTCGCGGTCAACCCCTATGATTACGGCGCTAAAATATCTGTCTAGCGCGTACCATGGTGAAATGC
 CCCCAGGATACAAAATCTGGCGTGCGCGGAGTTCTCGTTGGATGACCCAGTTAAGTACAAACATACCTGGGGGTTTGA
 TCGGATACAGCGTATCTGTATGAGTTACCCGAAACGGTGGGAGGATTACAATGATGCGTTTCTGTCGCGCCG
 GGAAGGGAATTTATAAGGCCACTGCCACCAGCTTGAAGTTTATTTTCCCGGGGCCCTGTCAATTGAACCAACTTTAG
 GCCTGAATTGAATGAATGGGTCCATGCAAGCCTTTTGTACAAAATGGCCAACTTTTGTGGATGCTTTACGGAG
 TTCTTGGTGTCCATTGTTGATATCATTATATTTTGGCCATTTTGTGTTGGCTTACCATCGCCGGTTGGCTGGTGGTCTT
 TTGCATCAGATTGGTTTGGCTCCGCGATACCTCGTACGCGCCCTGCCATTCACTCTGAGCAATTACAGAAGATCTTATGAG
 GCCTTCTTTCCAGTGCCAAAGTGGACATTCCACCTGGGGAACATAACATCCTTTGGGGATGCTTTGGCACCATAAGGT
 GTCACCCCTGATTGATGAAATGGTGTGCGCTCGAATGTACCGCATCATGGAAAAGCAGGGCAGGCTGCTGGAAACAGG
 TGGTGAGCGAGGCTACGCTGTCTCGCATAGTAGTTTGGATGTGGTGGCTCATTTTTCAGCATCTAGCCGCATTGAAGCC
 GAGACCTGTAATATTTGGCCTCCCGGCTGCCCATGCTACCAACCTGCGCATGACAGGGTCAAATGTAACCATAGTGTA
 TAATAGCACTTTGAATCAGGTGTTTGTATTTTCAACCCCTGGTTCCCGGCCAAAGCTTATGATTTTTCAGCAATGGT
 TAATAGCTGTACATCTCCTCATATTTTCTGTGACGCTTCTGTACTCTTTTGTGTTGCTGTGGTTGCGGGTTCCA
 ATACTACGTAAGTTTGGTTTCCGCTGGTTAGGGCAATTTTCTTTGAACTCACAGTGAATTACACGGTGTGTCCA
 CCTTGCTCACCCGGCAAGCAGCCACAGAGATCTACGAACCCGGTAGGTCTCTTGGTGAGGATAGGGTATGACCGATG
 TGGGGAGGACGATCATGACGAGCTAGGTTTATGATACCGCTGGCCTCTCCAGCGAAGGCCACTTGACTGGTGTTTACG
 CCGTGGTTGGCGTTCTGTCTCAGCTACACGGCCAGTTCCATCCCGAGATATTCGGGATAGGGAATGTGAGTCGAGTT
 TATGTTGACATCAACATCACTCATCTGCGCGAACATGACGGGCAGAACACCACCTTGCCCTCGTCATGACAACTTTC
 AGCCGTGTTTCAGACCTATTACCAACATCAAGTCGACGGCGCAATTTGGTTTCACTAGAATGGCTTCGCTCCCTTCTTT
 CCTCGTGGTTGGTTTTAAATGTCTCTTGGTTTCTCAGGCGTTCGCTGCAAAACCATGTTTCAGTTCGAGTCTTGACAGATA
 TTAAGACCAACACCACCGCAGCGGCAAGCTTTGCTGTCTCCAGACATCAGTTGCCTTAGGCATCGCGACTCGGCCCTCT
 GAGGCGATTTCGCAAAATCCCTCAGTCCGTACGGCGATAGGGACACCCGTGTATGTTACCATCACAGCCAATGTGACAGA
 TGAGAAATTATTACATCTCTGATCTCCTCATGCTTCTTCTTGCCTTTCTATGCTTCTGAGATGAGTGAAAAGGGAT
 TTAAGGTGGTATTGGCAATGTGTAGGCATCGTGGCTGTGTGTGCAATTTTACCAGCTACGTCCAACATGTCAAGGAG

Fig. 1A-4

TTTACCCAACGCTCCCTGGTGGTCGACCATGTGCGGTTGCTCCATTTTCATGACACCTGAGACCATGAGGTGGGCAACTGT
 TTTAGCCTGTCTTTTGGCAATTCGTGGCAATTTGAATGTTTAAGTATGTTGGAGAAATGCTTGACCGCGGGCTGTTGC
 TCGCGATTGCTTTCTTTGTGGTGTATCGTGCCGTTCTGTTTTGCTGTGCTCGCCAAACGCCAGCAACGACAGCAGCTCCCA
 TCTACAGCTGATTTACAACCTGACGCTATGTGAGCTGAATGGCACAGATTGGCTAGCTAACAAATTTGATTGGGCAGTGG
 AGAGTTTTGTTCATCTTTCCCGTTTTGACTCACATTGTCTCCTATGGTGCCCTCACTACCAGCCATTTCTTGACACAGTC
 GCTTTAGTCACTGTGTCTACCGCCGGGTTTGTTCACGGGCGGTATGTCTTAAGTAGCATCTACGCGGTCTGTGCCCTGGC
 TGGCTTGACTTGCTTCGTCAATAGGTTTGCAAAGAATTGCATGTCTGGCGCTACGCGTGTACAGATATACCAACTTTC
 TTCTGGACACTAAGGGCAGACTCTATCGTTGGCGGTGCGCTGTTCATCATAGAGAAAAGGGGCAAAGTTGAGGTGGAAGGT
 CATCTGATCGACCTCAAAGAGTTGTGCTTGATGGCTCCGTGGCAACCCCTATAACCAGAGTTTCAGCGGAACAATGGGG
 TCGTCCTTAGATGACTTCTGTACGATAGCACGGCTCCACAAAAGGTGCTTTTGGCGTTTTCTATTACCTACACGCCAGT
 GATGATATATGCCCTAAAGGTGAGTCGCGGCCGACTGCTAGGGCTTCTGCACCTTTTGATCTTCTGAATTGTGCTTTCA
 CCTTCGGGTACATGACTTTTCGCGCACTTTTCAGAGTACAAATAAGGTGCGCTCACTATGGGAGCAGTAGTTGCACTCCCTT
 TGGGGGGTGTACTCAGCCATAGAAACCTGGAAATTCATCACCTCCAGATGCCGTTTGTGCTTGCTAGGCCGCAAGTACAT
 TCTGGCCCCCTGCCCACCACGTTGAAAGTGCCGCAGGCTTTTCATCCGATTGCGGCAAAATGATAACCACGCATTTGTGCTCC
 GCGTCCCGGCTCCACTACGGTCAACGGCACATTGGTGCCCGGTTAAAAAGCCTCGTGTGGGTGGCAGAAAAGCTGTT
 AAACAGGGAGTGGTAAACCTTGTCAAATATGCCAAATAACAACGGCAAGCAGCAGAGAAGAGAAAGAGGGGATGGCCAGC
 CAGTCAATCAGCTGTGCCAGATGCTGGGTAAAGTATCGCTCAGCAAAACCACTCCAGAGGCAAGGGACCGGGAAGAAA
 AATAAGAAGAAAACCGGAGAAGCCCATTTTCCTCTAGCGACTGAAGATGATGTGACACATCACTTTACCCCTAGTGA
 GCGGCAATTGTGTCTGTGCTCAATCCAGACCGCCTTTAATCAAGGCGCTGGGACTTGCACCTGTGAGATTGAGGAGGA
 TAAGTTACACTGTGGAGTTTAGTTTTGCCTACGCATCATACTGTGCGCCTGATCCGCGTCACAGCATCACCTCAGCATGA
 TGGGCTGGCATTCTTGAGGCATCTCAGTGTGTAATTGGAAGAATGTGTGTTGAATGGCACTGATTGACATTGTGCTCT
 AAGTCACCTATTCATTAGGGCGACCGTGTGGGGTGAGATTTAATTGGCGAGAACCATGCGGCCGAAATTAACAAAAA

Fig. 1B-1

>VR-V5.seq

ATGACGTATAGGTGTTGGCTCTATGCCTTGGCATTGTATTGTGTCAGGAGCTGTGACCATTGGCACAGCCCAAACTTGCT
 GCACAGAAACACCCCTTCTGTGATAGCCTCCTTCAGGGGAGCTTAGGGTTTGTCCCTAGCACCTTGCTCCGGAGTTGCAC
 TGCTTTACGGTCTCTCCACCCCTTTAACCATGTCTGGGATACTTGATCGGTGCACGTGTACCCCAATGCCAGGGTGT
 ATGGCGGAGGGCCAAGTCTACTGCACACGATGCCTCAGTGCACGGTCTCTCCTTCCCTGAACCTCCAGGTTTCTGAGCT
 CGGGGTGCTAGGCCTATTCTACAGGCCCGAAGAGCCAACCTCCGGTGGACGTTGCCACGTGCATTCCCCACTGTTGAGTGCT
 CCCCCGCGGGGCTGCTGGCTTTCTGCAATCTTCCAATCGCACGAATGACCAGTGGAAACCTGAACCTTCAACAAAGA
 ATGGTACGGGTGCGAGCTGAGCTTTACAGAGCCGGCCAGCTCACCCCTGCAGTCTGAAGGCTCTACAAGTTTATGAACG
 GGGTTGCCGCTGGTACCCATTGTTGGACCTGTCCCTGGAGTGGCCGTTTTCGCCAATTCCTACATGTGAGTGATAAAC
 CCTTCCCGGGAGCAACTCAGTGTTGACCAACCTGCCGCTCCCGCAGAGACCAAGCCTGAAGACTTTTGCCCTTTGAG
 TGTGCTATGGTACTGTCTATGACATTGGTCATGACGCCGTGATGTATGTGGCCGAAAGGAAAGTCTCCTGGGCCCTCG
 TGGCGGGGATGAAGTGAATTTGAAGCTGTCCCGGGGAGTTGAAGTTGATTGCGAACCGGCTCCGCACCTCCTTCCCGC
 CCCACACACAGTGGACATGTCTAAGTTCGCCTTACAGCCCTGGGTGTGGTGTCTATGCGGGTGAACGCCAACAC
 GGCTGCCTTCCCGCTGACACTGTCCCTGAAGGCACTGCTGGTGGAGCTTGTGACTTGCTTCCACTGGAAGTTCAGAA
 CAAAGAAATTCGCCATGCTAACCAATTTGGCTACCAGACCAAGCATGGTGTCTTGGCAAGTACCTGCAGCGGAGGCTGC
 AAGTTAATGGTCTCCGAGCAGTAACTGACCTAAACGGACCTATCGTCGTACAGTACTTCTCGTTAAGGAGAGTTGGATC
 CGCCATTTGAAACTGGCGGGAGAACCCAGCTACTCTGGGTTTGAGGACCTCCTCAGAATAAGGGTTGAGCCTAACACGTC
 GCCATTGGCTGACAAGGAAGAAAAATTTCCGGTTTGGCAGTCACAAGTGGTACGGCGCTGGAAGAGAGCAAGAAAAAG
 CACGCTCTTGTGCGACTGCTACAGTCGCTGGCCGCGCTTTGTCCGTTCTGTGAAACCCGCGCAGGCCAAGGAGCACGAGGTT
 GCCCGCGCAACAAAGGCTGAGCACCTCAAACACTACTCCCCGCTGCGGAAGGGAATTGTGGTTGGCACTGCATTTCCGC
 CATCGCAACCCGGATGGTGAATTCCAAATTTGAAACCAACCTTCCCGAAAGAGTGAGACCTCCAGATGACTGGGCTACTG
 ACGAGGATCTTGTGAATGCCATCCAAATCCTCAGACTCCCTGCGGCCCTTAGACAGGAACGGTGCTTGTACTAGCGCAAG
 TACGTACTTAAGCTGGAAGGTGAGCATTGGACTGTCACTGTGACCCCTGGGATGTCCCTTCTTGTCTCCCTCTTGAATG
 TGTTACGGGCTGTTGTGGGCACAAGGGCGGTCTTGGTTCCCGAGATGCACTGAGGCTCTCCGATTTGACCTGCCTGCC
 TTGACCGGCTGGCTGAGGTGATGCACCTGCCTAGCAGTGCATCCCAGCCGCTGTGGCCGAAATGTCTGGCGATTCCGAT
 CGTTCCGCTTCTCCGGTCAACCCGTGTGGACTGTTTCGAGTTCCTTGGCCGTCACAGCGGAGGGAATCACCTGACCA
 AGTGGCCTTAGGGAAAATTATCAGCCTTTGTGAGGTGATTGAGGACTGCTGCTGTTCCAGAACAAAACCAACCGGGTCA
 CCCCCGAGGAGGTGCGACGAAAGATTGACCTGTACCTCCGTGGTGCAACAAATCTTGAAGAATGCTTGGCCAGGCTTGA
 AAAGCGCGCCGCCACGCGTAATCGACACCTCCTTGTATTGGGATGTTGTGCTCCCTGGGGTTGAGGCGGCAACCCAGAC
 GATCAAGCTGCCCCAGGTCAACCAAGTGTCTGCTGCTGGTCCCTGTTGTGACTCAAAAGTCTTGGACAACAACCTCGGTCC
 CCTGACCGCTTTTCACTGGCTAACTACTACTACCGTGCAGGAGTGAAGGTTCTGACCGTGAAAGACTAACCGCC
 GTGCTCTCAAGTTGGAAGAGGTTGTTGAGAAGAATATGGGCTCATGCCAACCGAGCCTGGTCCACGGCCCACTGCC
 ACGCGGGCTCGACGAACTCAAAGACCAGATGGAGGAGGACTTGCTGAAACTGGCTAACGCCAGACGACTTCGGACATGA
 TGGCCTGGGCGAGTCGAGCAGGTTGACCTAAAAAATTGGGTCAAGAACTACCCGCGGTGGACACCAACCCCTCCGCCA
 AAAGTTCAGCCTCGAAAAACGAAGCCTGTCAAGAGCTTGCCGGAGAGAAAGCCTGTCCCCGCCCGCGCAGGAAGGTTGG
 GTCGATTGTGGCAGCCCGGTTTCATTAGCGGCGATGTCCCTAACAGTTGGGAAGATTTGGCTGTTAGTAGCCCTTTG
 ATCTCCCGACCCACCTGAGCCGGCAACACCTTCAAGTGAGCTGGTGATTGTGCTCCTACCGCAATGCATCTTACGGCCG
 GCGACACCTTGAGTGAGCCGGCTCAATTCGCGACCTCGCGAACTGTGTCTCGACCGGTGACACCTTGAGTGAGCC
 GATCCTGTGCCCGACCGCGGCGTAAGTTTCAGCAGGTGAAAAGATTGAGTTCGGCGGCGGCAATCCCACCGTACCAGG
 ACAGCCCTTGATTTGTCTGCTTCTCACAGACTGAATATGAGGCTCTCCCCAGCACCGCGCAGAGCGGGGGCGTT
 CTGGGAGTAGAGGGGATGAAGCTGAGGAAACCTGAGTGAAATCTCGGACATGTGGGTAAACATTAAACCTGCGTCCGT
 GTCATCAAGCAGCTCCTTGTCCAGCGTGAGAATCACACGCCAAAAATACTCAGCTCAAGCCATCATCGACTCGGGCGGGC
 CCTGAGTGGGCATCTCAAGAGGTAAAGGAAACATGCCTTAGTGTGATGCGCGAGGCATGTGATGCGACTAAGCTTGAT
 GACCTGCTACGCGAAGATGGCTTTCTCGCATGTGGGATCGGGTGGACATGCTGACTTGGCGCAACAGTCTGTTACCA
 GCGGATTTGCACCTTAGATGGCAGGTTAAAGTTCCTCCCAAAATGATACTCGAGACACCGCGCCCTATCCGTGTGAGT
 TTGTGATGATGCTCACACGCTGCACCTCCGTAGGTGCGGAGAGCGACCTTACCATTGGCTCAGTTGCTACTGAAGAT
 GTTCCACGCATCCTCGAGAAAATAGAAAATGTGGCGAGATGGCAACCGGAGCCCTTGGCCTTCTCGAGGATAAACCC
 GGATGATGACCAACTTGTCAACGACCCCCGATATCGTCGCGGAGGCTGACGAGAGCACATCAGCTCCGTCCGCGAGGCA

Fig. 1B-2

CAGGTGGCGCCGGCTCTTTTACCGATTGCGCCCTTCAGATGGCGCGGATGCGGACGGGGGGGGCCGTTTCGGACGGTA
 AAAAGAAAAGCTGAAAGGCTCTTTGACCAACTGAGCCGTAGGTTTTTACCTCGTCTCCCATCTCCCTGTTTTCTCTC
 ACGCCTTTTCTACCTGGCGGTGGTTATTCTCGGGTGATTGGGGTTTTGACGTTTTACTCTATTGTGCCTCTTTTAT
 GTTACAGTTACCCAGCCTTTGGTATTGCTCCCTCTTGGGTGTGTTTTCTGGGTCTTCTCGGCGCGTTCGAATGGGGGT
 TTTGGCTGCTGGTTGGCTTTTGTGTTGGTCTGTTCAAGCCTGTGTCCGACCCAGTCGGCGCTGCTTGTGAGTTGACTC
 GCCAGAGTGTAGAAACATCTTCATTCTTTGAGCTTCTCAAACCTTGGGACCCTGTCGAGCCTTGTGTGGGCCCG
 TCGGTCTCGGTCTGCCATTCTTGGCAGGTTACTGGGCGGGGCACGCTGCATCTGGCACTTTTGTCTAGGCTTGGCATT
 GTTGCAGACTGTATCTTGGCTGGAGCTTACGTGCTTCTCAAGGTAGGTGTAAAAAGTCTGGGGATCTGTATAAGAAC
 TGCTCCCAATGAGGTGCTTTTAACTGTTTCTTTTACACGTGCGACCCAGGTGTCACCTTATCGACCTGTGCGATCGGT
 TTTGTGCGCAAAAGGAATGGACCCATTTTTCTCGCCACTGGGTGGCGCGGGTGTGGGCCGGCCGAAGCCCCATTGAG
 CAACCTCTGAAAAACCATCGCGTTTGGCCAGTTGGATGAAAAGAAGATTACGGCTAGGACTGTGGTCGCCAGCCTTA
 TGACCCCAACCAAGCCGTAAAGTGCTTGGGGTATTGCAGGCGGGTGGGGCGATGGTGGCTAAGGCGGTCCAAAAGTGG
 TCAAGTTTTCCGCTGTTCCATTCCGAGCCCCCTTCTTCCCACTGGAGTGAAAGTTGACCCCTGATTGCAGGGTCTGGTT
 GACCTGACACTTTCACTGCAGCTCTCCGGTCTGGCTACTCCACCACAAACCTCGTCTTGGTGTGGGGACTTTGCCCA
 GCTGAATGGATTAAAAATCAGGCAATTTCAAGCCTTCAAGGGGAGGCCACATCTCATGGCTGCCCTGCATGTTGCCT
 GCTCGATGGCTCTGCACATGCTTCTGGGATTATGTGACTGCGGTGGGTCTTGGCGCACCGGCACCAACGACCCGTGG
 TGGCTAACCCTTTGGCGTCCCTGGCTACGGACCTGGCTCTCTGACGTCAGATTGTGCATTTCCCAACACGGCCT
 TACCCTGCCCTTGACAGCACTTGTGGCGGGATTGGTATTCAAGAAATTGCCTTGGTGGTTTTGATTTTTGTTTCCATCG
 GAGGCATGGCTCATAGGTTGAGCTGTAAGGCTGACATGCTGTGTCTTGTGCAATTGCCAGCTATGTTGGGTACCT
 CTTACCTGGTTGCTTGTGTGTTTCTTGTGGTGGCTGTTTTCTTGCACCCCTCACCATCCTATGGTGGTGT
 TTTCTTGATTTCTGTGAATATGCCTTCAAGGAATCTTGGCATGGTGTGTTGGTTTCTTGGCTTCTGGTCTGTATA
 CTAATGTTGCTGGCCTTGTACCCCTACGACATTCATTACACAGTGGCCCCCGGGTGTGGCGCTTGGTACC
 GCACCAGATGGGACCTACTTGGCGCTGTCCGCGCGCTGCGTTGACTGGCCGCACCATGCTGTTTACCCGTCCAGCT
 TGGGTCTCTTCTGAGGTGCTTTCAGAACTCGAAAGCCCTCACTGAACACCGTCAATGTGATCGGGTCTCCATGGGCT
 CTGGCGGGGTGTTTACCATCGACGGGAAAGTCAAGTGCGTAACTGCCGCACATGTCCTTACGGGCAATTGAGCTCGGGT
 TCCGGGGTGGCTTCAATCAAATGCTTGACTTTGACGTAAGGGGAGATTTGCTATAGCTGATTGCCCGAATTGGCAAGG
 GGCTGCCCCCAAGACCAATTCTGCACGGATGGATGGAAGTGGCGTGCCTATTGGCTAACATCCTTGGCGTCGAACCCG
 GCGTCATTGGAAAAGGATTCGCTTCTGCTTCAACGATGTGGCGATTCCGGTCCCCAGTGATCACCAGGCGGGTGA
 CTTGTGGCGGTTACACGGGATCGAATAAACAAGGGGGGGCATTGTTACGCGCCCTCAGGCCAGTTTGTAAATGTGGC
 ACCCATCAAGCTAAGCGAATTAAGTGAATTTCTTGGTGGGCTAAGGTCCCGTCCGGTGATGTGAAGGTGGCAGCCACA
 TAATTAAGACATAAGCGAGGTGCCTTCAGATCTTGTGCTTGTGCTGCCAAACCTGAAGTGAAGGAGGCTCTCC
 ACCGTCAAACCTCTTGTGTGTTTTTCTCTGTGGAGAATGATGGGACATGCCTGGACGCCCTTGGTGTGTGAGTTT
 CTTATTTTGAATGAGGTTCTCCAGCGCTCTGGTCCGGAGTGTCTCTTCTTGAATGTTGTGCTATCCTGGCTCA
 CGCCATGGTCTGCGCAAGTTCTGATGATCAGGCTTCTGACAGCAGCTTAAACAGGAACAGATGGTCACTTGCCTTTTCT
 AGCCTCGGTGAGTGACCGTTTTGTGCGAGATCTTGGGCCACTCAGGGGCATCCGTTGCAGGCAGTGATGAATTTGAG
 CACCTATGCATTCCTGCCTCGGATGATGGTTGTGACCTCACCAGTCCAGTGATCACGTGTGGTGTCTGACCTACTTG
 CCATCATTTTGTACTTGTAAAGTACCGTGGCCTGACCATATCCTTGTGGCGATGGAGTGTCTGCGGCTTTCTC
 TTGAGATACTTGGCGAGGAAAGTTGAGGGAAGGGGTGTGCAATCCTGCGGAATGAATCATGAGTCTCTGACTGGTGC
 CCTCGCTATGAGACTCAATGACGAGGACTTGGATTTCTTATGAAATGGAAGTATTTAAGTGCTTGTCTGCGTCCA
 ACATGAGGAATGACGCGGGTCAATTTATCGAGGCTGCCTATGCTAAAGCACTAGAGTAGAACTGGCCAGTTGGTGACG
 GTTGATAAAGTTGAGGTAATTTGGCCAACTTGAAGCTTTTGTGATAACCGTGGCACCTCACTCTCGCCCGGTGACAT
 TGTGTGCGCTCTCGGCCACCGCTGTTGGCAGTATCTTGCACCTAAGGTTGGTAGCACCAGCATACCTCCAAGCCA
 TTGAGACCAGAGTCTTGTGGTCCAAATGACCGTGGCGCGCTGTCGACCCGACCCCAACGCCCCACCCGCAACCC
 GTGCCATCCCCCTCCACCGAAAGTTCTGGAGAATGGCCCAACGCTTGGGGGATGAGGACCGTTTGAATAAGAAGAA
 GAGGCGCAGGATGGAAGCCCTCGGCATCTATGTTATGGGCGGAAATAACAGAAATTTTGGGACAAGAAATCCGGTG
 ATGTGTTTTATGAGGAGGTCCATAATAACACAGATGAGTGGGAGTGTCTCAGAGTTGGCGACCTGCCGACTTTGACCT
 GAGAAGGAACTCTGTGTGGACATGTACCATTGGAACAAGGCTTACCATGTTTACACCTCCCATCTGGTAAGAAGTT
 CTTGGTCCCCGTCAACCCAGAGAATGGAAGAGTCCAATGGGAAGCTGCAAAGCTTTCGTGGAGCAGGCCCTAGGTATGA

Fig. 1B-3

TGAATGTCGACGGCGAACTGACTGCCAAAGAACTGGAGAACTGAAAAGAATAATTGACAACTCCAGGGCCTGACTAAG
 GAGCAGTGTAACTGCTAGCCGCCAGCGACTTGACCCGCTGTGGTCGCGGCGCTTGGTTGTTACTGAAACAGCGGTA
 AAAATAGTCAAATTTCAACACCGGACCTTCACCTGGGACCTGTGAATTTAAAGTGGCCAGTGAGGTTGAGCTAAAAGA
 CGCGTTGAGCACAACCAACACCCGGTTGCGAGACCGATCGATGGTGGAGTTGTGCTCCTGCGTTCCGCGGTTCTTCGC
 TTATAGACGTCTTGATCTCCGGTGCTGATGCATCTCCCAAGTTACTTGCCCATCACGGGCCGGGAAACACTGGGATCGAT
 GGCACGCTCTGGGATTTTGAAGTCCGAAGCCACTAAAGAGGAAGTCGCACTCAGTGCGCAAATAATACAGGCTTGTGACAT
 TAGGCGCGGCGACGCTCTGAAATTGGTCTCCCTTACAAGCTGTACCCTGTTAGGGGTAAACCTGAGCGGGTGAAGGAG
 TTCTGCAGAATACAAGTTTGGAGACATACCTTACAAAACCCCAAGTGACACTGGAAGCCCAAGTGACACGCGGCTGCCTGC
 CTTACGCCCAACGCCACTCCGGTGACTGATGGGCGCTCCGTCTTGGCCACGACCATGCCCCCGGGTTTGAAGTTATATGT
 ACCGACCATAACAGCGCTGTGCTTATTACCTTGACTCTAGGCTGACTGCCCTAAACAGCTGACAGAGCACGGCTGCG
 AAGATGCGCACTGAAAGACCTCTCTAATATGACTTGTCCACCAAGGCTTTGTTTTACCTGGAGTTCTTCGCTTGTG
 CGGAAATACCTGTTTGCCCATGTAGGTAAGTGCCACCCGTTCTCGGCTTCTACTTACCCTGCTAAGAATTCTATGGC
 TGGAATAAATGGGAACAGGTTCCCAACCAAGGACATTAGAGCGTCCCTGAAATCGACGTTCTGTGCGCACAGGCTGTGC
 GAGAAACTGGCAACTGTCAACCTTGTACTCTTAAGAAACAGTATTGCGGGAAGAAGAAGACTAGGACCATACTCGGC
 ACCAATAACTTCATCGCACTAGCCACCGAGCAGTGTGAGTGGTGTACCCAGGGCTTCATGAAAAAGGCGTTAACTC
 GCCATCGCCCTCGGAAAGAACAGTTTAAGGAGCTACAGACTCCGGTCTGGCAGGTGCCTTGAAGCTGATCTCGCAT
 CCTGCGATCGATCCACGCTGCAATTGTCGCTGGTTTGGCGCAACCTTCTTATGAACTTGCCTGTGCTGAAGAGCAT
 CTACCGTCTGACGTGCTGAACTGCTGCCACGACTTACTGGTCAACGAGTCCGGCGCAGTGACTAAGAGAGGTGGCCTGTC
 GTCTGGCGACCCGATCACCTCTGTGCTAACACCATTTATAGTTTGGTGTATGACACAGCATATGGTGCTTACTTACT
 TCAAAAGTGGTCAACCCCATGGCCTTCTGTTCTTACAAGACAGCTAAAGTTTGAAGCATGCTCAAGGTTCAACCCCTG
 ATCGTCTATTGCGACGACCTCGTGCTGTATGCCGAGTCTCCACCATGCCAACTATCACTGGTGGGTGAACATCTGAA
 TTTGATGCTGGGGTTTCAAGCGGACCCAAAGAGACAGCAATAACAGACTCGCCATCATTTCTAGGCTGTAGAATAATAA
 ATGGGCGCCAGCTAGTCCCAACCGTGACAGGATCCTCGCGGCCCTCGCCTATCACATGAAGGCGAGTAATGTTTCTGAA
 TACTATGCCTCAGCGGCTGCAATACTCATGGACAGCTGTGCTTGTGGAGTATGATCCTGAATGGTTTGAAGAACTTGT
 AGTTGGAATAGCGCAGTGCGCCCGCAAGGACGGCTACAGCTTTCCGGGACGCGGTTCTTATGTCCATGTGGGAAAAAC
 TCAGGTCCAATTATGAGGGGAAGAAGTCGAGAGTGTGCGGGTACTGCGGGGCCCGGGCCCGTACGCTACTGCCTGTGGC
 CTCGAGCTCTGCATTTACCACACCCACTTCCACGAGATTGTCCAGTCACAATCTGGTGTGGCCATCCAGCGGGTTCTGG
 TTCTTGTAGTGAGTGCAATCCCCTGTAGGGAAAGGCACAGCCCTTTAGACGAGGTGCTGGAACAAGTCCCGTATAAGC
 CCCCACGACCGTTATCATGCATGTGGAGCAGGGTCTCACCCCTTGTCCAGGTAGATACCAAACTCGCCGCGGATTGA
 GTCTCTGTGAGGCGTGAATTAGGGGAAATGAAGTTGGACTACCAGACGGTGATTATGCTAGCACCGCCTTGTCCCTAC
 CTGCAAGAGATCAACATGGTGTGCTGCTTCCAATGTATTGCGCAGCAGGTTCTATCATCGGCCACCCGGTGTGGGA
 AAACATACTGGCTCCTTCAACAGGTCCAGGATGGTGTATTTTACACACCAACTACCAGACCATGCTTGACATGATT
 AGGGCTTTGGGGACGTGCCGGTTCAACGTCCCGGACGGCACAACGCTGCAATTCCTCCCTCCCTCCCGCACCGGTCCGTG
 GGTTCCGATCCTAGCCGGCGGTTGGTGTCTGGCAAGAATTCCTTCTAGATGAAGCAGCGTATTGCAATCACCTTGATG
 TTTTGAAGCTTCTAGTAAACTACCCTCCTGTCTAGGAGACTTCAAGCAACTCCACCCAGTGGGTTTGAATCTCAT
 TGCTATGTTTTGACATCATGCCTCAAACTCAACTGAAGACCATCTGGAGGTTTGGACAGAATATCTGTGATGCCATTCA
 GCCAGATTACAGGGACAAACTCATGTCCATGGTCAACACAAACCGTGTGACCTACGTGGAAAAACCTGTGAGGTATGGGC
 AGGTCTCACCCCTACACAGGGACCGAGAGGACGACGCTTACTATTGACTCCAGTCAAGGCGCCACATTGATGTG
 GTTACATTGCATTTGCCACTAAAGATTCACTCAACAGGCAAGAGCCCTTGTGCTATCACCAGGGCAAGACACGCTAT
 CTTTGTGTATGACCCACACAGGACGCTGACGGGCTTGTGATCTTCTGCAAAAGGCACGCGCTCAACCTCGCAGTGC
 ACTGCGACGGGACGCTGATCGTGTGGATAGAAATAACAAAGATGACCGGTTGCTCAGGCTTAGGCAACGGGGATAAA
 TTTAGGGCCACAGACAAGCGTGTGTAGATTCTCTCGCGCCATTTGTGCTGATCTAGAAGGGTCGAGCTCTCCGCTCCC
 CAAGGTGCGACACAATTGGGATTTTATTTCTACCTGATTTAACACAGTTTGCTAACTCCAGTAGAATTGCACCTC
 ACTGGCCCGTGGTGTCAACCCAGAAATGAAAAGTGGCCGGATCGGCTGGTTGCCAGCCTTCGCCCTATCCATAAATAC
 AGCCGCGCGTGATCGGTGCCGGCTATATGGTGGGCCCTTCGGTGTCTAGGCACTCCTGGGGTCTGTGCTACTATCT
 CACAAAATTTGTTAAGGGCGGGCTCAAGTGCTTCCGGAGACGGTTTTCAGCACCGGCCGAATTGAGGTAGACTGCCGGG
 AATATCTTGATGATCGGGAGCGAGAAGTGTGCTGCTCCCTCCACACGCTTTCATTGGCGACGTCAAGGCACTACCGTT
 GGAGGATGTCATCATGTACCTCCAGATACCTCCCGCGCTCTTCCCAAGGAATCAGTTGCGGTAGTCGGGGTTTCAAG

Fig. 1B-4

CCCCAGAAAAGCCGCGAAAGCATTGTGCACACTGACAGATGTGTACCTCCAGATCTTGAAGCCTATCTCCACCCGGAGA
 CCCAGTCCAAGTGCTGGAAAATGATGTTGGACTTCAAAGAAGTTCGACTAATGGTCTGGAAAACAAAACAGCCTATTTT
 CAACTTGAAGTGCCTATTTACCTGGTATCAGCTTGCCAGCTATGCCTCGTACATCCGTGTTCCCGTCAACTCTACGGT
 GACTTGGACCCCTGCATGGCCCCGCCCTTTGCAACAGGAGAGTCTGCGGGTCCACCCACTGGGGGGTACCTCGCGG
 TCACCCCTTATGATTACGGCGCTAAAATTATCTGTCTAGCGCGTACCATGGTGAAATGCCCCCGGATACAAAATTCTG
 GCGTGCAGGAGTTCTCGTTGGATGACCCAGTTAAGTACAAACATACCTGGGGGTTTGAATCGGATACAGCGTATCTGTA
 TGAGTTCAACCGGAAACGGTGAGGACTGGGAGGATTACAATGATGCGTTTCGTGCGCGCCAGGAAGGGAAAAATTTATAAGG
 CCACTGCCACAGCTTGAAGTTTTATTTCCCGGGCCCTGTCAATTGAACCAACTTTAGGCCTGAATTGAAATGAAATG
 GGGTCCATGCAAGCCTTTTTGACAAAATTGGCCAACCTTTTTGTGGATGCTTTCACGGAGTTCTTGGTGTCCATTGTTGA
 TATCATTATATTTTTGGCCATTTTGTGGCTTACCATCGCCGGTGGCTGGTGGTCTTTGTCATCAGATTGGTTTGTCT
 CCGCGATACTCCGTACGCGCCCTGCCATTCCTCTGAGCAATTACAGAAGATCTTATGAGGCCTTTCTTCCAGTGCCA
 AGTGGACATTCACCTCGGGAACTAAACATCCTTTGGGGATGCTTTGGCACCATAAGGTGTCAACCCTGATTGATGAAA
 TGGTGTGCGGTGCAATGTACCGCATCATGAAAAAGCAGGGCAGGCTGCCTGAAACAGGTGGTGAGCGAGGCTACGCTG
 TCTCGATTAGTAGTTTGGATGTGGTGGCTCATTTTACGATCTAGCCGCCATTGAAGCCGAGACCTGTAAATATTTGGC
 CTCCCGGTGCCATGCTACACAACCTGCGCATGACAGGGTCAAATGTAAACATAGTGTATAATAGCACTTTGAATCAGG
 TGTGTTGCTATTTTTCAACCCCTGGTTCGCGGCCAAAGCTTCATGATTTTTCAGCAATGGTTAATAGCTGTACATTCCTCC
 ATATTTTCTCTGTTGCAAGCTTCTTGTACTCTTTTGTGTGCTGTGGTGGCGGTTCCAATACTACGTACTGTTTTGG
 TTTCCGCTGGTTAGGGGCAATTTTTCTTTCGAACTCACAGTGAATTACACGGTGTGTCCACCTTGCTCACC CGGCAAGC
 AGCCACAGAGATCTACGAACCCGGTAGGTCTCTTGGTGAGGATAGGGTATGACCGATGTGGGGAGGAGCATGACG
 AGCTAGGGTTTATGATACCGCTGGCTCTCCAGCGAAGGCCACTTGACTGGTGTTCACGCTGGTTGGCGTTCTTGTCC
 TTCAGCTACACGGCCAGTTCATCCCGAGATATTGCGGATAGGGAATGTGAGTCGAGTTTATGTTGACATCAAAACATCA
 ACTCATCTGCGCCGAACATGACGGGCAGAACACCACCTTGCTCGTCATGACAACATTTAGCCGTGTTTCAGACCTATT
 ACCAACATCAAGTCGACGGCGGCAATTGGTTTACCTAGAATGGCTTCGCTCCCTCTTTTCTCGTGGTTGGTTTAAAT
 GTCTCTTGGTTTCTCAGGCGTTCGCTGCAAAACCATGTTTCAGTTCGAGTCTTGAGATATTAAGACCAACACCACGCA
 GCGGCAAGCTTTGCTGTCTCAAGACATCAGTTGCCTTAGGCATCGGACTCGGCTCTGAGGCGATTGCAAAATCCC
 TCAGTGCCGTACGGCGATAGGGACACCCGTGTATGTTACCACTACAGCCAATGTGACAGATGAGAATTATTTACATTCTT
 CTGATCTCCTCATGCTTTCTTCTGCTTTTCTATGCTTCTGAGATGAGTGAAAAGGGATTTAAGGTGGTATTTGGCAAT
 GTGTGAGGCATCGTGGCTGTGTGTCAATTTTACCAGCTACGTCCAACATGTCAAGGAGTTTACCCAACGCTCCCTGGT
 GGTGCAACATGTGCGGTTGCTCCATTTATGACACCTGAGACCATGAGGTGGGCAACTGTTTATAGCTGTCTTTTGCCA
 TTCTGTTGGCAATTTGAATGTTAAGTATGTTGGAGAAATGCTTGACCGCGGGCTGTTGCTCGCGATTGCTTCTTTGTG
 GTGATCGTGCCGTTCTGTCTTGTGTGCTCGCCAACGCCAGCAACGACAGCAGCTCCCATCTACAGCTGATTTACAAC
 TGACGCTATGTGAGCTGAATGGCACAGATTGGCTAGCTAACAAATTTGATTGGGCAAGTGGAGAGTTTGTATCTTTCCC
 GTTTTGACTCACATTGTCTCTATGGTGCCCTCACTACCAGCCATTTCTTGACACAGTCGCTTTAGTCACTGTGTCTAC
 CGCCGGGTTTGTTCACGGGCGGTATGTCTAAGTAGCATCTACGCGGTCTGTGCCCTGGCTGCGTTGACTTGCTTCGTCA
 TTAGGTTTGCAAAGAATTGCATGTCTGGCGCTACGCGTGTACCAGATATACCACTTTCTTCTGGACACTAAGGGCAGA
 CTCTATCGTTGGCGGTGCGCTGTATCATAGAGAAAAGGGGCAAGTTGAGGTGGAAGGTCTATGATCGACCTCAAAAG
 AGTTGTGCTTGTGCTGCTGCGTGGCAACCCCTATAACCAGAGTTTCAGCGGAACAATGGGGTCTGCTTATAGTACTCTG
 TCACGATAGCACGGCTCCACAAAAGGTGCTTTTGGCGTTTCTATTACCTACACGCCAGTGATGATATATGCCCTAAAGG
 TGAGTCGCGGCCGACTGCTAGGGCTTCTGCACCTTTTATCTTCTGAATTGTGCTTTACCTTCGGGTACATGACTTTC
 GCGCACTTTAGAGTACAAATAAGGTGCGGCTCACTATGGGAGCAGTAGTTGCACTCCTTTGGGGGGTGTACTCAGCCAT
 AGAAACCTGGAAATTCATCACCTCCAGATGCCGTTTGTGCTTGTAGGCGCGAAGTACATTCTGGCCCTGCCACCCAG
 TTGAAAGTGCCGAGGCTTTATCCGATTGCGGCAATGATAACCACGCATTTGTCGTCGCGGCTCCCGGCTCCACTACG
 GTCAACGGCACATTTGGTGGCGGTTAAAAAGCCTCGTGTGGGTGGCAGAAAAGCTGTTAAACAGGGAGTGGTAAACCT
 TGTCAAATATGCCAAATAACAACGGCAAGCAGCAGAGAAGAGAAAGAGGGGGATGGCCAGCCAGTCAATCAGCTGTGCCAG
 ATGCTGGGTAAAGTATCACTCAGCAAAACAGTCCAGAGGCAAGGGACCGGGAAAGAAAAATAAGAAGAAAAACCCGGA
 GAAGCCCCATTTCTCTAGCGACTGAAGTGTGACACATCACTTTACCCCTAGTGAGCGGCAATTTGTGTCTGTCTG
 CAATCCAGACCGCCTTAAATCAAGGCGCTGGGACTTGACCCCTGTGAGTTTCAAGGAGGATAAGTTACTGTGGAGTTT
 AGTTTGCCTACGCATCATACTGTGCGCTGATCCGCGTCAACAGCATACCCTCAGCATGATGGGCTGGCATTCTTGAGGC

Fig. 1B-5

ATCTCAGTGTTTGAATTGGAAGAATGTGTGGTGAATGGCACTGATTGACATTGTGCCTCTAAGTCACCTATTCAATTAGG
GCGACCGTGTGGGGGTGAGATTTAATTGGCGAGAACCATGCGGCCGAAATTA
AA
AAAAAAAAAAAAAAAAAAAA

Fig. 1C-1

>VR-V5G7475A.seq

ATGACGTATAGGTGTTGGCTCTATGCCTTGGCATTTGTATTGTCAGGAGCTGTGACCATTGGCACAGCCAAAACCTTGCT
 GCACAGAAACACCCCTTCTGTGATAGCCTCCTTCAGGGGAGCTTAGGGTTTGTCCCTAGCACCTTGCTCCGGAGTTGCAC
 TGCTTTACGGTCTCTCCACCCCTTTAACCATGTCTGGGATACTTGATCGGTGCACGTGTACCCCAATGCCAGGGTGT
 ATGGCGGAGGGCCAAAGTCTACTGCACACGATGCCTCAGTGCACGGTCTCTCCTTCCCCTGAACCTCCAGGTTTCTGAGCT
 CGGGGTGCTAGGCTATTCTACAGGCCCGAAGAGCCACTCCGGTGGACGTTGCCACGTGCATTCCCCACTGTTGAGTGCT
 CCCCCCGGGGCTGCTGGCTTTCTGCAATCTTTCAATCGCACGAATGACCAGTGGAAACCTGAACTTCCAACAAAGA
 ATGGTACGGGTGCGAGCTGAGCTTTACAGAGCCGGCCAGCTCACCCCTGCAGTCTTGAAGGCTCTACAAGTTTATGAACG
 GGGTTGCCGTGGTACCCATTGTTGGACCTGTCCCTGGAGTGGCCGTTTTCGCCAATTCCTACATGTGAGTGATAAAC
 CTTCCCGGGAGCAACTCAGTGTTGACCAACCTGCCGCTCCCGCAGAGACCAAGCCTGAAGACTTTTGCCCTTTGAG
 TGTGCTATGGCTACTGTCTATGACATTGGTCATGACGCCGTGTATGTATGTGGCCGAAAGGAAAGTCTCCTGGGCCCCCTCG
 TGGCGGGGATGAAGTGAATTTGAAGCTGTCCCGGGGAGTTGAAGTTGATTGCGAACCGGCTCCGCACCTCCTTCCCGC
 CCCACACACAGTGGACATGTCTAAGTTCGCTTACAGCCCTGGGTGTGGTGTCTATGCGGGTCGAACGCCAACAC
 GGCTGCCCTCCCGTGACACTGTCCCTGAAGGCACTGTGGTGGAGCTTGTGTTGACTTGCTTCCACTGGAAGTTCAGAA
 CAAAGAAATTCCGCATGCTAACCAATTTGGTACACAGCAAGCATGGTGTCTCTGGCAAGTACCTGCAGCGGAGGCTGC
 AAGTTAATGGTCTCCGAGCAGTAACTGACCTAACCGACCTATCGTCTGACAGTACTTCTCCGTTAAGGAGAGTTGGATC
 CGCCATTTGAAACTGGCGGGAGAACCCAGCTACTCTGGGTTTGGAGACCTCCTCAGAATAAGGGTTGAGCCTAACACGTC
 GCCATTGGCTGACAGGAAGAAAAAATTTCCGGTTTGGCAGTCACAAGTGGTACGGCGCTGGAAAGAGAGCAAGAAAAAG
 CACGCTCTTGTGCGACTGCTACAGTGCCTGGCCGCGCTTGTCCGTTCTGTGAAACCCGGCAGGCCAAGGAGCACGAGGT
 GCCCGGCCAACAAAGGCTGAGCACCTCAACACTACTCCCCGCTGCCGAAGGGAATTGTGGTTGGCACTGCATTTCCGC
 CATCGCAACCGGATGGTGAATTCCAAATTTGAAACCAACCTTCCGAAAGAGTGAACCTCCAGATGACTGGGCTACTG
 ACGAGGATCTTGTGAATGCCATCAAATCCTCAGACTCCCTGCGGCTTAGACAGGAACGGTGCTGTACTAGCGCCAAG
 TACGTACTTAAGCTGGAAGGTGAGCATTGGAAGTGTACTGTGACCCCTGGGATGTCCCTTCTTGTCTCCTCTTGAATG
 TGTTCAGGGCTGTTGTGGCGCAAGGGCGGTCTTGGTTCCTCAGATGCAGTCGAGGTCTCCGGATTGACCTGCCTGCC
 TTGACCGGCTGGTGAAGTGTGACCTGCCTAGCAGTGCTATCCAGCCGCTCTGGCCGAAATGTCTGGCGATTCCGAT
 CGTTCCGCTTCTCCGGTCAACCCGTGTGGAAGTGTTCGAGTCTTTCGCCGTCACAGCGGAGGGAATCACCTGACCA
 AGTGGCTTAGGGAAAAATTTATCAGCCTTTGTGAGGTGATTGAGGACTGTGCTGTTCCAGAACAAAAACCAACCGGTCA
 CCCCCGAGGAGGTGCGAGCAAGATTGACCTGTACCTCCGTGGTGAACAAATCTTGAAGAATGCTTGGCCAGGCTTGA
 AAAGCGCGCCGCCACCGCTAATCGACACCTCCTTGATTGGGATGTTGTGCTCCTGGGTTGAGGCGGCAACCCAGAC
 GATCAAGCTGCCCGAGGTCAACAGTGTCTGCTGTGGTCCCTGTTGTGACTCAAAAGTCTTGGACAACAACCTCGGTCC
 CCTGACCGCTTTTCACTGGCTAACTACTACCGTGGCGCAAGGTGACGAAGTTCTGTACCGTGAAAGACTAACCGCC
 GTGCTCTCAAGTTGAAAAAGGTTGTTGAGAAGAATATGGGCTCATGCCAACCGAGCCTGGTCCACGGCCCACTGCC
 ACGCGGGCTCGACGAACTCAAAGACCAAGTGGAGGAGGACTTGTGAAACTGGCTAACGCCAGACGACTTCGGACATGA
 TGGCCTGGGAGTCGAGCAGGTTGACCTAAAACTTGGGTCAAGAACTACCCGCGGTGGACACCACACCCCTCCGCCA
 AAAGTTACGCTCGAAAAACGAAGCTGTCAAGAGCTTGGCGGAGAGAAAGCTGTCCCCGCCCGCGCAGGAAGGTTGG
 GTCCGATTGTGGCAGCCCGGTTTCATTAGCGGGGATGTCCCTAACAGTTGGGAAGATTGGCTGTAGTAGCCCTTTG
 ATCTCCCGACCCACCTGAGCCGGCAACACCTTCAAGTGAGCTGGTGATTGTCTCACCAGCAATGCATCTTACGGCCG
 GCGACACCTTGAGTGAGCCGGCTCAATTTCCCGCACTCGCGGAAGTGTGTCTGACCCGGTGACACCTTGAGTGAGCC
 GATCCCTGTGCCCGACCGCGGCTAAGTTTACGAGGTGAAAAGATTGAGTTCGGCGGGCGCAATCCACCGTACCAAG
 ACGAGCCCTGGATTTGTCTGCTTCTCAGAGACTGAATATGAGGCTCTCCCCAGCAACCGCCGAGAGCGGGGGCTT
 CTGGGAGTAGAGGGCATGAAGCTGAGGAAACCTGAGTGAAATCTCGGACATGTGGGTAACATTAAACCTGCGTCCGT
 GTCATCAAGCAGCTCCTTGTCCAGCGTGAGAATCACACGCCAAAATACTCAGCTCAAGCCATCATCGACTCGGGCGGGC
 CCTGCAGTGGGCATCTCAAGAGGTAAAGGAAACATGCCTTAGTGTATGCGCGAGGCATGTGATGCGACTAAGCTTGAT
 GACCTGCTACGCAGGAATGGCTTTCTCGCATGTGGGATCGGGTGGACATGCTGACTTGGCGCAACACGCTGTGTTACCA
 GCGGATTTGACCTTAGATGGCAGGTTAAAGTTCCTCCCAAAATGATACTCGAGACACCGCCGCCCTATCCGTGTGAGT
 TTGTGATGATGCTTCAACCGCTGCCTTCCGTAGGTGCGGAGAGCGACCTTACCATTGGCTCAGTTGCTACTGAAGAT
 GTTCCACGCATCTCTGAGAAAAATAGAAAATGTGCGCGAGATGGCCAAACAGGGACCTTGGCTTCTCCGAGGATAAAC
 GGATAGATGACCAACTGTCAACGACCCCGGATATCGTCGCGGAGGCTGACGAGAGCACATCAGCTCCGTCCGAGGCA

Fig. 1C-2

CAGGTGGCGCGGCTCTTTTACCGATTGCGCGCTTCAGATGGCGCGGATGCGGACGGGGGGGGCGCTTCGGACGGTA
 AAAAGAAAAGCTGAAAGGCTCTTTGACCAACTGAGCCGTGAGGTTTTGACCTCGTCTCCCATCTCCCTGTTTTCTCTC
 ACGCCTTTTCTACCTGGCGGTGGTTATTCTCCGGGTGATTGGGGTTTTGCAGCTTTTACTCTATTGTGCCTCTTTTAT
 GTTACAGTTACCCAGCCTTGGTATTGCTCCCTCTTGGGTGTGTTTTCTGGGTCTTCTCGGCGGTTGGAATGGGGGT
 TTTGGCTGCTGGTTGGCTTTTGTCTGTTGGTCTGTTCAAGCCTGTGTCCGACCCAGTCGGCGCTGCTTGAGTTTGACTC
 GCCAGAGTGTAGAAACATCCTTATTCTTTGAGCTTCTCAAACCTTGGGACCCTGTTGCGAGCCTTGTGTGGGCCCCG
 TCGGTCTCGGTCTTGCCATTCTTGGCAGGTTACTGGGCGGGGACGCTGCATCTGGCACTTTTTGCTTAGGCTTGGCATT
 GTTGACAGACTGTATCTTGGCTGGAGCTTACGTGCTTCTCAAGGTAGGTGTAAAAAGTCTGGGGATCTGTATAAGAAC
 TGCTCCCAATGAGGTGCTTTTAACTGTTTCTTTTACACGTGCGACCGAGTCTGCACTTATCGACCTGTGCGATCGGT
 TTTGTGCGCCAAAAGGAATGGACCCCATTTTTCTCGCCACTGGGTGGCGCGGGTGTGGGCGGCGGAAGCCCCATTGAG
 CAACCTCTGAAAAACCATCGCGTTTGGCCAGTTGGATGAAAAGAAGATTACGGCTAGGACTGTGGTCGCCAGCCTTA
 TGACCCCAACCAAGCCGTAAAGTGCTTGGGGTATTGCAGCGGGTGGGGCGATGGTGGCTAAGGCGGTCCAAAAGTGG
 TCAAGGTTTCCGCTGTTCCATTCCGAGCCCCCTTTTCCCACTGGAGTGAAAGTTGACCCTGATTGCAGGGTCGTGGT
 GACCCTGACACTTTCACTGCAGCTCTCCGGTCTGGCTACTCCACCACAAACCTCGTCTTGGTGTGGGGGACTTTGCCA
 GCTGAATGGATTAATAATCAGGCAAAATTCAGGCTTCAAGGGGAGGCCACATCTCATGGCTGCCCTGCATGTTGCCCT
 GCTCGATGGCTCTGCACATGCTTGGGATTATGTGACTGCGGTGGGTCTTGGCGCACCGGACCAACGACCCGTGG
 TGCGTAACCCGTTTGGCGTCCCTGGCTACGGACCTGGCTCTCTGCACGTCCAGATTGTGCATTTCCCAACACGGCT
 TACCCTGCCCTTGACAGCACTTGTGGCGGATTGGTATTCAAGAAATTCCTTGGTGGTTTTGATTTTTGTTTCCATCG
 GAGGCATGGCTCATAGGTTGAGCTGTAAGGCTGACATGCTGTGTCTTGTGCAATTGCCAGCTATGTTTGGGTACCT
 CTTACCTGGTTGCTTTGTGTGTTTCTTGGTGGTGGCTGTTTTCTTGCACCCCTCACCATCCTATGGTTGGTGT
 TTTCTTGATTTCTGTGAATATGCCTCAGGAATCTTGGCCATGGTGTGTGGTTTCTTTGGCTTCTTGGTGGTTATA
 CTAATGTTGCTGGCTTGTGACCCCTACGACATTCACTTAACACAGTGGCCCCGCGGTGTTGCCGCTTGGCTACC
 GCACAGATGGGACCTACTTGGCGCTGTCCGCCGCTGCGTTGACTGGCCGACCATGCTGTTTACCCGCTCCAGCT
 TGGGTCTCTTCTGAGGGTGCTTTCAGAACTCGAAAGCCCTCACTGAACACCGTCAATGTGATCGGGTCTCCATGGGT
 CTGGCGGGGTGTTTACCATCGACGGGAAAGTCAAGTGCGTAACTGCCGCACATGTCCTTACGGGCAATTAGCTCGGGT
 TCCGGGGTGGCTTCAATCAAATGCTTGACTTTGACGTAAGGGGAGATTCGCTATAGCTGATTGCCGAATTGGCAAGG
 GGCTGCCCCAAGACCAATTCTGCACGGATGGATGGACTGGCCGTGCCTATTGGCTAACATCCTTGGCGTCGAACCCG
 GCGTCAATTGAAAAGGATTGCGCTTCTGCTTCAACCGATGTGGCGATTCCGGGTCCCAAGTATCACCAGGCGGGTGA
 CTTGTGGCGTTTACACGGGATCGAATAAACAAGGGGGGGGATTGTTACGCGCCCTCAGGCCAGTTTTGTAATGTGGC
 ACCCATCAAGCTAAGCGAATTAAGTGAATCTTTGCTGGGCTAAGTCCGCTCGGTGATGTGAAGGTGGCAGGCCACA
 TAATTAAGACATAAGCGAGTGCTTTCAGATCTTTGCTGCTTGTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT
 ACCGTCCAACCTCTTTGTGTGTTTTCTCTGTGGAGAATGATGGGACATGCTGGACGCCCTTGGTGTGTGAGTTT
 CTTTATTTTGAATGAGGTTCTCCAGCGCTCTGGTCCGGAGTGTCTCTCTTGAATGTTTGTGCTATCTGGCTCA
 CGCCATGGTCTGCGCAAGTTCTGATGATCAGGCTTCTGACAGCAGCTCTAACAGGAACAGATGGTCACTTGCCTTTTCT
 AGCCTCGGTGCAGTGACCGTTTTGTGCGAGATCTTGGGCCACTCAGGGGCATCCGTTGCAGGCAGTGATGAATTTGAG
 CACCTATGCATTCTGCCTCGGATGATGGTTGTGACCTCACCAGTCCAGTGATCACGTGTGGTGTGTCGACCTACTTG
 CCATCATTTTGTACTTGTAAAGTACCGTGGCTGCACCATATCTTGTGGCGATGGAGTGTCTCTGCGGCTTTCTTC
 TTGAGATACTTTGCCGAGGGAAAGTTGAGGGAAGGGGTGTGCAATCTGCGGAATGAATCATGAGTCTCTGACTGGTGC
 CCTCGCTATGAGACTCAATGACGAGGACTTGGATTTCTTATGAAATGGACTGATTTTAAGTGCTTTGTTCTGCGTCCA
 ACATGAGGAATGCAGCGGGTCAATTTATCGAGGCTGCCTATGCTAAAGCACTAGAGTAGAACTGGCCAGTTGGTGCAG
 GTTGATAAAGTTCGAGGTACTTTGGCCAACTTGAAGCTTTTGTGATACCGTGGCACCTCAACTCTCGCCCGGTGACAT
 TGTGTGCTCTCGGCCACACGCTGTTGGCAGTATCTCGACCTAAAGGTTGGTAGCACCAGCATACCTCCAAGCCA
 TTGAGACCAGAGTCTTGTGGGTCCAAATGACCGTGGCGCGCTGTCGACCCGACCCCAACCCCAACCCGACCC
 GTGCCCATCCCCCTCCACCGAAAGTTCTGGAGAATGGCCCAACGCTTGGGGGGATGAGGACCGTTTGAATAAGAAGAA
 GAGGCGCAGGATGGAAGCCCTCGGCATCTATGTTATGGGCGGGAAAAATACCAGAAATTTGGGACAAGAATTCCGGTG
 ATGTGTTTTATGAGGAGGTCCATAATAACACAGATGAGTGGGAGTGTCTCAGAGTTGGCGACCTGCCGACTTTGACCT
 GAGAAAGGAACTCTGTGTGGACATGTACCATTTGAAACAAGGCTTACCATGTTTACACCTCCCATCTGGTAAGAAGTT
 CTTGGTCCCCGTCAACCCAGAGAATGGAAGAGTCAATGGGAAGCTGCAAAGCTTCCGTGGAGCAGGCCCTAGGTATGA

Fig. 1C-3

TGAATGTCGACGGCGAACTGACTGCCAAAGAACTGGAGAACTGAAAAGAATAATTGACAACTCCAGGGCCTGACTAAG
 GAGCAGTGTTTAACTGCTAGCCGCCAGCGACTTGACCCGCTGTGGTCGCGGCGGCTTGTTGTTACTGAAACAGCGGTA
 AAAATAGTCAAATTTACAACCGGACCTTACCCTGGGACCTGTGAATTTAAAGTGGCCAGTGAGGTTGAGCTAAAAGA
 CGCGGTTGAGCACAAACACCCGGTTGCGAGACCGATCGATGGTGGAGTTGTGCTCCTGCGTTCGCGGTTCTTCGC
 TTATAGACGCTTGATCTCCGGTGCTGATGCATCTCCAAGTTACTTGCCCATCACGGGCCGGGAAACTGGGATCGAT
 GGCACGCTCTGGGATTTTGTGTCGAAGCCACTAAAGAGGAAGTGCAGTCACTGCGCAAATAATACAGGCTTGACAT
 TAGGCGCGGCGAGCTCTGAAATTGGTCTCCCTTACAAGCTGTACCCTGTTAGGGGTAACCTGAGCGGGTGAAAGGAG
 TTCTGCAGAATAAAGGTTTGGAGACATACCTTACAAAACCCCACTGACACTGGAAGCCCACTGCACGCGGCTGCCTGC
 CTTACGCCCAACGCCACTCCGGTGAAGTGTGGGCGCTCCGTCTTGCCACGACCATGCCCCCGGGTTTGTGTTATATGT
 ACCGACCATACAGCGTCTGTCTTGATTACCTGACTCTAGGCTGACTGCCCTAAACAGCTGACAGAGCACGGCTGCG
 AAGATGCCGCACTGAAAGACCTCTAAATATGACTTGTCCACCCAAGGCTTTGTTTACCTGGAGTTCTTCGCTTGTG
 CGGAAATACCTGTTTGCCCATGTAGGTAAGTGCCACCCGTTTCATCGGCTTCTACTTACCCTGCTAAGAATTCTATGGC
 TGGAATAAATGGGAACAGGTTCCCAACCAAGGACATTCAGAGCGTCCCTGAAATCGACGTTCTGTGCGCACAGGCTGTGC
 GAGAAAACCTGGCAAATGTACCCCTTGTACTCTTAAGAAACAGTATTGCGGGAAGAAGAAGACTAGGACCATACTCGGC
 ACCAATAACTTCATCGCACTAGCCACCGAGCAGTGTGGTGGTGTACCCAGGGCTTCATGAAAAAGGCGTTTAACTC
 GCCCATCGCCCTCGGAAAGAACAAAGTTAAGGAGCTACAGACTCCGGTCTGGGCGAGTGCCTTGAAGCTGATCTCGCAT
 CCTGCGATCGATCCACGCTGCAATTGTCCGCTGGTTTGCCGCCAACCTTCTTTATGAAGTTCCTGTGCTGAAGAGCAT
 CTACCGTCTGACGTGCTGAAGTCTGCCACGACTTACTGGTCAAGCAGTCCGGCGCAGTGACTAAGAGAGGTGGCTGTC
 GTCTGGCGACCCGATCACCTCTGTGTCTAACACCATTTATAGTTTGGTGTCTATGCACAGCATATGGTGTCTAGTACT
 TCAAAAGTGGTCAACCCCATGGCTTCTGTTCTTACAAGACCAGCTAAAGTTTGAGGACATGCTCAAGGTTCAACCCCTG
 ATCGTCTATTGCGACGACCTCGTGTGTATGCCGAGTCTCCACCATGCCAACTATCACTGGTGGGTTGAACATCTGAA
 TTTGATGCTGGGGTTTCAGACGGACCCAAAGAAGACAGCAATAACAGACTCGCCATCATTTCTAGGCTGTAGAATAATAA
 ATGGGCGCCAGTAGTCCCAACCGTGACAGGATCCTCGCGGCCCTCGCTATCACATGAAGGCGAGTAATGTTTCTGAA
 TACTATGCCTCAGCGGCTGCAATACTCATGGACAGCTGTGCTTGGTGGAGTATGATCCTGAATGGTTTGAAGAACTTGT
 AGTTGGAATAGCGCAGTGCGCCCGCAAGGACGGCTACAGCTTCCCGGACGCGGTTCTTCATGTCCATGTGGGAAAAAC
 TCAGGTCCAATTATGAGGGGAAGAAGTCAAGAGTGTGCGGGTACTGCGGGGCCCGGCCCGTACGCTACTGCTGTGGC
 CTCGACGTCTGATTTTACCAACCCACTTCCACAGCATTGTCCAGTCACAATCTGGTGTGGCCATCCAGCGGGTCTGG
 TTCTTGATGAGTGAGTGCAATCCCTGTAGGGAAAGGCACAAGCCCTTAGACGAGGTGCTGGAACAAGTCCCGTATAAGC
 CCCCACGGACCGTTATCATGCATGTGGAGCAGGCTCACCCCTTGTATCCAGGTAGATACCAACTCGCCGCGGATTA
 GTCTCTGTGAGCGTGGAAATAGGGGAAATGAAGTTGGACTACCAGACGGTGATTATGCTAGCACCCTGCTCCCTAC
 CTGCAAGAGATCAACATGGTGTGCTGCTTCAATGTATTGCGCAGCAGGTTTCATCATCGGCCACCCGCTGCTGGGA
 AAACATACTGGCTCCTTCAACAGGTCCAGGATGGTGTATTATACACCAACTCACCAGACCATGCTTGACATGATT
 AGGGCTTTGGGACGTGCCGTTCAACGTCCCGGACGGCACAAGCTGCAATCCCCGCTCCCTCCCGCACCGGTCGGTG
 GGTTCCGATCCTAGCCGCGGTTGGTGTCTGGCAAGAATCTTCTAGATGAAGCAGCGTATTGCAATCACCTTGATG
 TTTTGAAGCTTCTAGTAAACTACCTCACCTGTCTAGGAGACTCAAGCAACTCCACCACTGGGTTTGTATTCTCAT
 TGCTATGTTTTGACATCATGCTCAAACTCACTGAAGACCATCTGGAGGTTTGGACAGAATATCTGTGATGCCATTCA
 GCCAGATTACAGGGACAACTCATGTCCATGGTCAACACAACCCGTGTGACCTACGTGAAAAACCTGTCAAGGATGGGC
 AGGTCTCACCCCTACCAAGGACCGAGAGGACGACGCTACTATTGACTCAAGTCAAGGCGCCACATTGATGTG
 GTTACATTGCATTTGCCCACTAAAGATTCACTCAACAGGCAAGAGCCCTTGTGCTATCACAGGGCAAGACAGCTAT
 CTTTGTGATGACCCACAGGAGCTGCAGGGCTTGTGTTGATCTTCTGCAAAAGGCACGCGCTCAACCTCGCAGTGC
 ACTGCGACGGGACGCTGATGCTGCTGGATAGAAATAACAAAGATGCACGTTGCTCAGGCTTAGGCAACGGGGATAAA
 TTAGGGCCACAGACAAGCGTGTGTAGATTCTCTCGCGCCATTGTGCTGATCTAGAAGGGTCGAGCTCTCGCTCCC
 CAAGGTGCGACACAACCTGGGATTTTATTTCTACCTGATTTAACACAGTTTGCTAAACTCCAGTAGAACTTGACCTC
 ACTGGCCCGTGGTGTCAACCCAGAACATGAAAAGTGGCCGGATCGGCTGGTGGCAGCTTTCGCCCTATCCATAAATAC
 AGCCGCGCGTGATCGGTGCCGGCTATATGGTGGGCCCTTCGGTGTCTAGGCACTCCTGGGGTGTGTCATACTATCT
 CACAAAATTTGTTAAGGGCGGGCTCAAGTGCTTCCGAGACGGTTTTGAGCACCAGGCGAATTGAGGTAGACTGCCGGG
 AATATCTTGATGATCGGGAGCGAGAAGTTGCTGCTCCCTCCACACGCTTTCATTGGCGACGTCAAGGCACTACCGTT
 GGAGGATGTCATCATGTACCTCCAGATACCTCCCGCGCTCTTCCCAAGGAATCAGTTGCGGTAGTCGGGTTTCAAG

Fig. 1C-4

CCCCAGAAAAGCCGCGAAAGCATTGTGCACACTGACAGATGTGTACCTCCAGATCTTGAAGCCTATCTCCACCCGGAGA
 CCCAGTCCAAGTGCTGGAAATGATGTTGGACTTCAAAGAAGTTCGACTAATGGTCTGGAAAGACAAAACAGCCTATTTTC
 CAACTTGAAGGTCGCTATTTACCTGGTATCAGCTTGCCAGCTATGCCTCGTACATCCGTGTTCCCGTCAACTCTACGGT
 GTACTTGGACCCCTGCATGGGCCCCGCCCTTTGCAACAGGAGAGTCTGCGGGTCCACCCACTGGGGGGCTGACCTCGCGG
 TCACCCCTTATGATTACGGCGCTAAAATTATCCTGTCTAGCGCGTACCATGGTGAAATGCCCCCGGATACAAAATTCTG
 GCGTGCGCGGAGTTCGTTGGATGACCCAGTTAAGTACAAACATACCTGGGGGTTTGAATCGGATACAGCGTATCTGTA
 TGAGTTACCCGGAACGGTGAGGACTGGGAGGATTACAATGATGCGTTTCGTGCGCGCCAGGAAGGGAAAATTTATAAGG
 CCACTGCCACAGCTTGAAGTTTTATTTCCCGGGGCCCTGTATTGAACCACTTTAGGCCTGAATTGAAATGAAATG
 GGGTCCATGCAAAGCCTTTTGACAAAATTGGCCAACCTTTTGTGGATGCTTTCACGGAGTTCCTGGTGTCCATTGTTGA
 TATCATTATATTTTGGCCATTTTGTGGCTTACCATCGCCGGTGGCTGGTGGTCTTTGCATCAGATTGGTTTGCT
 CCGCGATACTCCGTACGCGCCCTGCCATTCCTCTGAGCAATTACAGAAGATCTTATGAGGCCTTTCTTTCCAGTGCCA
 AGTGGACATTTCCACCTGGGGAACATAACATCTTTGGGGATGCTTGGCACCATAAGGTGTCAACCTGATTGATGAAA
 TGGTGTGCGGTGCAATGTACCGCATCATGAAAAAGCAGGGCAGGCTGCCTGGAACAGGTGGTGAGCGAGGCTACGCTG
 TCTCGCATTAGTAGTTTGGATGTGGTGGCTCATTTTTCAGCATCTAGCCGCCATTGAAGCCGAGACCTGTAATATTTGGC
 CTCGCGGTGCCCATGTACACAACCTGCGCATGACAGGGTCAAATGTAACCATAGTGATAATAGCACTTTGAATCAGG
 TGTTTGCTATTTTCCAACCCCTGGTCCCGGCCAAAGCTTCATGATTTTCAGCAATGGTTAATAGCTGTACATTCTCTCC
 ATATTTCTCTGTTGCAGCTTCTTGTACTCTTTTGTGTGCTGTGGTTGCGGGTCCAATACTACGTACTGTTTTTGG
 TTTCCGCTGGTTAGGGGCAATTTTCTTTCGAAGTACAGTGAATTACACGGTGTGTCCACCTTGCTCACCCTGCAAGC
 AGCCACAGAGATCTACGAACCCGGTAGGTCTCTTGGTGCAGGATAGGGTATGACCGATGTGGGGAGGATCATGACG
 AGCTAGGGTTTATGATACCGCTGGCTCTCCAGCGAAGGCCACTTGACTGGTGTACGCTGGTTGGCGTTCTTGTCC
 TTCAGCTACACGGCCAGTTCATCCCGAGATATTCGGGATAGGGAATGTGAGTCGAGTTTATGTTGACATCAAACATCA
 ACTCATCTGCGCCGAACATGACGGGCGAAGACACCACTTGCTCGTCATGACAACATTTACGCCGTGTTTACAGCCTATT
 ACCAACATCAAGTCGACGGCGGCAATTGGTTTACCTAGAATGGCTTCTGCTCTTCTCTCTGTTGGTTTAAAT
 GTCTCTTGGTTTCTCAGGCGTTGCGCTGCAACCATGTTTTCAGTTTCGAGTCTTGCAAATATTAAGACCAACACCAACGCA
 GCGGCAAGCTTTGCTGTCTCCAAGACATCAGTTGCTTAGGCATCGCGACTCGGCCTGAGGGCGATTGCAAAATCCC
 TCAGTGCCGTACGGCGATAGGGACACCCGTGTATGTTACCATCACAGCCAATGGGACAGATGAGAATTATCTACATTCTT
 CTGATCTCTCATGCTTCTTCTTGCCTTTCTATGCTTCTGAGATGAGTGAAAAGGGATTAAAGGTGGTATTTGGCAAT
 GTGTGAGGATCGTGGCTGTGTGTCAATTTTACCAGCTACGTCCAACATGTCAAGGAGTTTACCAACGCTCCCTGGT
 GGTCGACCATGTGCGGTGCTCCATTTTCATGACACCTGAGACCATGAGGTGGGCAACTGTTTACGCTGTCTTTTGGCA
 TTCTGTTGGCAATTTGAATGTTAAGTATGTTGGAGAAATGCTTGACCGCGGGCTGTTGCTCGCGATTGCTTCTTTGTG
 GTGTATCGTGCGCTTCTGTCTGTGCTCGCCAACGCCAGCAACGACAGCAGCTCCCATCTACAGCTGATTTACAAC
 TGACGCTATGTGAGCTGAATGGCACAGATTGGTAGCTAACAAATTTGATTGGGAGTGGAGAGTTTGTATCTTTCCC
 GTTTTGACTCACATTGTCTCTATGGTGGCTCACTACCAGCATTTCCTTGACACAGTCGCTTTAGTCACTGTGTCTAC
 CGCCGGGTTTGTTCACGGGCGGTATGTCCTAAGTAGCATCTACGCGGTCTGTGCCCTGGCTGCGTTGACTTGCTTCGTCA
 TTAGGTTTGCAAAGAATTGCATGTCTGGCGCTACGCGTGTACAGATATACCACTTTCTTCTGGACACTAAGGGCAGA
 CTCTATCGTTGGCGGTGCGCTGTATCATAGAGAAAAGGGGCAAAGTTGAGGTGGAAGGTCTGCTCTAGATGACTTCTG
 AGTTGTGCTTGATGGCTCCGTGGCAACCCCTATAACCAGAGTTTCAGCGGAACAATGGGGTCTGCTCTAGATGACTTCTG
 TCACGATAGCAGGCTCCACAAAAGGTGCTTTTGGCGTTTCTATTACCTACACGCCAGTGATGATATATGCCCTAAAGG
 TGAGTCGCGGCCGACTGCTAGGGCTTCTGCACCTTTTGTCTTCTGAATTGTGCTTTCACCTTCGGGTACATGACTTTC
 GCGCACTTTCAGAGTACAAATAAGGTGCGGCTCACTATGGGAGCAGTAGTTGCACTCTTTGGGGGTGTAAGTCAAGCAT
 AGAAACCTGGAATTCATCCTCCAGATGCCGTTTGTGCTTGTAGGCCGCAAGTACATTCTGGCCCTGCCACCACG
 TTGAAAGTGCCGCGAGGCTTTCATCCGATTGCGGCAATGATAACCAGCATTTGTCGTCCGGCGTCCCGGCTCCACTACG
 GTCAACGGCACATTGGTGCCCGGTTAAAAAGCCTCGTGTGGGTGGCAGAAAAGCTGTTAAACAGGGAGTGGTAAACCT
 TGTCAAATATGCCAAATAACAACGGCAAGCAGCAGAAGAGAAAGAGGGGGATGGCCAGCCAGTCAATCAGCTGTGCCAG
 ATGCTGGGTAAAGTATCACTCAGCAAAACAGTCCAGAGGCAAGGACCGGGAAAGAAAAATAAGAAGAAAAACCCGGA
 GAAGCCCCATTTTCTCTAGCGACTGAAGATGATGTACAGATCACTTTACCCCTAGTGAGCGGCAATTGTGTCTGCTGT
 CAATCCAGACCGCTTAAATCAAGCGCTGGGACTTGACCTGTGAGATTGAGGAGGATAAGTTACACTGTGGAGTTT
 AGTTTGCCTACGCATCACTGTGCGCTGATCCGCGTCACAGCATCACCTCAGCATGATGGGCTGGCATTCTTGAGGC

Fig. 1C-5

ATCTCAGTGTTTGAATTGGAAGAATGTGTGGTGAATGGCACTGATTGACATTGTGCCTCTAAGTCACCTATTCAATTAGG
GCGACCGTGTGGGGGTGAGATTTAATTGGCGAGAACCATGCGGCCGAAATTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAACCC

Fig. 1D-1

>VR-V6.seq

ATGACGTATAGGTGTTGGCTCTATGCCTTGGCATTGTATTGTGTCAGGAGCTGTGACCATTGGCACAGCCCAAACTTGCT
 GCACAGAAACACCTTCTGTGATAGCCTCCTTCAGGGGAGCTTAGGGTTTGTCCCTAGCACCTTGCTTCCGGAGTTGCAC
 TGCTTTACGGTCTCTCCACCCCTTTAACCATGTCTGGGATACTTGATCGGTGCACGTGTACCCCAATGCCAGGGTGT
 ATGGCGGAGGGCCAAAGTCTACTGCACACGATGCCTCAGTGACGGTCTCTCCTTCCCCTGAACCTCCAGGTTTCTGAGCT
 CGGGGTGCTAGGCTATTCTACAGGCCCGAAGAGCCACTCCGGTGGACGTTGCCACGTGCATTCCCACTGTTGAGTGCT
 CCCCCCGGGGCTGCTGGCTTCTGCAATCTTCCAATCGCACGAATGACCAGTGGAACCTGAACCTTCAACAAAGA
 ATGGTACGGTGCAGCTGAGCTTTACAGAGCCGGCCAGCTCACCCTGCAGTCTTGAAGGCTCTACAAGTTTATGAACG
 GGGTGGCGTGTACCCCATTTGTTGGACCTGTCCCTGGAGTGGCCGTTTTCGCCAATTCCTACATGTGAGTGATAAAC
 CTTCCCGGAGCAACTCAGTGTGACCAACCTGCCGCTCCCGCAGAGACCAAGCCTGAAGACTTTGCCCTTTGAG
 TGTGCTATGGCTACTGTCTATGACATTGGTCATGACGCCGTCTGTATGTGGCCGAAAGGAAAGTCTCTGGGCCCTCG
 TGGCGGGGATGAAGTGAATTTGAAGCTGTCCCGGGGAGTTGAAGTTGATTGCGAACCGGCTCCGCACCTCCTCCCGC
 CCCACACACAGTGGACATGTCTAAGTTCGCTTACAGCCCTGGGTGTGGTGTCTATGCGGGTGAACGCCAACAC
 GGCTGCCTTCCCGTGACACTGTCCCTGAAGGCAACTGTCTGGTGAGCTTGTGACTTGTCTTCACTGGAAAGTTGAGAA
 CAAAGAAATTCGCATGCTAACCAATTTGGTACCAGACCAAGCATGGTGTCTCTGGCAAGTACCTGCAGCGGAGGTGC
 AAGTTAATGGTCTCCGAGCAGTAAGTGAACCTAAACGAGCTATCGTGTACAGTACTTCTCCGTTAAGGAGAGTTGGATC
 CGCCATTTGAAACTGGCGGGAGAACCCAGCTACTCTGGGTTTGAAGGACCTCTCAGAATAAGGGTTGAGCCTAACACGTC
 GCCATTGGCTGACAAGGAAGAAAAATTTCCGGTTTGGCAGTCAAGTGGTACGGCGCTGGAAAGAGAGCAAGAAAAG
 CACGCTCTTGTGCACTGCTACAGTGCCTGGCCGCGCTTGTCCGTTCTGTAAACCCGGCAGGCCAAGGAGCACGAGGT
 GCCGGCGCAACAAAGGCTGAGCACCTCAACACTACTCCCGCCTGCCGAAGGGAATTGTGGTTGGCACTGCATTTCCGC
 CATCGCAACCGGATGGTGAATTCAAAATTTGAAACACCCCTTCCCGAAGAGTGAGACCTCCAGATGACTGGGCTACTG
 ACGAGGATCTTGTGAATGCCATCCAAATCCTCAGACTCCCTGGCGCTTAGACAGGAACGGTGTGTACTAGCGCCAAG
 TACGTACTTAAGCTGGAAGGTGAGCATTGGACTGTCACTGTGACCCCTGGGATGTCCCTTCTTGTCTCCCTCTTGAATG
 TGTTCAGGGCTGTGTGGGCACAAGGGCGGTCTTGGTCCCAGATGCAGTCGAGGTCTCCGGATTTGACCTGCCTGCC
 TTGACCGGCTGGCTGAGGTGATGCACCTGCCTAGCAGTGTATCCAGCCGCTCTGGCCGAATGTCTGGCGATTCCGAT
 CGTTCGGCTTCTCCGGTCAACCCGTGTGGACTGTTTCGAGTCTTTGCCCGTCAAGCGGAGGGAATCACCTGACCA
 AGTGCCTTAGGGAAATATCAGCTTTGTGAGGTGATTGAGGACTGCTGCTGTCCAGAACAAACCAACCGGGTCA
 CCCCCGAGGAGTGCAGCAAGATTGACCTGTACCTCCGTGGTGCAACAAATCTTGAAGAATGCTTGGCAGGCTTGAG
 AAAGCGCGCCCGCACGCGTAATCGACACCTCCTTTGATTGGGATGTTGTCTCCCTGGGGTTGAGCGGCAACCCAGAC
 GATCAAGCTGCCCAAGTCAACAGTGTCTGCTGCTGCTGCTGTTGTGACTCAAAAGTCTTGGACAACAACCTCGGTCC
 CCCTGACCGCTTTTCACTGGCTAACTACTACTACCGTGCAGCAAGGTGACGAAGTTCGTACCGTGAAAGACTAACCGCC
 GTGCTCTCAAGTTGGAAGAGTTGTTGAGAAGAATATGGGCTCATGCCAACCGAGCCTGGTCCACGGCCCACTGCC
 ACGCGGGCTCGACGAATCAAGACCAAGTGGAGGAGGACTTGTGAAACTGGCTAACGCCAGACGACTTCGGACATGA
 TGGCTGGGGCAGTCGAGCAGGTTGACCTAAAACTTGGGTCAAGAACTACCCGCGGTGGACACCAACCCCTCCGCCA
 AAAGTTCAAGCTCGAAAAACGAAGCCTGTCAAGAGCTTGGCGGAGAGAAAGCCTGTCCCGCCCGCGCAGGAAGGTTGG
 GTCCGATTGTGGCAGCCCGTTTATTAGCGGCGATGTCCCTAACAGTTGGGAAGATTGGCTGTTAGTAGCCCTTTG
 ATCTCCCGACCCACCTGAGCGGCAACACCTTCAAGTGAGCTGGTATTGTGTCTCACCAGCAATGCATCTTCAGGCCG
 GCGACACCTTGAGTGAGCGGCTCCAATTCGCCACCTCGCGGAACGTGTCTCGACCGGTGACACCTTGAGTGAGCC
 GATCCCTGTGCCCGACCGCGGCTAAGTTTCAAGAGGTGAAAAGATTGAGTTGGCGGCGGCAATCCACCGTACCAGG
 ACGAGCCCTGGATTGTCTGCTTCTCAGAGTGAATATGAGGCTCTCCCCAGCACCGCGCAGAGCGGGGCGTT
 CTGGGAGTAGAGGGGATGAAGCTGAGGAAACCTGAGTGAATCTCGGACATGTGGGTAAACATTAACCTGCGTCCGT
 GTCATCAAGCAGCTCCTGTCCAGCGTGAGAATCACAGCCCAAAATACTCAGCTCAAGCCATCATCGACTCGGGCGGGC
 CCTGCAGTGGGATCTCAAGAGGTAAAGGAAACATGCTTAGTGTATGCGCGAGGATGTGATGCGACTAAGCTTGAT
 GACCTGCTACGCAAGGAATGGCTTCTCGCATGTGGGATCGGGTGGACATGCTGACTTGGCGCAACAGCTGTGTTACCA
 GCGGATTGCACCTTAGATGGCAGGTTAAAGTTCCTCCAAAAATGATACTCGAGACACCGCCGCTTATCCGTGTGAGT
 TTGTGATGATGCCTCACAGCCTGCACCTTCCGTAGGTGGGAGAGGACCTTACCATTGGCTCAGTTGCTACTGAAGAT
 GTTCCAGCATCTCGAGAAATAGAAAATGTGGCGAGATGGCCAACCAAGGACCTTGGCCTTCTCCGAGGATAAACCC
 GGATAGTACCAACTTGTCAACGACCCCGGATATCGTGGCGAGGCTGACGAGACACATCAGCTCCGTCCGAGGCA

Fig. 1D-2

CAGGTGGCGCCGGCTCTTTTACCGATTTGCCGCCTTCAGATGGCGCGGATGCGGACGGGGGGGGCCGTTTCGGACGGTA
 AAAAGAAAAGCTGAAAGGCTCTTTGACCAACTGAGCCGTAGGTTTTTACCTCGTCTCCCATCTCCCTGTTTTCTTCTC
 ACGCCTTTTCTACCTGGCGGTGGTTATTCTCCGGGTGATTGGGGTTTTGCAGCTTTTACTCTATTGTGCCTCTTTTTAT
 GTTACAGTTACCCAGCCTTTGGTATTGCTCCCTCTTGGGTGTGTTTTCTGGGTCTTCTCGGCGCGTTTCAATGGGGGT
 TTTGGCTGCTGGTTGGCTTTGCTGTTGGTCTGTTCAAGCCTGTGTCCGACCCAGTCGGCGCTGCTTGTGAGTTTGACTC
 GCCAGAGTGTAGAAACATCCTTCATTCTTTGAGCTTCTCAAACCTTGGGACCTGTTGCGAGCCTTGTGTGGGCCCG
 TCGGTCTCGGTCTTGCCATTCTTGGCAGGTTACTGGGCGGGGACGCTGCATCTGGCACTTTTGTAGGCTTGGCATT
 GTTGCGAGCTGTATCTTGGCTGGAGCTTACGTGCTTTCTCAAGGTAGGTGTAAAAAGTGTGGGGATCTTGATAAAGAAC
 TGCTCCCAATGAGGTGCGTTTTAACGTGTTTCTTTACACGTGCGACCAAGTCGTCACTATCGACCTGTGCGATCGGT
 TTTGTGCGCCAAAAGGAATGGACCCATTTTCTCGCCACTGGGTGGCGCGGGTGTGGCCGGCCGAAGCCCCATTGAG
 CAACCTCTGAAAAACCATCGCTTTGCCAGTTGGATGAAAAGAAGATTACGGCTAGGACTGTGGTCGCCAGCCTTA
 TGACCCCAACCAAGCCGTAAAGTGCTTGCGGGTATTGCAGCGGGTGGGGCGATGGTGGCTAAGGCGGTCCAAAAGTGG
 TCAAGGTTCCGCTGTTCCATTCCGAGCCCCCTTCTTCCACTGGAGTGAAAGTTGACCTGATTGCAAGGTCGTGGT
 GACCTGACACTTTCACTGCAGCTCTCGGTCTGGCTACTCCACCACAAACCTCGTCCTTGGTGTGGGGGACTTTGCCCA
 GCTGAATGGATTAAAAATCAGGCAAATTTCAAGCCTTCAAGGGGAGGCCACATCTCATGGCTGCCCTGCATGTTGCCT
 GCTCGATGGCTCTGCACATGCTTGTGGGATTTATGTGACTGCGGTGGGTCTTGCGGCACCGGCACCAACGACCCGTGG
 TGCGCTAACCCGTTTGGCTCCCTGGCTACGGACCTGGCTCTCTGCACGTCCAGATTGTGATTTCCCAACACGGCT
 TACCCTGCCCTTGACAGCACTTGTGGCGGGATTGGGTATTCAAGAAATTGCTTGGTGTGTTTTGATTTTGTTCATCG
 GAGGCATGGCTCATAGGTTGAGCTGTAAGGCTGACATGCTGTGTCTTGTGCAATTGCCAGCTATGTTGGGTACCT
 CTTACCTGGTGTGTTGTGTGTTTCTTGTGTTGCGCTGTTTTCTTGCACCCCTCACCATCCTATGGTGGTGT
 TTTCTTGATTTCTGTGAATATGCCTCAGGAATCTTGGCCATGGTGTGTTGGTTCTCTTGGCTTCTTGGTCGTTATA
 CTAATGTTGCTGGCCTTGTACCCCCCTACGACATTCATCATTACACCAAGTGGCCCCGCGGTGTTGCCGCTTGGCTACC
 GCACCAAGATGGGACCTACTTGGCCGCTGTCCGCCGCGCTGCGTTGACTGGCCGACCATGCTGTTTACCCCGTCCAGCT
 TGGGTCTCTTCTGAGGGTGCTTTCAGAACTCGAAAGCCCTCACTGAACACCGTCAATGTGATCGGGTCTCCATGGGCT
 CTGGCGGGGTGTTTACCATCGACGGGAAAGTCAAGTGCGTAACTGCCGCACATGTCCTTACGGGCAATTGAGCTCGGGT
 TCCGGGGTGGCTTCAATCAAATGCTTGACTTTGACGTAAAGGGAGATTTGCTATAGCTGATTGCCGAATTGGCAAGG
 GGCTGCCCCCAAGACCCAATTCTGCACGGATGGATGGACTGGCGTGCCTATTGGCTAACATCCTCTGGCGTCGAACCG
 GCGTCATTGGAAAAGGATTCGCTTCTGCTTACCAGCATGTGGCGATTCCGGGTCCCAAGTGATCACCAGGCGGGTGAG
 CTTGTGCGGCTTCAACGGGATCGAATAAACAAGGGGGGGGATTGTTACGCGCCCTCAGGCCAGTTTGTAAATGTGGC
 ACCATCAAGCTAAGCGAATTAAGTGAATCTTTGCTGGGCTAAGTCCCGCTCGGTGATGTGAAGTGGCAGCCACA
 TAATTAAGACATAAGCGAGGTGCCTTCAGATCTTTGTGCTTGTGCTGCCAAACCTGAACTGGAAGGAGGCTCTCC
 ACCGTCCAATCTTTGTGTGTTTTCTCTGTGGAGAATGATGGGACATGCTGGACGCCCTTGGTGTGAGTTT
 CTTTATTTTGAATGAGGTTCTCCAGCGCTCTGGTCCGGAGTGTTTTCTCTTTGGAATGTTTGTGCTATCCTGGCTCA
 CGCCATGGTCTGCGCAAGTTCTGATGATCAGGCTTCTGACAGCAGCTCTTAACAGGAACAGATGGTCACTTGCCTTTTT
 AGCCTCGGTGCAGTGACCGGTTTTGTGCGAGATCTTGGGCCACTCAGGGGCATCCGTTGACGGCAGTGATGAATTTGAG
 CACCTATGCATTCCTGCCTCGGATGATGGTTGTGACCTCACCAGTCCAGTGATCACGTGTGGTGTGCTGCACCTACTTG
 CCATCATTTTGTACTTGTAAAGTACCGTGGCCTGCACCATATCCTTGTGGCGATGGAGTGTCTCTGCGGCTTTCTC
 TTGAGATACTTTGCCGAGGAAAGTTGAGGGAAGGGGTGTGCAATCCTGCGGAATGAATCATGAGTCTCTGACTGGTGC
 CCTCGCTATGAGACTCAATGACGAGGACTTGGATTTCTTATGAAATGGAATGATTTTAAAGTCTTTGTTCTGCGTCCA
 ACATGAGGAATGACGCGGGTCAATTTATCGAGGCTGCCTATGCTAAAGCACTTAGAGTAGAACTGGCCAGTTGGTGCAG
 GTTGATAAAGTTGAGGTACTTTGGCCAACTTGAAGCTTTGCTGATACCGTGGCACCTCAACTCTCGCCCGGTGACAT
 TGTTGTGCTCTCGGCCACACGCTGTTGGCAGTATCTTGCACCTAAAGGTTGGTAGCACAAGCATACCCTCCAAGCCA
 TTGAGACCAGAGTCTTGTGGGTCCAAAATGACCGTGGCGCGCTGCTGACCCGACCCCAACGCCCCACCCGACCC
 GTGCCATCCCCCTCCACCGAAAGTTCTGGAGAATGGCCCAACGCTTGGGGGATGAGGACCGTTGAATAAGAGAA
 GAGGCGCAGGATGGAAGCCCTCGGCATCTATGTTATGGGCGGGAATAACAGAAATTTGGGACAAGAATTCGGGTG
 ATGTGTTTTATGAGGAGGTCCATAATAACACAGATGAGTGGGAGTGTCTCAGAGTTGGCGACCCTGCCGACTTTGACCT
 GAGAAGGAACTCTGTGTGGACATGTACCATTTGGAACAAGGCTTACCATGTTTACACCTCCCATCTGGTAAGAAGTT
 CTTGGTCCCCGTCAACCCAGAGAATGGAAGAGTCCAATGGGAAGCTGCAAGCTTTCCGTGGAGCAGGCCCTAGGTATGA

Fig. 1D-3

TGAATGTCGACGGCGAACTGACTGCCAAAGAACTGGAGAACTGAAAAGAATAATTGACAACTCCAGGGCCTGACTAAG
 GAGCAGTGTTTAACTGCTAGCCGCCAGCGACTTGACCCGCTGTGGTCGCGCGGCTTGGTTGTTACTGAAACAGCGGTA
 AAAATAGTCAAATTTACAACCGGACCTTCACCCTGGGACCTGTGAATTTAAAAGTGGCCAGTGAGGTTGAGCTAAAAGA
 CGCGGTTGAGCACAACCAACACCCGTTGCGAGACCGATCGATGGTGGAGTTGTGCTCCTGCGTTCCGCGGTTCTTCGC
 TTATAGACGCTTGATCTCCGGTGCTGATGCATCTCCCAAGTTACTTGCCCATCACGGGCCGGGAAACACTGGGATCGAT
 GGCACGCTCTGGGATTTTGTAGTCCGAAGCCACTAAAGAGGAAGTGCAGCTCAGTGCGCAAATAATACAGGCTTGACAT
 TAGGCGCGGCGACGCTCTGAAATTTGGTCTCCCTTACAAGCTGTACCCTGTAGGGGTAACCTGAGCGGGTGAAAGGAG
 TTCTGCAGAATAACAAGTTTGGAGACATACCTTACAAAACCCCCAGTGACACTGGAAGCCCAAGTGACGCGGCTGCCTGC
 CTTACGCCCCAACGCCACTCCGGTGACTGATGGGCGCTCCGTCTTGCCACGACCATGCCCCCGGGTTTGAGTTATATGT
 ACCGACCATAACAGCGTCTGTCTTGATTACCTTGACTCTAGGCCTGACTGCCCTAAACAGCTGACAGAGCACGGCTGCG
 AAGATGCCGCACTGAAAGACCTCTCTAAATATGACTTGTCACCAAGGCTTTGTTTTACCTGGAGTTCTTCGCTTGTG
 CGGAAATACCTGTTTGCCCATGTAGGTAAAGTGCCACCCGTTTATCGGCCTTCTACTTACCCTGCTAAGAATTCTATGGC
 TGGAAATAAATGGGAACAGGTTCCCAACCAAGGACATTAGAGCGTCCCTGAAATCGACGTTCTGTGCGCACAGGCTGTGC
 GAGAAAACCTGGCAAACCTGTACCCCTTGACTCTTAAGAAACAGTATTGCGGGAAGAAGAAGACTAGGACCATACTCGGC
 ACCAATAAATTCATCGCACTAGCCACCGAGCAGTGTTGAGTGGTGTACCCAGGGCTTCATGAAAAAGCGGTTTAACTC
 GCCCATCGCCCTCGGAAAGAACAAGTTTAAAGAGCTACAGACTCCGGTCTGGGACAGTGCTTGAAGCTGATCTCGCAT
 CCTGCGATCGATCCACGCTGCAATTGTCCGCTGGTTTGGCGCAACCTTCTTTATGAACCTGCTGTGCTGAAGAGCAT
 CTACCGTGTACGTGCTGAAGTCTGCCACGACTTACTGGTCACGCACTCGGCGCAGTGACTAAGAGAGGTGGCCTGTC
 GTCTGGCGACCCGATCACCTCTGTGTCTAACACCATTTATAGTTTGGTGTCTATGCACAGCATATGGTGCTTAGTTACT
 TCAAAGTGGTCACCCCATGGCCTTCTGTTCTTACAAGACCAGCTAAAGTTTGGAGCATGCTCAAGGTTCAACCCCTG
 ATCGTCTATTGCGACGACCTCGTGCTGTATGCCGAGTCTCCACCATGCCAACTATCACTGGTGGGTTGAACATCTGAA
 TTTGATGCTGGGGTTTCAAGCGACCCAAAGAAGACGAATAACAGACTCGCCATCATTTCTAGGCTGTAGAATAATAA
 ATGGGCGCCAGCTAGTCCCAACCGTGACAGGATCCTCGCGGCCCTCGCTATCACATGAAGCGAGTAATGTTTCTGAA
 TACTATGCCTCAGCGGCTGCAATACTCATGGACAGCTGTGCTTGTGGAGTATGATCCTGAATGGTTTGAAGAAGTGT
 AGTTGGAATAGCGAGTGCGCCCGCAAGGACGGCTACAGCTTCCCGGCACGCCGTTCTTCATGTCCATGTGGGAAAAAC
 TCAGGTCCAATTATGAGGGGAAGAAGTCGAGAGTGTGCGGGTACTGCGGGGCCCCGGCCCCGTACGCTACTGCCTGTGGC
 CTCGACGTCTGCATTTACCACACCCACTTCCACCAGCATTGTCCAGTCACAATCTGGTGTGGCCATCCAGCGGGTCTGG
 TTCTGTAGTGAGTGCAAATCCCCTGTAGGGAAAGGCACAAGCCCTTAGACGAGGTGCTGGAAACAAGTCCCGTATAAGC
 CCCACGGACCGTTATCATGCATGTGGAGCAGGCTCACCCCTTGATCCAGGTAGATACCAAACTCGCCGCGGATTA
 GTCTCTGTGACGGCGTGAATTAGGGGAAATGAAGTTGGACTACACAGACGGTGATTATGCTAGCACCGCTTGTCTCTAC
 CTGCAAAGAGATCAACATGGTGTGCTGTGCTTCCAATGTATTGCGCAGCAGGTTATCATCGGCCACCCGGTGTGGGA
 AAACATACTGGCTCCTTCAACAGGTCCAGGATGGTGTATTATACACCAACTCACCAGACCATGCTTGACATGATT
 AGGGCTTTGGGGACGTGCCGTTCAACGTCCCGGCAGGCACAACGCTGCAATTCCTCGTCCCTCCCGCACCGGTCCGTG
 GGTTGCGATCCTAGCCGGCGGTTGGTGTCTGGCAAGAAATCCTTCTAGATGAAGCAGCGTATTGCAATCACCTTGATG
 TTTTGAGGCTTCTTAGTAAACTACCTCACCTGTCTAGGAGACTTCAAGCAACTCCACCCAGTGGGTTTGATTCTCAT
 TGCTATGTTTTGACATCATGCCTCAAACCTCAACTGAAGACCATCTGGAGGTTGGACAGAATATCTGTGATGCCATTCA
 GCCAGATTACAGGGACAACTCATGTCCATGGTCAACACAACCCGTGTGACCTACGTGGAAAAACCTGTAGGTATGGGC
 AGGTCCTCACCCCTACCACAGGGACCGAGAGGACGACGCCATCACTATTGACTCCAGTCAAGGCGCCACATTGATGTG
 GTTACATTGCAATTTGCCCACTAAAGATTCACTCAACAGGCAAGAGCCCTTGTGCTATCACCAGGGCAAGACACGCTAT
 CTTTGTGTATGACCCACACAGGAGCTGCAAGGCTTGTGATCTTCTGCAAAAGGCACGCCGTCAACCTCGCAGTGC
 ACTGCGACGGGCGAGTGTGCTGGATAGAAATAACAAAGAATGCACGGTTGCTCAGGCTTAGGCAACGGGGATAAA
 TTTAGGGCCACAGACAAGCGTGTGTAGATTCTCTCCGCGCCATTTGTGCTGATCTAGAAGGGTCGAGCTCTCCGCTCCC
 CAAGGTGCGACACAACCTTGGGATTTTATTTCTCACCTGATTTAACACAGTTTGCTAAACTCCAGTAGAAGTTGCACCTC
 ACTGGCCCGTGGTGTCAACCCAGAACAAATGAAAAGTGGCCGGATCGGCTGGTGCCAGCCTTCGCCCTATCCATAAATAC
 AGCCGCGCTGCATCGGTGCCGCTATATGGTGGGCCCTCGGTGTTTCTAGGCACTCTGGGGTGTGTCTACTATCT
 CACAAAATTTGTTAAGGGCGGGCTCAAGTGCTTCCGGAGACGGTTTTCAGCACCGGCCGAATTGAGGTAGACTGCCGGG
 AATATCTTGATGATCGGGAGCGAGAAGTTGCTGCGTCCCTCCACACGCTTTTATTGGCGACGTCAAAGGCACTACCGTT
 GGAGGATGTATCATGTACCTCCAGATACCTCCCGCGCTCCTTCCAAGGAATCAGTTGCGGTAGTCGGGGTTTCAAG

Fig. 1D-4

CCCCAGAAAGCCGCGAAAGCATTGTGCACACTGACAGATGTGTACCTCCAGATCTTGAAGCCTATCTCCACCCGGAGA
 CCCAGTCCAAGTGCTGGAAAATGATGTTGGACTTCAAAGAAGTTCGACTAATGGTCTGGAAAGACAAAACAGCCTATTTT
 CAACTTGAAGGTCGCTATTTACCTGGTATCAGCTTGCCAGCTATGCCTCGTACATCCGTGTTCCCGTCAACTCTACGGT
 GTACTTGGACCCCTGCATGGGCCCCGCCCTTGAACAGGAGAGTCTGCGGGTCCACCCACTGGGGGGTGACCTCGCGG
 TCACCCCTTATGATTACGGCGCTAAAATTATCTGTCTAGCGCGTACCATGGTGAAATGCCCCCGGATACAAAATTCTG
 GCGTGCGCGGAGTTCTCGTTGGATGACCCAGTTAAGTACAAACATACCTGGGGGTTTGAATCGGATACAGCGTATCTGTA
 TGAGTTACCCGGAACGGTGAGGACTGGGAGGATTACAATGATGCGTTTCTGCGCGCCAGGAAGGGAATTTATAAGG
 CCACTGCCACCAGCTTGAAGTTTTATTTTCCCCGGGCCCTGTATTGAACCAACTTTAGGCCTGAATTGAAATGAAATG
 GGGTCCATGCAAAGCCTTTTTGACAAAATTGGCCAACTTTTTGTGGATGCTTTCACGGAGTTCTTGGTGTCCATTGTTGA
 TATCATTATATTTTTGGCCATTTTGTGGCTTACCATCGCCGGTGGCTGGTGGTCTTTTGCATCAGATTGGTTTGTCT
 CCGCGATACTCCGACGCGCCCTGCCATTCACTCTGAGCAATTACAGAAGATCTTATGAGGCCTTTCTTCCAGTGCCA
 AGTGGACATTTCCACCTGGGGAACATAAATCCTTTGGGGATGCTTGGCACCATAAGGTGTCAACCTGATTGATGAAA
 TGGTGTGCGGTGCAATGTACCGCATCATGAAAAAGCAGGGCAGGCTGCCTGGAAACAGGTGGTGAGCGAGGCTACGCTG
 TCTCGCATTAGTAGTTTGGATGTGGTGGCTATTTTCAAGCATCTAGCCGCCATTGAAGCCGAGACCTGTAATATTTGGC
 CTCCCGGCTGCCATGCTACACAACCTGCGCATGACAGGGTCAAATGTAACCATAGTGATAATAGCACTTTGAATCAGG
 TGTGTCTATTTTTTCAACCCCTGGTCCCGGCCAAAGCTTCATGATTTTCAAGCAATGGTTAATAGCTGTACATTCCTCC
 ATATTTTCTCTGTGTCAGCTTCTTGTACTCTTTTGTGTGCTGTGGTTCGCGGTTCGAATACTACGTAAGTGTCTTGG
 TTTCCGCTGGTTAGGGGAATTTTTCTTTCGAACCTCACAGTGAATFACACGGTGTGTCCACCTTGCCTCACCCGGCAAGC
 AGCCACAGAGATCTACGAACCCGGTAGGTCTCTTGGTGACGGATAGGGTATGACCGATGTGGGGAGGACGATCATGACG
 AGCTAGGGTTTATGATACCGCTGGCTCTCCAGCGAAGGCCACTTGACTGGTGTTCACGCTGGTTGGCGTTCTTGTCC
 TTCAGCTACACGGCCAGTTCATCCCGAGATATTCGGGATAGGGAATGTGAGTCGAGTTTATGTTGACATCAAAATCA
 ACTCATCTGCGCGAATGACGGGCGAAGACCACTTGCCTCGTCATGACAACATTTAGCCGTGTTTCAGACCTATT
 ACCAATCATCAAGTCGACGGCGCAATTGGTTTCACTAGAATGGCTTCGTCCTTCTTTCTCGTGGTGGTTTAAAT
 GTCTCTTGGTTTCTCAGGCGTTGCGCTGCAAAACCATGTTTCAGTTCGAGTCTTGAGATATTAAGACCAACACCAACCGCA
 GCGGCAAGCTTTGCTGCTCCTCAAGACATCAGTTGCCTTAGGCATCGCGACTCGGCCTCTGAGGCGATTTCGCAAAATCCC
 TCAGTGCCGTACGGCGATAGGGACACCCGTGTATGTTACCATCACAGCCAATGTGACAGATGAGAATTATTTACATTCTT
 CTGATCTCCTCATGCTTTCTTCTGCTTTCTATGCTTCTGAGATGAGTGAAAAGGGATTTAAGGTGGTATTTGGCAAT
 GTGTGAGGCATCGTGGCTGTGTGTCAATTTTACCAGCTACGTCCAACATGTCAAGGAGTTTACCAACGCTCCCTGGT
 GGTGCAACATGTGCGGTTGCTCCATTTATGACACCTGAGACCATGAGGTGGGCACTGTTTAGCCTGTCTTTTGCCA
 TTCTGTTGGCAATTTGAATGTTAAGTATGTTGGAGAAATGCTTGACCGCGGGCTGTTGCTCGCGATTGCTTTCTTTGTG
 GTGTATCGTGCCGTTCTGTTTGTGTGCTCGCAACGCCAGCAACGACAGCAGCTCCCATCTACAGCTGATTTACAAC
 TGACGCTATGTGAGCTGAATGGCACAGATTGGCTAGCTAACAAATTTGATTGGGCAGTGGAGAGTTTGTCTCTTTCCC
 GTTTTGACTCACATTGTCTCCTATGGTGCCCTCACTACCAGCCATTTCTTGACACAGTCGCTTTAGTCACTGTGTCTAC
 CGCCGGGTTTGTTCACGGGCGGTATGTCTAAGTAGCATCTACCGGCTGTGTGCCCTGGCTGCGTTGACTTGCTTGTCA
 TTAGGTTTGCAAAGAAATGATGTCTGGCGCTACGCGTGTACCAGATATACCAACTTTCTTCTGGACACTAAGGGCAGA
 CTCTATCGTTGGCGGTGCGCTGTATCATAGAGAAAAGGGGCAAGTTGAGGTGGAAGGTCTATGATCGACCTCAAAAG
 AGTTGTGCTTGATGGCTCCGTGGCAACCCCTATAACCAGAGTTTACGCGGAACAATGGGGTCTGCTTAGATGACTTCTG
 TCACGATAGCACGGCTCCACAAAAGGTGCTTTTGGCGTTTTCTATTACCTACACGCCAGTGATGATATATGCCCTAAAGG
 TGAGTCGCGGCCGACTGCTAGGGCTTTCGACCTTTTGATCTTCTGAATTTGTGCTTTCACCTTCGGGTACATGACTTTC
 GCGCACTTTTCAAGTACAAATAAGGTGCGGCTCACTATGGGAGCAGTAGTTGCACTCCTTTGGGGGGTGTACTCAGCCAT
 AGAAACCTGGAAATTCATCACCTCCAGATGCCGTTTGTGCTTGTAGGCCGCAAGTACATTCTGGCCCTGCCACCACG
 TTGAAAGTGCCGAGGCTTTTATCCGATTGCGGCAATGATAACCACGCATTTGTCTGCGGCGTCCCGGCTCCACTACG
 GTCAACGGCACATTGGTGCCCGGTTAAAAAGCCTCGTGTGGTGGCAGAAAAGCTGTTAAACAGGGAGTGGTAAACCT
 TGTCAAATATGCCAAATAACAACGGCAAGCAGCAGAGAGAAAGAAAGGGGATGGCCAGCCAGTCAATCAGCTGTGCCAG
 ATGCTGGGTAAAGATCATCGCTCAGCAAAACAGTCCAGAGGCAAGGGACCGGGAAGAAAAATAAGAAAGAAAAACCCGGA
 GAAGCCCCATTTTCTCTAGGCACTGAAGATGATGTGACAGATCACTTTACCCCTAGTGAGCGGCAATTGTGTCTGTCTG
 CAATCCAGACCGCCTTAATCAAGGCGCTGGGACTTGACCCCTGTGAGTTACAGGGAGGATAAGTTACACTGTGGAGTTT
 AGTTTGCTACGCATCATACTGTGCGCTGATCCGCTCACAGCATCACCTCAGCATGATGGGCTGGCATTCTTGAGG

Fig. 1D-5

ATCTCAGTGTGTTGAATTGGAAGAATGTGTGGTGAATGGCACTGATTGACATTGTGCCTCTAAGTCACCTATTCAATTAGG
GCGACCGTGTGGGGGTGAGATTTAATTGGCGAGAACCATGCGGCCGAAATTAACAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAA

Fig. 1E-1

>MN184A_DQ176019.seq

ATGACGTATAGGTGTTGGCTCTATGCCACGACATTTGTATTGTACAGGAGCTGTGACCACTGGCAGACGCCAAAGCTTGCT
 GCACAGAAACACCTTCTGTGACGGCTCTTTCAGGGGAGTTTAGGGTTTATCCCTAGCACCTTGTCTGGAGTTGCA
 CTGCTTTACGGTCTCTCCACCCCTTAACCATGTCTGGGATTCTTGATCGGTGCACGTGCACCCCAATGCCAGGGTGT
 TATGGCAGAGGGCCAAAGTCTACTGCACACGATGTCTAGTGCACGGTCCCTCCTTCCCTGAATCTCAAGTCCCTGAGC
 TCGGAGTGTGGGCTTGTATATAGGCCGAAGAGCCGCTCCGGTGGACGTTGCCACGCGATTCCCACTGTTGAGTGC
 TCCCTGCTGGGGCTTGTGGCTTCTGCAATTTTCCAATTGCACGAATGACCACTGGAAACCTGAACCTTCAACAAAG
 ATTAGTGGGGTGCAGCTGAGCTTTACAAAGCCGGCTGCCTCACCCCTATAGTCTAAAGAATCTACAAGTCTATGAAC
 GGGTTGCCGATGGTACCCATCGTTGGACCTGTCCCTGGAGTTGCCGTTTTCGCCAACTCCCTACATGTGAGTGATAGA
 CCTTCCAGGGGTAACCTGACGTGCTAACCAACCTGCCGCTCCCGCAGAGACCTAAGCCTGAAGATTTTGGCCCTTTGA
 GTGTGCTATGGCTGMCCTCTATGAYATTGGTCATGACGCCGTTATGTTCTGGCCGAAGGGAGAGTCTCTGGGCTCCGC
 GTGGTGGGGGAAAAGGAAATTTGAACTGTTCCCGAGGAGTTGAGGTTGATTGCAGAGCACTTTATACCTCCTTCCCG
 CCCACACGTTGGTGGACATGTGAAATTCACCTTACCGCCCTGAGTGTGGTGCTTCCATGCGAGTGCAGCGCATTA
 TGGCTGCTCCCGCCGGCACTGTCCCTGACGGCAATTGCTGGTGGAGTTTGTAGCTCGCTCCCACTGGAAATCCAGT
 ACAAGAAATTCGCCACGCCACCAATTTGGCTATCAAACTAAGCATGGCGTTGCTGGCAAGTACCTACAGCGGAGGCTG
 CAAGTTAATGGTCTCCGAGCATGTTGACTCGAATGGACTATCGTCATACAGTACTTCTCTGTTAAGGAGAGCTGGA
 CCGCCACGTGAACTGGCGGAAGAGTTGACTACCTGGGTTTGAAGTCTCTCAGGATAAGAGTCGAGCCCAACACGT
 TGCCATTGTCCAACAAGGACGAGAAAATCTTCCGGTTTGGTGGTGCAAGTGGTACGGTGCTGGGAAGAGGGCAAGGAGG
 GCACGTGCAAGTGCAGTACCGCAGTCGCCGGTCACTCGCTCCGCTACTCGTGAACCCAGCAAGCCAAAGAACACGAGGC
 TGCTAGTGCCAAACAAGGCTGAGCTTCTTGAACGCTACTCCCGCCTGCTGAAGGGAATTGCGGCTGGCACTGTATTTCCG
 CCATCGCAATCGGATGGTAAATCTAAGTTTGAAGTCTGCCCTTCCGAAAGAGTGAGATCCCAAGAGCTGGGCTACT
 GATGAGGATCTTGTGAATACTATCCARATCCTCAGGCTCCCGGCGCTTAGACAGGAACGGCGCTGTGCAAGCGCCAA
 GTACATCTTAAAGCTGGAAGGTGAGCACTGGACTGTTTCACTGATTCCCGGAATGYCCCTTCTTGTCTCCCTTGAAT
 GCGTTTCAAGGTTGCTGTGAGCATAGGGTAATCTTGGTTCTCCGAACGCGGTGGGGTTTTTGGATTGACCTGCCAGC
 CTTGACCGACTTGTGGGGTGTGCACTGCCAGCAGTCCATCCAGCGCTTGGCCGAGTTGTCTGGCGACCTTGA
 TCGTCAAATTTCCCGGCGCCACTGTGTGGACTGTCTCGAGTTTATGCTCGTCATAGTGGAGGRGAGCATCTGATC
 AAAAGTGTATAAAAAAATATCAGTCTCTGTGAGGTGATCGAGAGTTGTTGCTGTTCTCGAAACAAACTAACCGGGTC
 ACCCGGAAGAGGTGACAGCAAGATTGATCTGTACCTTTTTGGTGCAGCAAGTCTTGAAGAAATGCTTGGCCAGGCTTGA
 RAAAGCTCGCCGCCAAGCGTATTARACACCTCCTTGTATTGGGATGTTGTCTCCCTGGTGTGGGRCGGCTGCTCAAG
 CAGCAAACTGCCCTCACCACACGAGCTCACGCTCTAGCCTGTTGTGACYCAAAGGTCTTGGCGAAATTTCAACCT
 CGAAAAGCGAGTCTGTCAAGAGCCTACCAGAGAGCAGGCCACTCCCTGCCCGCGCAAAAAGATTAGGTCAGGTGTGG
 TAGTCCGATTTTATTGGGCGCAATCTCCCTGACAGCCAGGAAGACTTGGCCGGTGGTTCTTTGATTTCCCAACCTAC
 CTGAGTTGGTGGTAAGCTCGAGTGAAGTGTGCTGTCCCTGCACCGCGCAGGGTTGTGTCCCGATTAGTGTCTCCG
 ATAGTGTGACCCCTGTGCCCGCACCGCTGGGCTTCCGCGAGTGGAGGGAATGAATTTGGCGGCACTGACTCTAGC
 GTGCCAGGACGAGCCCTCGATTTGTCTGCGTCTCGCACTGAATATGAGGCGTCCCTTGGCATTGCCGTGAGTG
 AGGATGTCTGGCGGTGGAGAGACGAGAAGTTGAAGAAGTCTGAGCGGAATATCGGGCATGTGAGTGAATCAGGTTG
 GCGCCGTGTATCAAGTAGCTCCCTGTCAAGCATAGAGATCACACGTCCAAGTACTCAGCTCAAGCCATCATTAATC
 AGGTGGGCTTGTGTGGGCACTCCAGGAGGTGAAAGAGAAATACCTTAATGTGATGCGTGAGGCATGTGATGCGACCA
 AGCTTGATGACCTGCCAGCAAGAAATGGCTTCCCGTATGTGGGATAGGGTAGACATGCTAACCTGGCGCAACAGCTCC
 ATTTTTCAGGCGCTTTCACCTTGGTGACAAGTTTAAAGTCCCTCCCGAAGATGATACTCGAAACACCGCGCCCTACCC
 TTGTGGGTTTGTGATGATGCCCCGACGCTGCACCTTCTGTAGGTGCGGAGAGCGACCTCACCGTTGGCTCAGTTGCTA
 CTGAAGATGTCCCGCATTTCTCGGAAGGTACAAGTGTGGCGAAACGACCGACAGGAGCCCTTGGCACTCTTCGA
 GATGAATTGGCAGATGACCACTGTCTAGAGAACCCCGACACAAACCTCTGCAAGCGAGGTGGCGCGGCTTGTAGT
 TTTGATTCTGGAGGTCGCGGAGCTCACTGACCTCCGCTTCCARACGGTACAGACGCGGCGGAGGGGACCGTTAC
 ACACGGTCAAGAAAGAAAGCTGAGAGGTGCTTTGACCAGCTGAGCCGTGGGTTTTTGAATTGTCTCCCATCTCCCTGTC
 TTCTTCTACGCTTTTCAAGCTGACAGTCACTACTCTTCCGGTGAAGTGTGAGCTTTTACTTTATTTGTGCTC
 CTCTCTATGTTACAGTTACCGGCTTTGGTGTGCTCCCTATTGGGTGATTTTCTGGGCTTCTCGCGCGCTTCGCA
 TGGGGTTTTTGGCTGCTGGTTGCTTTCGCTGTTGGTTTGTCAAGCCTGCACCCGACCCAGTCCGTGCTGTTGTGAG

Fig. 1E-2

TTTGATTGCCAGAGTGTAGAGACATCCTTCATTCTTTGAGCTCCTGCAACCTTGGGATCCTGTTGCGAGCCTTGTTGGT
 GGGACCCGTCGGTCTCGGTCTTGCCATTATTGGCAGGTTACTGGGCGGGGCACGCTACGCTGCTGCTTTTGCTTAGGC
 TTGGCATCGTTTCAGACTGTATCTTGGCTGGAGCTTACGTGCTTTCGCAAGGTAGGTGTAAGGTTGGGGATCTTGT
 ATAAGAACTGCCCCAGTGAGGTCGCCTTCAATGTGTTCCCTTCACACGTGCAACCAGATCGTCACTTGTGACCTGTG
 CGACCGGTTTTGTGCGCCCAAGGGCATGGACCCCATCTTCTCGCCACTGGATGGCGCGGATGCTGGTCCGGCCAGAGCC
 CCGTTGAGCAACCCACTGAGAAACCCATTGCATTGCCCCAGTTGGATGAGAAGAAAATCACGGCAAGGACTGTGGTTGCC
 CAACCTTATGACCCCAACCAAGCTGGAAGTGTTACGAGTCTTGACGGCGGGTGGGGCGATGGTGGCTGAGGCGATTCC
 AAAAGTGGTTAAGGTCTCTGCTGTCCCATTTGAGCCCCCTTCTCCCAACCGGAGTGAAAGTTGATCCTGAATGCAGGG
 TCGTGGTTGACCCAGACACCTTCACAACTGCTCTCGGTCCGGCTACTCCACCACAAACCTCATTCTTGGTGTGGGGAT
 TTTGCCAGCTGAATGGGTTGAAAATCAGACAAATTTCCAAGCCTTCAGGAGGAGGCCATACCTCATGGCGGCTTACA
 TGTGCTTGTCTGATGGCCTTGACATGCTCGTTGGGATTATGTTACCGCGGTGGGTTCTTGTGGTCTGGCACTAACG
 ATCCGTGGTGCACTAACCCGTTTGCCGTCCCTGTCTACGGGCTGGCTCTCTTGCACGTCCAGGTTGTGATCTCCAG
 CATGGCCTTACTCTGCCTTAAACAGCGCTTGTGGCGGGGTTGGTATTAGGAAGTTGCTTTGGTGTGTTTAACTTTGC
 TTCCATCGGGGATGGCTCACAGGTTGAGTTGCAAGGCCGATGTGCTGTGCATTCTGCTTGAATGGCAGCTATGTTT
 GGTACCCCTTCACTGGTTGCTTTGTGTGTTCTTGTGTTGCGCTGGTTTTCTTGCATCCCCTCACCATTCTATGG
 TTGGTGTGTTTTTTGATTCTGTGAACATGCCCTCAGGAATCTTGGCTTTAGTGTGTTGATCTCTCTGGCTCTTGG
 TCGCTATACCAATGTGCTGGCCTTGTACCCCTTATGACATTACCAATACACCAACGGCCCCGCGGCTTGGCGCT
 TGGCCACTGCCCCGATGGGACCTATTTGGCTGCTGTCCGCGCGCTGCGTTGACTGGCCGTACCATGCTGTTTACCCG
 TCTCAACTTGGTGACTCCTTGAGGGCGCTTTAGAACCACAAAGCCTTCACTGAATACCGTCAATGTGGTTGGGCTCTC
 CATGGGCTCGGGCGGGTGTTCACCATGACGGGAAAATTAATGCGTGACCGCGGCACATATCTCACGGGTAACCTG
 CTAGGGTCTCTGGGGTTGGCTTCAATCAAATGTTGGATTTGATGTAAAGGGGATTTGCCATAGCCGATTGTCCGGGT
 TGGCAAGGAGTCGCTCCCAAGTCCAGATCTGCAAGGATGGGTGGAAGTGGCCGCGCTTATTGGCTAACGTCTCTGGCGT
 CGAACCCGGCGTCATTGGTAGGGGATTGCGCTTTGTTTACCAGCGTGGCGGATTCGGGTCCCAAGTATCACCGAGG
 CCGGAGAGCTTGTGAGGATCCACACGGGATCAACAAACAAGGAGGAGGCAATGTACGCGCCCTTCAGGCCAGTTTGT
 AATGTGACACCCCAACCAAGTGAATTGAGTGAATTTCTCGCGGACCCAGGGTCCCGCTTGGTGTGTAAGGTTGG
 CAACCACATAATCAAAGATACAGATGAGGTGCCCTCAGATCTTTCGCGCTTGTGCTGCCAAGCCCGAGTTGGAAGGAG
 GCCTCTCCACCGTTCAACTTCTGTGCGTGTGTTTTCTCCTATGGAGAATGATGGGACATGCCTGGACGCCCTTGGTTGCT
 GTTGGTTTTTTCATCTTGAATGARATCTCCAGCGGTCTGGTCCGGAGTGTGTTCTCCTTTGGAATGTTYGYKTRTC
 YTGGCTCACCCATGGTCTGCGCAAGTCTGATGATCAGGCTTCTGACAGCAGCTTAAACAGGAACAGATGGTCACTTG
 CCTTTTTCAGCCTCGGTGCAAGTACCCGTTTTGTGCGAGATCTTTCGCGCACTCAGGGGCATCCGTTGCAGGCAAGTATG
 AATTTGAGCACCTATGCATTCTGCCTCGGATGATGGTTGTGACCTCACCACTCCAGTATCACGTGTGGTGTGCTGCA
 CCTACTTGCCATCATTTTGTACTTGTAAAGTACCGTGGCCTGCACCATATCCTTGTGGCGATGGAGTGTCTCTGCGG
 CTTTCTTCTTGAGATACTTTCGCGAGGGAAGTTGAGGGAAGGGGTGTCGCAATCTGCGGAATGAATCATGAGTCTGT
 ACTGGTGCCCTCGCTATGAGACTCAATGACGAGGACTTGGATTTCTTATGAAATGGACTGATTTTAAAGTCTTTGTTTC
 TGGCTCAACATAGGAATGCGGCGGGTCAAGTTATCGAGGCCCTTATGCGAAGCGATCAGGGTGGAACTTGCCAGT
 TAGTGCAGGTGCAAGGTTGCGGGTGTGTTAGCCAACTTGAAGCTTTGCTGACACCGTGGCGCCCATCTTACCC
 GCGGACATTGTTGTTGTTCTTGGTCATACGCCCCGTTGGCAGCATCTTGACTTAAAGATTGGCRATGCCAAGCACCCCT
 ACAAGCCATCGAGACCAGAGTCTTGTGGTCCAGGATGACCGTGGCGCGTGTGCTTATCCGACTCCCGCGCCGCCAC
 CCGTACCCGTGCCGTTCTCTCCCAACGAAAGTTTTAGAGAACGGCCCCAGTGCCTGGGGGATGAAGACCGCCTGAAC
 AAAAGAAGCGGCGCAAGATGGAAGCCGTTGGCGTTTACGTATGGGCGGGAAAAAGTACCAGAAATTTGGGATAAGAA
 TTCTGGTGTGTTCTATGAGGAAGTCCACGACAACACAGATGCGTGGGAATGCCTTAGAGCTGACGACCTGCCGACT
 TGGATCCTGAGAGGGGAACCTTGTGTGGACAGTCACCATAGAGAATAGGCCTTACCATGTTTACGCCTCCCGCTGCTGGT
 AGGAAGTTCTGGTCCCTGCCGACCCAGAGAAATGGGAAGCCAGTGGGAAGCTGCAAAGCTTCCATAGAGCAGGCCCT
 TGGTATGATGAACGTTGACGCGGAGCTGACCGCAAGAACTGGAGAAATTAAGAGAATAATTGACAACTCCAGGGCC
 TGACTAAGGAGCAGTGTGTTAACTGTTAGCGCGCAGCGGCTTACCCGCTGTGGTGGCGCGGCTTGGTTATTACTGAGA
 CAGCGGTAAGAAATAGTCAGATCCACAATCGGACCTTACCCTGGGGCTGTGAATTTGAAAGTGGCCAGCGAAGTTGAG
 TTGAAAGACCGGTCGAGCACAACCAACACCCGTTGCAAGACCAAGTTGACGGTGGCGTTGTGCTCTGCGCTCTGCAAGT
 TCCTTCGCTTATAGACGCTTGTATCTCCGTGCGGACGCATCTCCCAAGTTGCTCGCCATCACGGTCCAGGAAACACTG

Fig. 1E-3

GGATTGATGGCACGCTCTGGGATTTT GAGTCCGTAGCCACTAAAGAGGAAGTCGCACTTAGTGACAAATAATACAGGCT
 TGTGGCATTAGGCGTGGCGATGCTCCTGAGATTGGCCTCCCTTACAAGCTGCACCCTGTTAGGGACAACCTGAACGTGT
 AAAAGGGGTTTTGAAAAACACAAGGTTTGGAGACATACCTTACAAGACCCTAGCGACACTGGGAGCCCACTACATGCGG
 CCGCCTGCCTTACGCCTAATGCCACCCCGGTGACTGATGGGCGCTCCGTCTTGGCCACGACTATGCCCTCCGGGTTTGAG
 TTGTATGTGCCGACCATTCAGCGCTGTGCTTGATTACCTTGATTCCAGGCCAGACTGCCCTAAACAGTTGACGGAGCA
 CGGGTGTGAAAATGCTGCATTGAGAGACCTCTCAAATATGACTTGCCACCCCAAGGTTTTGTTTTGCCCGGAGTCTCC
 GCCTCGTGCGGAAATACTTGTGTTGCCACGTGGGCAAGTGCCACCTGTCCATCGGCCCTCCACCTACCCGGCAAGAAT
 TCCATGGCTGGAATAAACGGGAATAGGTTCCCGACCAAGGACATTGAGAGCATCCCTGAGATCGACGTTCTGTGTGCACA
 GGCTGTACGAGAGAACTGGCAGACGTTACCCCTTGCAACCCTCAAGAAGCAGTATTGCGGGAAGAAGAAAACCAAGGACCA
 TACTCGGTACCAATAACTTCATTGCGCTGGCCACCCGGGCAGCACTGAGTGGTGTCAACCAGGGCTTCATGAAAAAGCGG
 TTTAACTCGCCCATCGCCCTCGGGAAGAACAATTCAGAGAGGTACAGACTCCGGTCTGGGCAGATGTCTTGAGGCTGA
 TCTTGCCTCTTGGCATCGGTCCACTCCCGCGATTGTCCGCTGGTTTGGCGCCCATCTCCTTTATGAACCTGCCTGCGCTG
 AGGAGCACCTACCGTGTATGTGTGAATTGCTGCCATGACCTATTGGTCACGCAGTCCGGTGCGGTGACTAAGAGAGGT
 GGCCTGTCTGCTGGTATCCGATCACCTCTGTATCCAACACCATTTACAGTCTGGTAATTTATGCGCAGCATGTTGCT
 CAGTTACTTCAAAGTGGTCACCCACATGGTCTCCTGTATCTCCAGGACCAGCTAAAGTTTGAGGACATGCTTAAGGTTT
 AGCCCTGATYGTCTACTCGGATGATCTTGTGCTGTATGCCGAGTCCCCACCATGCCAACTACCACTGGTGGGTTGAG
 CATCTGAACCTTGATGCTAGGGTTTCAAGCGGACCCAAAGAAGCAACCACTTACTGACTCGCCATCTTTCTGGGCTGAG
 GATAATGAATGGGCGTCAGCTAGTCCCAAACCGTGACAGGATTCTCGCAGCTCTTGCCTACCACATGAAGGCGAATAATG
 TTTCTGAGTACTACGCTCCGCTGCTGCAATACTCATGGACAGTTGTGCTTGTCTGGAGTACGACCCTGAATGGTTTGAA
 GAACCTTGTGGTTGGAATGGCGCTATGCGCCCGCAAGGACGGCTATAGCTTCCCCGGCCCGCGCTTCTTCTTATCCATGTG
 GGAGAACTTAAGTCCAATTATGAGGGGAAGAAGTCAAGGGTATGTGGTACTGCGGAGCTTCGGCCCGTATGCCACTG
 CCTGTGGTCTTGACGTCTGTGTTTACCACACTCACTTTACCAGCATTGTCCAGTCATAATCTGGTGTGGCCACCCTGCA
 GGTTCCAGGTCTGTGATGAGTGCAAATCCCCATAGGGAAAGGCACAAGCCCTCTGGATGAGGTTTGGAGACAAGTCCC
 GTATAAGCCTCCACGGACGCTCCTCATGCATGTGGAGCAGGGCTCACCCCTTGACCCAGGCAGATATCAGACCCGCC
 GTGGGTTGGTTGCCGTTAGGCGCGGGATCAGGGGAAATGAAGTTGACCTACCAGATGGTGATTATGCTAGCACCGCCTTA
 CTCCCAACCTGTAAAGAGATCAACATGGTTGCTGTTGCTTCTAATGTGTTGCGCAGCAGATTTATCATCGGTCCACCCGG
 TGCTGGGAAAACATACTGGCTCCTTCAACAGGTCCAGGATGGTGATGTATATACACCCGACCCATCAGACCATGCTTG
 ACATGATCAAAGCTTTRGGGACGTGCCGGTTAACGTCCCGGCAGGCACAACGCTGCAATTCCCCGTCCCCTCCCGCAC
 GGTCCGTGGGTTTCGCATCCTGGCCGGCGGGTGGTGTCTGGCAAACTCCTTCTGGACGAAGCTGCGTATTGTAATCA
 TCTTGATGTCTTGAGGCTTCTAGCAAAACCACTCTCACCTGTTTGGGGGACTTCAAACAACCTCCACCCAGTGGGTTTG
 ATTCTATTGCTATGTCTTTGACATTATGCCTCAGACTCAATTGAAGACCATCTGGAGATTGGACAGAACATCTGTGAT
 GCCATCCAACAGACTACAGAGACAAGCTTATGTCATGGTCAACACAACCTCGTGAACCTTATGTGGAAAAACCTGTCAA
 ATATGGGCAAGTCTCACCCCTTACCATAGGGACCGAGAGGATAGCGCCATTACCATGACTCCAGTCAAGGCGCCACAT
 TTGATGTGGTTACACTGCATTTGCCACGAAAGATTCACTCAACAACAAGGGCCCTTGTGCTATTACCAGGGCAAGA
 CATGCCATCTTTGTGTATGACCCATATAGGCAACTGCAGAGCCTATTTGATCTTCTGCAAAAAGCACGCCCGTCAACTT
 GGCCGTGCACCACGATGGGCAACTGATTGTGCTAGATAGAAATAACAAAGAATGCACGGTTGCCAAGCTCTGGGTAATG
 GTGACAAATTTAGGGCCACAGACAAGCGCTTGTGGATTCTCTCCGCGCCATTGTGCTGACCTAGAAGGTCGAGCTCT
 CCACTCCCCAAGTTGCACATAATTTGGGGTTTTATTTCTCACCTGATTTGATACAGTTTGCCAAGCTTCCAATAGAACT
 TGCGCCACACTGGCCAGTAGTGACGACCCAGACAATAAACTGGCCAGATCGGCTGGTTGCCAGCCTACGCCCTATTC
 ACAACATAGCCGTGCGTGATCGGTGCCGGCTATATGGTGGGCCCTCGGTGTTTTAGGCACCCCTGGGGTTGTGTCA
 TACTATCTTACAAAATTTGTTAAGGGCGAGGCTCAAGTGCTCCGGAAACGGTCTTCACTACCGGCCGAATTGAGGTGGA
 TTGCCGGGAATATCTTGACGACCGGGAGCGGGAAGTTGCAGCGTCCCTCCACACGCCTTTATCGGCGACGTCAAAGGCA
 CTACCGTCGGAGGGTGTATCACATCACCTCAAATACCTTCCGCGCTTCTCCCAAGGAATCAGTTGCGGTAGTCGGG
 GTTTCAAGCCCCGAAAAGCAGCGAAAGCAGTGTGTACATTGACAGATGTGTACCTCCAGACCTTGAAGCTTACCTCCA
 TCCTAAGACCTGTCCAAGTGTGAAAAATGATGTTGGACTTCAAAGAAGTTGCGCTGATGGTCTGGAAGGACAAGACGG
 CCTATTTCCAACCTCGAAGGTGCGCAATTCACCTGGTATCACTGTGCTAGCTATGCCTCGTACATCCGTGTTCTTTAAAC
 TCCACGGTGTACCTGGACCCCTGCATGGGCCCGCCCTTTGCAACAGAAAAAGTTGTTGGGTCCACTCATTGGGGAGCTGA
 CCTCGCAGTCACCCCTTATGATTATGGGGCAAGAATTATTTGTCTAGTGCCTACCATGGTGAGATGCCTCCTGGGTACA

Fig. 1E-4

AGATTCTGGCGTGC GCGGAGTTCTCGCTGGACGACCCAGTCAGATACAAGCACACTTGGGGGTTTGAGTCGGGATACAGCG
 TACTTGTACGAGTTCACTGGAAACGGTGAGGACTGGGAGGATTATAACGACGCGTTTCGTGCGCGACAGAAAGGAAAGAT
 TTACAAGGCCACTGCCACCAGCCTGAAGTTCATTTTCTCCGGGTCAACCGTTGAACCAACTTTGGGCTAGACTGAA
 ATGAAATGGGGGCTGTGCAGAGCCTATTTGATAAAATTGGCCAACTGTTTGTGGACGCTTTCACGGAGTTCTTGGTGCC
 ATTGTTGATATCATCATATTTTGGCCATTTGTTGGGCTTCACAATCGCCGGTGGCTGGTGGTCTTTGCATCAGATT
 GGTGTTGCTCCGCGATACTCGTTGCGGCTCTGCCGTTCAACCTGAGCAATTACAGAAGATCCTATGAGGCATTTCTCTCC
 CAGTGCCGGACGGAYATCCACCTGGGGAATAAACATCCCTGGGGATGCTCTGGCACCACAAGGTGTCGACCCTAAT
 TGATGAAATGGTGTGCGCTCGAATGTACCGCATCATGGAACAAGCAGGGCAGGCTGCCTGGAACAGGTGGTGACCGAGG
 CAACGTTGTCTCGTATTAGTAGCTTGATGTGGTGGCTCATTTCCAGCACCTTGCCGCCATAGAAGCCGAGACTTGTA
 TACTTGGCCTCCCGGTGCCAATGCTGCACAACCTGCGCATGACAGGGTCAAATGTAACCATAGTGTATAATAGCTCTCT
 AGAACAGGTGTTGCTGTTTTCCGACCCCTCAGTTCGCGGCAAGCTTCATGATTTTCGGCAATGGTAATAGCTGTGC
 ATCTCTCATATTCTCTTCTGTTGCGGCTTCTGTACCCCTTTCTGCTGCTGTGGTTCGGGCTTCCAATAATACGTA
 GTTTTGGTTTCCACTGGTTAGGGCAATTTTCTTTCGAGCTCACAGTGAACACAGGTGTCTCTCCCTGCCCTCACC
 CGGCAGGCGGCGCAGAGATCTACGAACCTAGTGGTCTCTTTGGTGAGGATAGGGCAGATCGATGCTCGGAGGACGA
 TCACGACGAGCTAGGATTTCTGGTGCCGCTGGCCTCTCCAGCGAAGGCCACTTGACCACTGTTTACGCTGGTGGCGT
 TCTTGTCTTTCAGTTACACGGCCAGTTTCAACCCGAGATATTGGGATAGGGAATGTGAGTAAAGTTTATGTTGACATC
 AAGCATCAATTTATTTGCGCTGTTGATGACGGGCAAAACACCACCTTGCTCGCCATGACAACGCTCAGCCGTGTTCCA
 GACTTATTACCAGCATCAGGTGACGGCGGCAATTGGTTTCACTGGAATGGCTGCGCCCTTCTTCTCTCTGTTGG
 TTTTGAACGTCTCTTGGTTTCTCAGGCGTTCGCTGTAAGCCGTGTTTCACTGAGTCTCTCAGACATTAAGACCAACA
 CCACCGCAGCTGACGGCTTGTCTCTCAAGACATCAGTTGTCTTAGGCATGGCCACTCGTCTCTGAGGCGACTCGC
 AAAAGCCGTCAATGTGCGCAGGCGATAGGAACGCCGTATACATTACTGTACAGCCAATGTAACAGATGAGAATTATTT
 GCATTCCTCTGACCTTCTCATGCTTCTCTTGCCTTTCTACGCTTCCGAGATGAGTGAAAAGGGATTTGAAGTGATAT
 TTGGCAATGTGTAGGCATAGTGGCTGTGTGTGCACTTTACCAGCTATGTCCAACATGTCAAGGAGTTCACCCAGCGC
 TCCTTGGTGGTTGACCATGTGCGGTTACTTCATTTTATGACACCTGAGACTATGAGGTGGGCGACCGTTTACGCTGTCT
 TTTTGCATTTCTGTTGGCATTGGAATGTTGAGATATGTTGGGAAATGCTTGACCGCGGGCTATTGCTCGCAATTGCTT
 TTTTGTGGTGATCGTGCGGTTCTGTCTTGTGCTCGCTCGTCAACGCCGACAGCAACAGCAGCTCCCATTTACAGTTGAT
 TTATAAMTTAACGATATGTGAGCTGAATGGCACAGACTGGCTGAACAATCATTTTAGTTGGGAGTGGAGACTTTCGTTA
 TCTTCTCTGTGTTGACTCATATTGTTTCTACGGCGCCCTCACTACCAGCCACCTCTTGACACGGTGGGCTGATCACT
 GTGTCCACCGCGGATACTGCCATAAGCGGTATGTCTTGAGTAGCATCTATGCTGTCTGCGCCCTGGTGCGCTGATTTG
 CTTCGTATCAGGTTGACGAAAAATGTATGTCCTGGCGCTACTCATGTACCAGATATACCACTTTCTTCTGGACACCA
 AGGGCAGACTCTATCGCTGGCGGTCAACCGTCATCATAGAGAAAAGGGGTAATGAGGTTGAGGTGACCTGATCGAC
 CTCAAGAGAGTTGTGCTTGATGGTTCCGCGGCAACCCCTGTAACCAAGTTTCAGCGGAACAATGGGGTCTGCTTAGAC
 GACTTCTGCAATGACAGCACGGCTCCACAAAAGGTGATCTTGGCATTTTCTATCACCTACACACCACTGATGATATATGC
 CCTAAAGGTGAGTCGTGGCCGGCTGCTAGGGCTTTTACACCTTTTATTTTAACTGTGCTTTTACCTTCGGGTATA
 TGACATTTGTGCACTTTAGAGCACAACAGAGTTGCACTCACTATGGGAGCAGTAGTCGCGCTCCTTTGGGGGTGTAC
 TCAGCTATAGAAACCTGGAAATTCATCACTCCAGATGCCGTTTGTGCTTGCTAGGCCGCAAGTACATTCTGGCCCTGC
 CCACCAGTTGAGAGTGCCGACGGCTTTTATCCGATTGCGGCAAGTGATAACCAACGCAATTTGCTGCGCGCTCCCGTT
 CCACTACGGTTAACGGCACATTGGTGCCCGGTTGAAAAGCCTCGTGTGGGTGGCAGAAGAGCTGTCAAACGGGGAGTG
 GTAAACCTCGTTAAATATGCCAAATAACAACGGCAGGCAGCAGAGAAAAAGAAAGGGGACGGCCAGCCAGTCAATCAGC
 TGTGCCAAATGTTGGGAGGATCATGCCAGCAAAACAGTCCAGAGGTAAGGGACCGGGGAGAAAGTAAGAAGAAA
 AGCCCGGAGAAGCCCCATTTTCTCTCGGACTGAAGATGACGTTAGACATCACTTCAACCTAGTGAGCGGCAATTGTG
 TCTGTGCTCAATCCAGACTGCCTTAACCAAGGCGCTGGAATTTGACCCTGTCGGATTAGGGAGAATAAGTTACGCTG
 TGGAGTTTAGTTTGCCTACGCATCATACTGTGCGCTAATTCGCGTCACAGCATCACCTCAGCATGATGAGCTGGCATT
 CTTGAGACATCCAGTGTTTGAATTGGAAGGATGTGTGGTGAATGGCACTGATTGATTTGTGCTTYAAGTCACCTATT
 CAATTAGGGCGACCGTATGGGGTAATTTAATTGGCGTGAACCATGCGGCCGAAATT

Fig. 1F-1

>MN184B_DQ176020.seq

ATGACGTATAGGTGTTGGCTCTATGCCACGACATTTGTATTGTGAGGAGCTGTGACCACTGGCACAGCCCAAAGCTTGCT
 GCACAGAAACACCTTCTGTGACGGCCTCTTCAGGGGAGTTAGGGGTTTGTCCCTAGCACCTTGTTCTGGAGTTGCA
 CTGCTTTACGGTCTCTCCACCCCTTTAACCATGTCTGGGATTCTTGATCGGTGCACGTGCACCCCAATGCCAGGGTGT
 TATGGCAGAGGGCCAAAGTCTACTGCACACGATGTCTCAGTGCACGGTCCCTCCTTCCCCTGAATCTCCAAGTCTCTGAGC
 TCGGAGTGTGGGCTTGTGCTTTATAGGCCTGAAGAGCCGCTCCGGTGGACGTTGCCACGCGCATTCCCCACTGTTGAGTGC
 TCCCCTGCTGGGGCTTGTGGCTTTCTGCAATTTTCCAATTGCACGAATGACCAGTGGAAACCTGAACCTTCAACAAAG
 ATTAGTGCGGGTGCGAGCTGAGCTTTACAAAGCCGGCTGCCTCACCCCTACAGTCTAAAGAGTCTACAAGTCTATGAAC
 GGGGTTGCCGCTGGTACCCCATCGTTGGACCTGTCCCTGGAGTTGCCGTTTTCGCCAACTCCCTACATGTGAGTGATAGA
 CCTTCCCAGGTGCTACTCACGTGCTAACCAACCTGCCGCTCCCGCAGAGACCTAAGCCTGAAGATTTTGGCCCTTTGA
 GTGTGCTATGGCTGCCGTCTATGACATTGGTCATGACGCCGTTATGTTCTGTGGCCGAAGGGAGAGTCTCTGGGCTCCGC
 GTGGTGGGGAAGGAAATTTGAAACTGTTCCCGAGGAGTTGGGGTTGATTGCAGAGCACTTTATACCTCCTTCCCG
 CCCCACTTGGTGGACATGTGAAATTCACCTTACGGCCCTGAGTGTGGTCTTCCATGCGAGTGAACGCCAGTA
 TGGCTGCCTCCCCCTGGCACTGTCCCTGACGGCAATTTGGTGGAGCTTGTAGCTCGCTCCCATTTGGAAGTCCAGT
 ATAAAGAAATTCGCTACGCCACCAATTTGGCTATCAAACTAAGCATGGCGTTGCTGGCAAGTACCTACAGCGGAGGCTG
 CAAATTAATGGTCTCCGAGCAGTGGTTGACTCGAATGGACCATCGTCATACAGTACTTCTCTGTTAAGGAGAGTGGAT
 CCGCCACGTGAACTGGCGGAAGAGTTGACTACCTGGGTTTGGAGTCTCCTCAGGATAAGAGTCGAGCCCAACAGT
 TGCCATTGTCCAACAGGACGAGAAATCTTCGGTTTGGTGGTGCAAGTGGTACGGTCTGGGAAGAGGGCAAGGAGG
 GCACGTGCAAGTGCAGTACCCGAGTCCCGGTGACGCTCCGCTACTCGTGAACCCAGCAAGCAAGAAACACGAAGC
 TGCTAGTGCCAACAGGCTGAGCTTCTGAACGCTACTCCCGCTGCTGAAGGGAATTGCGGCTGGCACTGTATCTCCG
 CCATCGCAACCGGATGGTRAATTCYAARTTTGAAACYRCCCTTCCCGAAAGAGTGAGACCTCCAGATGACTGGGCTACT
 GACGAGGATCTTGGAATGCCATCCAAATCCTCAGACTCCCTGCGGCCTTAGACAGGAACGGTGCTTGTACTAGCGCAA
 GTACGTAATTAAGCTGGAAGGTGAGCATTGGACTGTCACTGTGACCCCTGGGATGTCCCTTCTTTGCTCCCTCTTGAAT
 GTGTTGAGGCTGTTGTGGGCACAAGGGCGGTCTTGGTTCCCGAGATGCAGTCGAGGTCTCCGGATTTGACCCTGCCTGC
 CTTGACCGGCTGGCTGAGGTGATGCACCTGCCTAGCAGTGTATCCAGCCGCTCTGGCCGAAATGTCTGGCAGTCCGA
 TCGTTGGCTTCTCCGGTCAACCTGTGGACTGTTTCGAGTTCTTTGCCGTCAACGCGGAGGGAATCACCTGACC
 AAGTGGCTTAGGGAAATTTATCAGCCTTTGTGAGGTGATTGAGGACTGCTGCTGTTCCAGAACAAACCAACCGGCTC
 ACCCGGAGGAGGTGCGAGCAAGATTGAYCAGTACCTTTTGGTGCAGCAAGTCTTGAAGAATGCTTGGCCAGGCTTGA
 RAAAGCTCGCCCGCAAGCGTATTARACACCTCCTTTGATTGGGATGTTGTGCTCCCTGGTGTGGGGCGGTGCTCAAG
 CAGCAAACTGCCCTCACCAACAGCGTCACGCTCTAGCCACTGTTGTGACTCAAAGGTCTTGGCGAAATTTCAACCT
 CGAAAGCGGAGTCTGTCAAGAGCTACCAGAGAGCAGGCCCCCTCCCTGCCCGCGCAAAAGATTGGGTCCAGGTGTGG
 TAGTCCGATTTATTGGGCGCAATCTCCCTGACAGCCGGGAAGACTTGGCCGGTGGTCTTTGATTTCCAACCTTAC
 CTGAGTTGGTGGCAAGCTCGAGCGAGCTGTGCTGTCCCTGCACCGCGCAGGGTTGTGTCCCGATTAGTGTCTCGG
 ATAGTGTGACCCCTGTGCCCGCACACGACGTGGGCTTGGCAGGTGGAGGGAATGAATTTGGCGGCGGTGACTTAGC
 GTGCCAGGACGAGCCCTCGATTTGTCTGCGTCTCGCAGACTGAATATGAGGCGTCCCTTGGCATTGCCGTGAGTG
 AGGATGTCTGGCGGTGGAGAGACGAGAAAGTTGAAGAAGTCTGAGCGGAATATCGGGCATGCCAGATGACATCAGGTTG
 GCGCCCGTGTCAAGTAGCTCCCTGTCAAGCATAGAGATCACACGTCAAAGTACTCAGCTCAAGCCATCATTAACCTC
 AGGTGGGCTGTTGTGGGCACCTCCAGGAGGTAAGAGAAATACCTTAATGTGATGCGTGAGGCATGTGATGCGACCA
 AGCTTGATGACCTGCCACGCAAGAAATGGCTTTCCCGCATGTGGGATAGGGTAGACATGCTAACCTGGCGCAACACGTCC
 ATTTTTCAGGCGCCTTTACCTTGGCTGACAAGTTTAAGACCTCCCGAAGATGATACTCGAAACACCGCCGCTTACCC
 TTGTGGGTTTGTGATGATGCCCCGACGCTGACCTTCTGTAGGTGCGGAGAGCGACCTCACCGTTGGCTCAGTTGCTA
 CTGAGGATGTCCCGCATTTCTCGGAATGTACAAGGTGTTGGCGAAACGACCGACCGGAGGACCTTGGCACCTTCGCA
 GACGAATTGGCAGATGACCACTTGTAGAGAACCCCGACACAAACCTCCTGCAAGCACAGGTGGCGCCGGCTTGGT
 TTCGGATTCTGGAAGGTGCGCGGAGCTCACTGACCTGCCGCTTCAAACGGTACAGACGCGGGCGGAGGGGGCGGTAC
 ACACGGTCAAGAAGAAAGCTGAGAGGTGCTTGAACAGCTGAGCCGTGGGTTTTGACATTGTCTCCATCTCCCTGTT
 TTCTTCTACGCCCTTTCAAGCTGACAGTCACTACTTTCGGGTGACTGGAGTTTGCAGCTTTTACTTTATTGTGCTT
 CTTCTATGTTACAGTTACCAAGCCTTGGTGTGCTCCCTATTGGGTGATTTTCTGGGTCTTCTCGGCGGCTTCGCA
 TGGGGTTTTTGGCTGCTGGTGGCTTTCGCTGTTGGTTGTTCAAGCCTGCACCCGACCCAGTCCGTGCTGCTGTGAG

Fig. 1F-2

TTTGATTCGCCAGAGTGTAGAGACATCCTTCATTCTTTTGAGCTTCTGCAACCTTGGGACCCTGTTGCGAGCCTTGTGGT
 GGGGCCCGTCGGTCTCGGTCTTGCCATTATTGGCAGGTTACTGGGCGGGGCACGCTACGTCTGGCTGCTTTTGCTTAGGC
 TTGGCATCGTTTCAGACTGTATCTTGGCTGGAGCTTATGTGCTTTCGCAAGGTAGGTGTAAAAAGTGTGGGGATCTTGT
 ATAAGAACTGCTCCAGTGAGGTGCGCTTCAATGTGTTTCCCTTCACACGTGCAACCAGATCGTCACTTGTGACCTGTG
 CGACCGGTTTTGTGCGCCCAAGGGCATGGACCCCATCTTCCTCGCCACTGGATGGCGCGGATGTTGGTCCGGCCAGAGCC
 CCATTGAGCAACCCACTGAGAAACCCATTGCGTTCGCCCAGTTGGATGAAAAGAAAATCACGGCAAGGACTGTGGTTGCC
 CAACCTTATGACCCCAACCAAGCTGTGAAGTCTTACGAGTCTTGACGGCGGGTGGGGCGATGGTGGCTGAGGCGGTTCC
 AAAAGTGGTTAAGGTCTCTGCTGTCCCATTTGAGGCCCTTCTTCCCCCGGAGTGAAAGTTGATCCTGAATGCAGGG
 TCGTGGTTGACCCAGACACCTTCACAACTGCTCTCCGGTCCGGCTACTCCACCACAAACCTCATTCTTGGTATGGGGGAT
 TTTGCCCAACTGAATGGGTTGAAAATCAGACAAATTTCCAAGCCTTCAGGAGGTGGTCCATACCTCATGGCGGCTTACA
 TGTGCTTGTCTGATGGCTTGCACATGCTGTTGGGATTATGTTACCGCGGTGGGTTCTTGTGGTTCTGGCACTAACG
 ATCCGTGGTGCACTAACCCGTTTGCCGTCCCTGTCTACGGGCTGGCTCTTTGACGTCAGGTTGTGCATCTCCAG
 CATGGCCTTACTCTGCCTTTAACAGCGCTTGTGGCGGGGTTGGCATTGAGGAAGTTGCTTGGTTGTTTAACTTTAC
 TTCCATCGGGGGTATGGCTCAGAGTTGAGCTGCAAGGCCGATGTGCTGTGATTCTGCTTGAATTGCCAGCTATGTT
 GGGTACCCTTCACTGGTTGCTTGTGTGTTTCTTGTGGTGGCTGGTCTTCTTGCATCCCTCACCATTCTATGG
 TTGGTGTGTTTTCTTGATTTCTGTGAACATGCCCTCAGGAATCTTGGCTTGTGTTGATCTCTCTGGCTCCTTGG
 TCGCTATACCAATGTGCTGGCCTGTACCCCTTATGACATTACCATTAACCAACGGCCCCCGGGCGTTGCCGCT
 TGGCCACTGCCCCGGATGGGACCTATTTGGTGTCTGTCCGCCGCTGCGTTGACTGGCCGCACCATGCTGTTTACCCCG
 TCTCAACTTGGGTCACTCCTTGAGGGCGCTTTAGAACCACAAAGCCTTCACTGAATACCGTCAATGTGGTTGGGTCTC
 CATGGGCTCCGGCGGGGTGTTACCATTGACGGGAAAATTAAGTGCCTGACCGCCGCACATATCTCACGGGTAACCTGTG
 CTAGGTCTCTGGGTTGGCTTCAATCARATGTTGGATTTTGTGTAAGGGGATTTTGCCATAGCCGATTGTCCGGGT
 TGGCAGGGAGTCTGCTCCAAGTCCAGTTCTGCAAGGATGGGTGGACTGGCCGCGCTTATTGGCTAACGTCTCTGGCGT
 CGAACCCGGCGTCATTGGTAGGGGATTGCGCTTTTGTTCACCGCTGCGGCGATTCCGGGTCCCCAGTGATCACCGAGG
 CCGGAGAGCTTGTGCGAGTCCACACGGGATCAACAAACAAGGAGGAGGATTGTACGCGCCCTTACGGCCAGTTTGT
 AATGTGRCACCCAYCAARYTAAGYGAATTRAGTGAATTTCTTYGCGYGRCTARGGTCCCGCTYGGTGAYGTGAAGTCCG
 CAGCCACATAATTAAGACATAAGCGAGGTGCCCTCAGATCTTTGTGCTTGTGCTGCCAAACCTGAAGTGAAGGAG
 GCCTCTCCACCGTCCAACCTCTTTGTGTGTTTTTCTCCTGTGGAGAATGATGGGACATGCCTGGACGCCCTTGGTGT
 GTGAGTTTCTTATTTTGAATGAGGTTCTCCAGCCGCTCTGGTCCGGAGTGTTTTCTCCTTGGAAATGTTTGTGCTATC
 CTGGCTCACGCCATGGTCTGCGCAAGTTCTGATGATCAGGCTTCTGACAGCAGCTTTAACAGGAACAGATGGTCACTTG
 CCTTTTTCAGCTCGGTGCACTGACCGGTTTTGTGCGAGATCTTGGGCCACTCAGGGGCATCCGTTGACGGCAGTGATG
 AATTTGAGCACCTATGCATTCTGCTCGGATGATGGTTGTGACCTCACCAGTCCAGTGATCACGTGTGGTGTGCTGCA
 CCTACTTGCCATCATTTTGTACTTGTGTTAAGTACCGTGGCTGCACCATATCCTTGTGGCGATGGAGTGTCTCTGCGG
 CTTTCTTCTGAGATACTTTGCCGAGGGAAGTTGAGGGAAGGGGTGTGCAATCCTGCGGAATGAATCATGAGTCTCTG
 ACTGGTGCCCTCGCTATGAGACTCAATGACGAGGACTTGGATTTCTTATGAAATGGAAGTATTTAAGTCTTTGTTT
 TGCGTCCAACATGAGGAATGCAGCGGGTCARTTTATCGAGGCGCTATGCGAAAGCGATCAGGGTGGAACTTGCCAGT
 TAGTGCAGGTGCAAGGTTGCGGGTGTGTTAGCCAAACTGAAGCTTTTGTGACACCGTGGCGCCCATCTTTCACCC
 GCGGACATTGTTGTTGTTCTTGGTCATACGCCGTTGGCAGCATCTTTGACTTAAAGATTGGCAATGCCAAGCACACCT
 ACAAGCCATCGAGACAGAGTCTTGTGGTCCAGGATGACCGTGGCGCGTGTGTTGATCCGACTCCCGCGCCGCCAC
 CCGTACCCGTCGCCGTTCTCTCCACCGAAAGTTTTAGAGAACGGCCCAAGTGCCTGGGGGATGAAGACCGCTGAAC
 AAAAAGAACGCGCGCAAGATGGAAGCCGTTGGCATTACGTTATGGGCGGGAAAAAGTACCAGAAATTTGGGATAAGAA
 TTCTGGTGATGTGTTCTATGAGGAAGTCCACGACAAACACAGATGCGTGGGAATGCCTTAGAGCTGACGACCCCGCGACT
 TGGATCCTGAGAGGGGAACCTTGTGTGGACACGTACCATAGAGAATAGGCCTTACCATGTTTATGCCTCCCGCTGGT
 AGGAAGTTCTGGTCCCTGCCGACCCAGAGAATGGGAAAGCCAGTGGGAAGCTGCAAAGCTTCCATGGAGCAGGCCCT
 TGGTATGATGAACGTTGACGGCGAGCTGACCGCAAAGAAGTGGAGAAATTGAAGAGAATAATTGACAACTCCAGGGCC
 TGAATAAGGAGCAGTGTGTTAACTGTAGCCGCCAGCGGCTTGAACCGCTGTGGTCCGGCGGCTTGGTTATTACTGAGA
 CAGCGGTAAAAATAGTCAGATTCACAATCGGACCTTCAACCTGGGGCTGTGAATTTGAAAGTGGCCAGCGAAGTTGAG
 TTGAAAGACGCCGTCGAGCACAACCAACCCGGTTGCAAGACAGTTGACGGTGGCGTTGTGCTCTGCGCTCTGAGT
 TCCTTCGCTTATAGACGTCTTGATCTCCGGTCCGACGCATCTCCAGTTGCTCGCCATCACGGGCCAGGAAACACTG

Fig. 1F-3

GGATTGATGGCACGCTCTGGGATTTT GAGTCCGTAGCCACTAAAGAGGAAGTCGCACTTAGTGACAAATAATACAGGCT
 TGTGGCATTAGGCGTGGCGATGCTCCTGAGATTGGCTCCCTTACAAGCTGCACCCTGTTAGGGGCAACCCTGAACGTGT
 GAAAGGGGTTTTGAAAAACAAGGTTTGGAGACATACCTTACAGGACCCTAGCGACACTGGGAGCCAGTACATGCGG
 CCGCTGCTTACGCCTAACGCCACCCGGTGACTGATGGGCGTCCGTCTTGGCCACGACTATGCCCTCCGGGTTT GAG
 TTGTATGTGCCGACCATTCCAGCATCTGTCTTGATTACCTTGATTCCAGGCCAGACTGCCCTAAACAGTTGACGGAGCA
 CGGGTGTGAAGATGTGCATTGAGAGACCTCTCCAAATATGACTTGTCCRCCCAAGGTTTTGTTTTGCCCGGAGTCTCC
 GCCTCGTGCGAAATACTTGTTTGCCACGTGGGCAAGTGCCACCTGTCCATCGGCCCTCCACCTACCCGGCCAAGAAT
 TCCATGGCTGGAATAAACGGGAATAGGTTCCCAACCAAGGACATTGAGAGCATCCCTGAGATCGACGTTCTGTGTGCACA
 GGCTGTACGAGAGAACTGGCAGACCGTTACCCCTTGACCCTCAAGAAGCAGTATTGCGGGAAGAAGAAAAACAGGACCA
 TACTCGGTACCAATAACTTCATTGCGCTGGCCACCGGGCAGCACTGAGTGGTGTCAACCAGGGCTTCATGAAAAAGGCG
 TTTAACTCGCCCATCGCCCTCGGGAAGAACAATTCAAGGAGCTACAGACTCCGGTCTGGGAGATGCCTTGAGGCTGA
 TCTTGCTCTTTGCGATCGATCACTCCGCGATTGTCCGCTGGTTTGGCGCCATCTCTTTATGAATCTGCTGCGCTGA
 AGGAACACCTACCGTGTATGTGCTGAATTGCTGCCATGACCTATTGGTCACGCAGTCCGGTCCGGTACTAAGAGAGGT
 GGCCTGTCTGCTGGTATCCGATCACCTCGGTATCCAACACCATTTACAGTCTGGTATTATGCGCAGCACATGGTGTCT
 CAGTTATTTCAAAAGTGGTCACCCACATGGTCTCTGTTTCTCCAGGACCAGCTAAAGTTTGAGGACATGCTTAAGGTTT
 AGCCCTGATTGTCTACTCGGATGATCTTGCTGTATGCCGAGTCTCCACCATGCCAACTATCACTGGTGGGTTGAG
 CATCTGAATTTGATGCTAGGTTTTCAGACGGACCCAAAGAAGACAACCATTAAGTCTGCTGCGCATCTTTCTGGGCTGTAG
 GATAATGAATGGGCGTCAGCTAGTCCAAACCGTGATAGGATTCTCGCAGCTTTCCTACCATGAAGGCGAATAATG
 TTTCTGAGTACTACGCTCCGCTGCTGCAATACTCATGGACAGTTGTGCTTGTCTGGAGTACGACCTGAATGGTTTGA
 GAATTTGTGGTTGGAATGGCGCAATGCGCCCGCAAGGACGGCTATAGCTTCCCGGCGCCGCTTCTTTATCCATGTG
 GGAGAACTCAGGTCCAATTATGAGGGGAAGAAGTCAAGGTTGTGGGTACTGCGGAGCTTCGGCCCCGTATGCCACTG
 CCTGTGGTCTTGACGTCTGTGTTTACCACACTCACTTTCACCAGCATTGTCCAGTCATAATCTGRTGTGGCCACCCTGCA
 GGTTCCAGGTCTGTGATGAGTGCAAATCCCCATAGGGAAAGGTACAAGCCCTCTGGATGAGGTTTTAAGACAAGTCCC
 GTATAAGCCTCCACGGACCGTCTCATGATGTGGAGCAGGGCTCACCCCCCTTGACCCAGGCAGATATCAGACCCGCC
 GTGGGTTGGTTGCCGTTAGGCGCGGGATCAGGGGAATGAAGTTGACCTACCAGATGGTGATTATGCTAGCACCGCCTTA
 CTCCCAACCTGTAAGAGATCAACATGGTGTGTTGCTTAAATGTGTGCGCAGCAGATTTATCATCGGTCCACCCGG
 TGCTGGGAAAACTACTGGCTCTTCAACAGGTCCAGGATGGTGATGTATATACACACCGACCCATCAGACCATGCTTG
 ACATGATCAAGCTTTGGGGACGTGCCGTTTAACTGCCCCGAGGCAACGCTGCAATCCCCGCCCTTCCCGCACT
 GGCCCGTGGGTTTCGATCTTGCCGCGCGGGTGGTGTCTGGCAAAAACCTCTTCTGGACGAAGCTGCGTATTGTAATCA
 TCTTGATGTCTTGAGGCTTCTAGCAAAACCACTCTCACCTGTTTAGGGGACTTCAAACAACCTCACCCAGTGGGTTTTG
 ATTCTCATTGCTATGTCTTTGACATTATGCCTCAGACTCACTGAAGACCATCTGGAGATTTGGACAGAACATCTGTGAT
 GCCATCCAACCAAGTACAGAGACAAGCTTATGTCCATGGTCAACACAACCTCGTGAACTTATGTGGAACCACTGTCAA
 ACATGGGCAAGTCTCACCCCTTACCATAGGGACCGAGAGGATAGCGCCATTACCATTGACTCCAGTCAAGGCGCCACAT
 TTGATGTGGTTACTGCAATTTGCCACGAAAGATTCACTCAACAAACAAAGGGCCCTTGTGCTATTGCCAGGGCAAGA
 CATGCCATCTTTGTGTATGACCCACATAGGCAACTGCAGAGCCTATTTGATCTTCTGCAAAAAGCACGCCCGTCAACTT
 GGCCGTGCACCAAGATGGTCAACTGATTGTGCTAGATAGAAATAACAAAGAATGCACGGTTGCCCAAGCTCTGGGTAATG
 GTGACAAATTTAGGGCCACAGACAAGCGCTTGTGGATTCTCTCCGCGCATTTGTGCTGACCTAGAAGGGTCGAGCTCT
 CCACTCCCCAAGGTTGCACATAATTTGGGGTTTTATTTCTCACCTGATTTGACACAGTTTGCCAAGCTTCCAATAGAAT
 TGCGCCACACTGGCCAGTAGTGACGACCCAAGACAATAAAACTGGCCAGATCGGCTGGTGGCAGCCTGCGCCCTATTTC
 ACAACATAGCCGTGCGTGATCGGTGCCGGCTATATGGTGGGCCCCCTCGGTGTTTTAGGCACCCCTGGGGTTGTGTCA
 TACTATCTTACAAAATTTGTTAAGGGCGAGGCTCAAGTGCTTCCGGAACCGTCTTCACTAGTGGCCGAATTTAGGTTAGA
 TTGCCGGGAATATCTTGACGACCGGGAGCGGGAAAGTTGCAGCGTCCCTCCACACGCCTTTATCGGCGACGTCAAAGGCA
 CTACCGTCGAGGGGTGTCATCATCACCTCCAAATACCTTCCGCGCTTCTCCCAAGGAATCAGTTGCGGTAGTCCGG
 GTTCAAGCCCCGGAAAGCAGCGAAAGCAGTGTGTACATTGACAGATGTGTACCTCCAGACCTTGAAGCTTACCTCCA
 TCCTAAGACCCTGTCAAAGTCTGGAAAATGATGTTGGACTTCAAAGAAGTTCGGCTGATGGTCTGGAAGGACAAGACGG
 CCTATTTCAAACCTGAAGGTGCGCATTTTCACTGGTATCACTTGTAGCTATGCCTCGTACATCCGTGTTCTTTAAAC
 TCCACGGTGACTGGACCCCTGCATGGGCCCCGCCCTTGAACAGAAAAGTGGTGGTCCACTATTGGGGAGCTGA
 CCTCGCAGTACCCCTTATGATTATGGGGCAAGAATTATTTGTCTAGTGCGTACCATGGTGAGATGCCTCTGGGTACA

Fig. 1F-4

AGATTCTGGCGTGCGCGGAGTTCTCGCTGGACGACCCAGTCAGATACAAGCACACTTGGGGGTTTGAGTCGGATACAGCG
 TACTTGACGAGTTCACTGGAAACGGTGAGGACTGGGAGGATTATAACGACGCGTTTCGTGCGCGACAGAAAGGAAAGAT
 TTACAAGGCCACTGCCACCAGCCTGAAGTTCCATTTTCTCCGGGTCATACCGTTGAACCACTTTGGGCTTAGACTGAA
 ATGAAATGGGGGCTGTGCAGAGCCTATTTGATAAAATTGGCCAACTGTTTGTGGACGCTTTCACGGAGTTCTTGGTATCC
 ATTGTTGATATCATCATATTTTGGCCATTTTGTTCGGCTTACAATCGCCGGTTGGCTGGTGGTCTTTTGCATCAGATT
 GGTTTGCTCCGCGATACTCCGTTCCGCGCTCTGCCGTTACCCCTGAGCAATTACAGAAGATCCTATGAGGCATTTCTCTCC
 CAGTGCCGGACGGACATTCCACCTGGGGAACTAAACATCCCTTGGGGATGCTCTGGCACCACAAGGTGTGACCCCTAAT
 TGATGAAATGGTGTGCGTGAATGTACCGCACCATTGGAACAAGCAGGGCAGGCTGCCTGGAGACAGGTGGTGACCGAGG
 CAACGTTGTCTCGTATTAGTAACCTGGATGTGGTGGCTCATTTCCAGCACCTTGGCGCCATAGAAGCCGAGACTTGTA
 TACTTGGCCTCCCGGCTGCCAATGCTGCACAACCTGCGCATGACAGGGTCAAATGTAACCATAGTGATAATAGCTCTCT
 AGAACAGGTGTTTGCTATTTTCCGACCCCTCGATTCCCGGCCAAAGCTTCATGATTTTCCGCAATGGCTAATAGCTGTGC
 ATTCTCCATATTCTCTCTGTTGCGGCTTCTGTACCTTTTCTGTGCTGTGGTTCGGGCTTCCAATAACGTA
 GTTTTGGTTTCACTGGTCAGGGCAATTTTCTTCGAGCTCACAGTGAACACAGGTGTCTCTCCCTGCCTCACC
 CGGCAGGCGGCCGAGAGATCTACGAACCTGGTGGTCTCTTTGGTGCAGGATAGGGCACGATCGATGCTCGGAGGACGA
 TCACGACGAGCTAGGATTTCTGGTCCGCGCTGGCCTCTCCAGCGAAGGCCACTTGACCAAGTGTACGCTGTTGGCGT
 TCTTGCTCTTCAGTTACACGGCCAGTTTCACCCCGAGATATTCGGAATAGGGAATGTGAGCCAAGTTTATGTTGACATC
 AAGCATCAATTTATTTGTGCTGTTGATGACGGGCAAAACACCACTTGCCTCGCCATGACAACGCTCTCAGCCGTGTTCCA
 GACTTATFACCAGCATCAGGTGACGGCGGCAATTGGTTTCACTGGAATGGCTGCGCCCTTCTTCTCCTCTGGTTGG
 TTTTGAACGTCTCTGGTTTCTCAGGCGTTCGCTGTAAGCCGTGTTTCACTTCGAGTCTTTCAGACATTAAGACCAACA
 CCACCGCAGCTGCAGGCTTTGCTGTCTCCAAGACATCAGCTGTCTTAGGCATGGCCACTCGTCTCTGAGGCGACTCGC
 AAAGGCCGCAATGCCGACGGCGATAGGAACGCCGATACATTACTGTACAGCCAATGTAACAGATGAGAATTATTT
 GCATTCCTCTGACCTTCTCATGCTTCTCTTGGCTTTTCTACGCTTCCGAGATGAGTGAAAAGGGATTTGAGGTGATAT
 TTGGCAATGTGTAGGCATAGTGGCTGTGTGTGCACTTTACCAGCTATGTCCAACATGTTAAGGAGTTTCAACCGCGC
 TCCTTGGTGGTTGACCATGTGCGGTTACTTCATTTTGTGACACCTGAGACTATGAGGTGGGCGACCGTTTATGCTGTCT
 TTTTGCATTCTGTTGGCCATTTGAATGTTGAGATATGTTGGGGAAATGCTTGACCGCGGCTATTGCTCGAATTGCCT
 TTTTGTGGTGTATCGTGCGGTTCTGTCTGCTGCGCTCGTCAACGCCAGCAGCAACAGCAGCTCCCACTTACAGTTGAT
 TTATACTTAACGATATGTGAGCTGAATGGCAGACTGGCTGAATGATCATTTTAGTTGGGAGTGGAGACTTTCGTTA
 TCTTCTGTTGACTCACATTGTTTCTACGGCGCCCTCACTACCAGCCACTTCTTGACACGGTGGCCTGATCACT
 GTGTCCACCGCCGATACTACCATGCGCGGTATGTCTGAGTAGCATCTATGCCGTCTGCGCCCTGGCTGCGCTGATTTG
 CTTGTCATCAGGTTGACGAAAAATTGTATGTCTGGCGCTACTCATGTACCAGATATACCACTTTCTTCTGGACACCA
 AGGGCAGACTCTATCGCTGGCGGTACCCGTCATCATAGAGAAAAGGGTAAAAATTGAGGTTGGAGGTGACCTGATCGAC
 CTCAAGAGAGTTGTGCTTGATGGCTCCGCGGCAACCCCTGTAACCAAAAGTTTCAAGCGGAACAATGGGGTCTGCTTAGAC
 GACTTCTGCAATGACAGCAGGCTCCACAAAAGGTGATCTTGGCATTTTCTATCACCTACACTCCAGTGATGATATATGC
 CCTAAAGGTGAGTCGTGGCCGGCTGTAGGGCTTTACACCTTTTGATTTTCTAACTGTGCTTTTACCTTGGGTATA
 TGACATTTGTGACTTTTCAAGACAAAAAGAGTTGCACTCACTATGGGAGCAGTAGTCGCGCTCCTTTGGGGGTGTAC
 TCAGCTATAGAACTGGAATTCATCACTTCCAGATGCCGTTTGTGCTTGTAGGCCGCAAGTACATTCTGGCCCTGC
 CCACACGTTGAGAGTGCCGACGGCTTTCATCCGATTGCGGCAAGTGATAACCACGCATTTGCTGTCGGCGCTCCCGTT
 CCACTACGGTTAACGGCACATTGGTTCCCGGTTGAAAAGCCTCGTGTTGGGTGGCAGAAGAGCTGTCAAACGGGGAGTG
 GTAAACCTCGTTAAATATGCCAAATAACAACGGCAGGCAGCAGAAGAAGAAGAAAGGGGACGGCCAGCCAGTCAATCAGC
 TGTGCCAAATGTTGGGAGGATCATCGCCAGCAAAACAGTCCAGAGGTAAGGGACCGGGGAGAAAGTAAGAAGAAA
 AGCCTGGAGAAGCCCCATTTTCTCTCGGACTGAAGATGACGTTAGACATCACTTCAACCTAGTGAGCGGCAATTGTG
 TCTGTGCTCAATCCAGACTGCCTTTAACCAAGGCGCTGAACTTGTACCTGTGCGATTGAGGAGAAATAGTTTACCTG
 CGGAGTTTAGTTTGCCTACGCATCACTGTGCGCTAATTGCGGTACAGCATCACCTCAGCATGATGAGCTGGCATT
 CTTGAGACATCCAGTGTTGAATTGGAAGGATGTGTGGTGAATGGCACTGATTGATATTGTGCCTTAAGTACCTATT
 CAATTAGGGCGACCGTATGGGGTAATTTAATTGGCGTGAACCATGCGGCCGAAAT

Fig. 1G-1

>V7-Nsp2d324-434.seq

ATGACGTATAGGTGTTGGCTCTATGCCTTGGCATTGTATTGTGTCAGGAGCTGTGACCATTGGCACAGCCCAAACTTGCT
 GCACAGAAACACCTTCTGTGATAGCCTCCTTCAGGGGAGCTTAGGGTTTGTCCCTAGCACCTTGCTTCCGGAGTTGCAC
 TGCTTTACGGTCTCTCCACCCCTTAACCATGTCTGGGATACTTGATCGGTGCACGTGTACCCCAATGCCAGGGTGT
 ATGGCGGAGGGCCAAGTCTAETGCACACGATGCCTCAGTGACGGTCTCTCCTCCCTGAACCTCCAGGTTTCTGAGCT
 CGGGGTGCTAGGCCTATTCTACAGGCCGAAGAGCCACTCCGGTGGACGTTGCCACGTGCATTCCCCTGTTGAGTGCT
 CCCCCGCGGGGCTGCTGGCTTCTGCAATCTTCCAATCGCACGAATGACCAGTGGAACCTGAACCTCCAACAAAGA
 ATGGTACGGGTGCGAGCTGAGCTTTACAGAGCCGGCCAGCTCACCTCGAGTCTTGAAGGCTCTACAAGTTTATGAACG
 GGGTTGCCGCTGGTACCCATTGTTGGACCTGTCCCTGGAGTGGCCGTTTTTCGCCAATTCCCTACATGTGAGTGATAAAC
 CTTTCCCGGAGCAACTCACGTGTTGACCAACCTGCCGCTCCCGCAGAGACCAAGCCTGAAGACTTTTGGCCCTTGAG
 TGTGCTATGGCTACTGTCTATGACATTGGTCATGACGCCGTGATGTATGTGGCCGAAAGGAAAGTCTCCTGGGCCCTCG
 TGGCGGGGATGAAGTGAATTTGAAGCTGTCCCGGGGAGTTGAAGTTGATTGCGAACCGGCTCCGCACCTCCTTCCCGC
 CCCACCACACAGTGGACATGTCTAAGTTCGCTTACAGCCCTGGGTGTGGTGTCTATCGGGTGAACGCCAACAC
 GGCTGCCTTCCCGTGACACTGTCCCTGAAGGCACTGCTGGTGGAGCTTGTGACTTGCTTCCACTGGAAGTTGAGAA
 CAAAGAAATTCGCTATGCTAACCAATTTGGCTACAGACCAAGCATGGTGTCTCTGCAAGTACCTGCAGCGGAGGCTGC
 AAGTTAATGGTCTCCGAGCAGTAACCTGACCTAAACGGACCTATCGTCGTACAGTACTTCTCCGTTAAGGAGAGTTGGATC
 CGCCATTTGAAACTGGCGGGAGAACCAGCTACTCTGGGTTTGAGGACCTCCTCAGAATAAGGGTTGAGCCTAACACGTC
 GCCATTGGCTGACAAGGAAGAAAAATTTCCGGTTTGGCAGTCACAAGTGGTACGGCGCTGGAAGAGAGCAAGAAAAAG
 CACGCTCTTGTGCGACTGCTACAGTCGCTGGCCGCGCTTGTCCGTTCTGTAAGCCGCGCAGGCCAAGGAGCAGAGGTT
 GCGGCGCCAAAGGCTGAGCACCTCAAACTACTCCCCGCTGCCGAAGGAATTGTGGTTGGCACTGCAATTCGCG
 CATCGCAACCGGATGGTGAATTCAAATTTGAAACCACTTCCCGAAAGAGTGAGACCTCCAGATGACTGGGCTACTG
 ACGAGGATCTTGTGAATGCCATCCAAATCCTCAGACTCCCTGCGGCTTAGACAGGAACGGTGTGACTAGCGCAAG
 TACGTACTTAAGCTGGAAGGTGAGCATTGGACTGTCACTGTGACCCCTGGGATGTCCCTTCTTGTCTCCTCTTGAATG
 TGTTAGGGCTGTTGTGGGCACAAGGGCGGTCTGGTTCCCGAGATGACGTGAGGTCTCCGGATTTGACCTGCCTGCC
 TTGACCGGCTGGCTGAGGTGATGCACCTGCCTAGCAGTGCTATCCAGCCGCTCTGGCCGAAATGTCTGGCGATTCCGAT
 CGTTCCGCTTCTCCGTCACACCGTGTGGACTGTTTTCGAGTTCTTTGCCCGTCACAGCGGAGGGAATCACCTGACCA
 AGTGGCTTAGGGAAATATCAGCCTTGTGAGGTGATTGAGGACTGCTGCTGTTCCAGAACAACCAACCGGGTCA
 CCCCAGGAGGTCGAGCAAGATTGACCTGTACCTCCGTGGTGCAAAATCTTGAAGAATGCTTGGCCAGGCTTGAG
 AAAGCGCGCCGCCACGCGTAATCGACACCTCCTTGTGGGATGTTGTCTCCCTGGGGTTGAGGCGGCAACCCAGAC
 GATCAAGCTGCCCCAGGTCAACAGTGTCTGCTGCTGGTCCCTGTTGTGACTCAAAAGTCTTGCCTTGAAGCTTCAAGCTC
 GAAAAACGAAGCTGTCAAGAGCTTCCCGGAGAGAAAGCTTCTCCCGCCCGCGCAGGAAGGTTGGGTCCGATTGTGGC
 AGCCCCGTTTCTATTAGGCGGCGATGTCCCTAACAGTTGGGAAGATTTGGCTGTAGTAGCCCCCTTGATCTCCGACCCC
 ACCTGAGCCGCGCAACCTTCAAGTGAAGTGGTATTGTCTCACCAGCAATGCATCTTCAAGCCGCGACACCTTGA
 GTGAGCCGGCTCAATTTCCCGCACCTCGCGGAAGTGTCTCGACCGGTGACACCTTGAAGTGAAGCCGATCCCTGTGCC
 GCACCGCGGCTAAGTTTCAAGAGTGAAGAGATTGAGTTCCGGCGCGCAATCCACCGTACCAGGACGAGCCCTGGA
 TTTGTCTGCTTCTCACAGACTGAATATGAGGCTCTCCCCAGCACCGCGCAGAGCGGGGGCTTCTGGGAGTAGAGG
 GGCATGAAGCTGAGGAAACCTGAGTGAATCTCGGACATGTCCGGTAACATTAAACCTGCGTCCGTGTATCAAGCAGC
 TCCTTGTCCAGCGTGAGAATCACACGCCAAAATACTCAGCTCAAGCCATCATCGACTCGGGCGGCTTCAAGTGGGA
 TCTCCAAGAGGTAAGGAAACATGCCTTAGTGTGATGCGCGAGGATGTGATGCGACTAAGCTTGATGACCTGTACGC
 AGGAATGGCTTCTCGCATGTGGGATCGGGTGACATGCTGACTTGGCGCAACACGTCTGTTTACCAGGCGATTGCAAC
 TTAGATGGCAGGTTAAAGTCTCCCAAAATGATACTCGAGACACCGCGCCCTATCCGTGTGAGTTGTGATGATGCC
 TCACACGCTGCACCTTCCGTAGGTGCGGAGAGCAGCTTACCATTGGCTCAGTTGCTACTGAAGATGTTCCACGCATCC
 TCGAGAAAATAGAAAATGTCCGGGAGATGGCCAACCAAGGACCTTGGCTTCTCCGAGGATAAACCGGTAGATGACCA
 CTTGTCAACGACCCCGGATATCGTCGCGGAGGCTGACGAGAGCACATCAGCTCCGTCCGAGGACAGGTGGCGCCGG
 CTCTTTTACCGATTGCGGCTTCAAGTGGCGGATGCGGACGGGGGGGGCGTTTCCGACGGTAAAAAGAAAGCTG
 AAAGGCTCTTGACCACTGAGCCGTGAGTTTGTGACCTCGTCTCCATCTCCCTGTTTTCTTCTACGCTTTTCTAC
 CCTGGCGGTGTTATTCTCCGGGTGATTGGGGTTTGCAGCTTTTACTCTATTGTGCTCTTTTATGTTACAGTTACCC
 AGCCTTTGGTATTGCTCCCTCTTGGGTGTGTTTCTGGGTCTTCTCGGCGGTTGGAATGGGGTTTTTGGCTGCTGGT
 TGGCTTTGCTGTTGGTCTGTTCAAGCCTGTGTCCGACCCAGTCGGCGCTGCTGTGAGTTTGAAGTCCGAGAGTGTAGA
 AACATCCTTCATTCTTTGAGCTTCTCAACCTTGGGACCTGTTGCGAGCTTGTGTGGGCCCGTCCGTCTCGGTCT

Fig. 1G-2

TGCCATTCTTGGCAGGTTACTGGGCGGGGCACGCTGCATCTGGCACTTTTGGCTTAGGCTTGGCATTGTTGCAGACTGTA
 TCTTGGCTGGAGCTTACGTGCTTTCTCAAGGTAGGTGTAAAAAGTGTGGGGATCTTGATAAGAACTGCTCCCAATGAG
 GTCGCTTTTAACGTGTTTCTTTTACACGTGCGACCGAGTCTGCACTTATCGACCTGTGCGATCGGTTTTGTGCGCCAAA
 AGGAATGGACCCCATTTTTCTCGCCACTGGGTGGCGCGGGTGTGGGCCGGCCGAAGCCCATTTAGCAACCCCTCTGAAA
 AACCCATCGCGTTTGGCCAGTTGGATGAAAAGAAGATTACGGCTAGGACTGTGGTCGCCAGCCTTATGACCCCAACCAA
 GCCGTAAAGTGCTTGGCGGTATTGACGGCGGGTGGGGCGATGGTGGCTAAGGCGGTCCAAAAAGTGGTCAAGGTTTCCGC
 GTTCCATTCCGAGCCCCCTCTTTCCCACTGGAGTGAAAGTTGACCCGATTGACGGGTGCTGGTGACCCGACACTT
 TCACTGCAGCTCTCCGGTCTGGCTACTCCACCACAAACCTCGTCTTGGTGTGGGGGACTTTGCCAGCTGAATGGATTA
 AAAATCAGGCAATTTCCAAGCCTTCAAGGGGAGGCCACATCTCATGGCTGCCCTGCATGTTGCCCTGCTCGATGGCTCT
 GCACATGCTTGTGGGATTATGTGACTGCGGTGGGTTCTTGGCGACCGGCACCAACGACCCGTGGTGGCTAACCCTG
 TTGCCGTCCCTGGCTACGGACCTGGCTCTCTGACAGTCCAGATTGTGCATTTCCCAACACGGCCTTACCCCTGCCCTTG
 ACAGCACTTGTGGCGGGATTCCGTTATTCAAGAAATTGCCCTTGGTCGTTTTGATTTTTGTTTCCATCGGAGGCATGGCTCA
 TAGGTTGAGCTGTAAAGCTGACATGCTGTGCTTGGTGTGCAATTGCCAGCTATGTTTGGGTACCTTACCTGGTTGC
 TTTGTGTGTTTCTTGTGCTGGTGGCTGTTTTCTTGTGACCCCTCACCATCCTATGGTGGTGTGTTTTCTTGATTTCT
 GTGAATATGCCTTCAGGAATCTTGGCCATGGTGTGTTGGTGTCTTGTGGCTTCTTGGTGTGTTACTAATGTTGCTGG
 CCTGTGACCCCTACGACATTCATCATTACACAGTGGCCCCCGGTGTTGCCGCTTGGCTACCGCACCAGATGGGA
 CCTACTTGGCGCTGTCCGCCGCTGCGTTGACTGGCCGACCATGCTGTTTACCCCGTCCAGCTTGGGTCTCTTCTT
 GAGGGTGCTTTCAGAACTCGAAAGCCCTCACTGAACACCGTCAATGTGATCGGGTCTCCATGGGCTCTGGCGGGGTGT
 TACCATCGACGGGAAAGTCAAGTGCCTAAGTGCACATGCTCTTACGGGCAATTGAGCTCGGGTTTCCGGGGTCTGGCT
 TCAATCAAATGCTTGACTTTGACGTAAAGGGAGATTTGCTATAGCTGATTGCCCGAATTGGCAAGGGGCTGCCCAAG
 ACCCAATTCTGCACGGATGGATGGACTGGCCGTGCCTATTGGCTAACATCCTCTGGCGTCAACCCGGCGTATTGAAA
 AGGATTCGCTTCTGCTTACCGCATGTGGCGATTCCGGGTCCCAAGTATCACCAGGCGGGTGGCTGTGCGGCTTC
 ACACGGGATCGAATAAACAAGGGGGGGCATTGTACGCGCCCTCAGGCCAGTTTTGTAATGTGGCACCACATCAAGCTA
 AGCGAATTAAGTGAATTTTGTGCTGGGCTAAGGTCCCGTCTGGTGATGTGAAGGTGGCAGCCACATAATTAAGACAT
 AAGCGAGGTGCCTTCAAGTCTTGTGCTTGTGCTGCTGCAAACTGAACTGGAAGGAGGCTCTCCACCGTCAACTTC
 TTTGTGTGTTTTCTCCTGTGGAGAATGATGGGACATGCTGGACGCCCTTGGTTGCTGTGAGTTTCTTATTTGAAT
 GAGGTTCTCCAGCGCTCTGGTCCGGAGTGTCTTCTCCTTGGAAATGTTGTGCTATCCTGGCTCACGCCATGGTCTGC
 GCAAGTTCTGATGATCAGGCTTCTGACAGCAGCTCTTAACAGGAACAGATGGTCACTTGCCTTTTTCAGCCTCGGTGCAG
 TGACCGGTTTTGTGCGAGATCTTGGGCCACTCAGGGGATCCGTTGCAGGCAGTGATGAATTTGAGCACCTATGCATTC
 CTGCTCGGATGATGGTTGTGACCTCACCAGTCCAGTGATCAGGTGGTGTGCTGCACCTACTTGCATCATTGTA
 CTTGTTAAGTACCGTGGCTGCACCATATCCTGTTGGCGATGGAGTGTCTGCGGCTTTCTTCTGAGATACTTTG
 CCGAGGGAAAGTTGAGGGAAGGGGTGTGCAATCCTGCGGAATGAATCATGAGTCTCTGACTGGTGCCTCGCTATGAGA
 CTCAATGACGAGGACTTGGATTTCTTATGAAATGGAGTATTTAAGTGTCTTGTCTGCGTCCAACATGAGGAATGC
 AGCGGGTCAATTTATCGAGGCTGCCTATGCTAAAGCACTTAGAGTAGAACTGGCCAGTTGGTGCAGGTTGATAAAGTTC
 GAGGTACTTTGGCCAAACTTGAAGCTTTTGTGATACCGTGGCACCTCAACTCTCGCCCGGTGACATTGTTGTGCTCTC
 GGCCACACGCTGTTGGCAGTATCTTGCACCTAAAGGTTGGTAGCACCAGCATACCTCCAAGCCATTGAGACCAGAGT
 CTTGCTGGGTCCAAATGACCGTGGCGCGCTGTCGACCCGACCCACGCCCCACCCGACCCGTCGCCATCCCC
 TCCACCGAAAGTTCTGGAGAATGGCCCCAACGCTTGGGGGATGAGGACCGTTGAATAAGAAGAAGAGGCGCAGGATG
 GAAGCCCTCGGCATCTATGTTATGGGCGGAAAAAATACCAGAAATTTGGGACAAGAATTCGGTGATGTGTTTTATGA
 GGAGGTCCATAATAACACAGATGAGTGGGAGTGTCTCAGAGTTGGCGACCCCTGCCGACTTTGACCCCTGAGAAGGGAATC
 TGTGTGGACATGTACCATTTGAAAAAAGGCTTACCATTGTTTACACCTCCCATCTGGTAAGAAGTTCTTGGTCCCGTC
 AACCCAGAGAATGGAAGAGTCCAATGGGAAGCTGCAAGCTTTCCGTGGAGCAGGCCCTAGGTATGATGAATGTCGACGG
 CGAACTGACTGCCAAAGAACTGGAGAACTGAAAAGAATAATTGACAAATCCAGGGCTGACTAAGGAGCAGTGTAA
 ACTGCTAGCCGCCAGCGACTTGACCCGCTGTGGTGGCGGGGCTTGGTGTGTTACTGAAACAGCGGTAAAAATAGTCAAAT
 TTCACAACCGGACCTTACCCTGGGACCTGTGAATTTAAAGTGGCCAGTGAGGTTGAGCTAAAGACGCGGTTGAGCAC
 AACCAACACCCGTTGCGAGACCGATCGATGGTGGAGTTGTGCTCTGCTTCCGCGGTTCTTCTGCTTATAGACGCTCT
 GATCTCCGGTGTGATGCATCTCCAAGTTACTTGCCCATCAGGGCGGGAAACACTGGGATCGATGGCACGCTCTGGG
 ATTTTGTAGTCCGAAGCCACTAAAGAGGAAGTGCAGTCACTGCGCAATAATACAGGCTTGTGACATTAGGCGCGGCGAC
 GCTCCTGAAATGGTCTCCCTTACAAGCTGTACCCTGTTAGGGGTAAACCCTGAGCGGGTGAAGGAGTCTGCAGAATAC
 AAGGTTTGGAGACATACCTTACAAAACCCCAAGTACACTGGAAGCCAGTGACGCGGCTGCCTGCCTTACGCCCAACG

Fig. 1G-3

CCACTCCGGTGACTGATGGGCGCTCCGTCTTGGCCACGACCATGCCCCCGGGTTTGAGTTATATGTACCGACCATACCA
 GCGTCTGTCTTGATTACCTTGACTCTAGGCGTGACTGCCCTAAACAGCTGACAGAGCACGGCTGCGAAGATGCCGCACT
 GAAAGACCTCTCTAAATATGACTTGTCCACCAAGGCTTTGTTTTACCTGGAGTTCTTCGCTTGTCGCGAAATACCTGT
 TTGCCCATGTAGGTAAGTGCCACCCGTTTCATCGGCTTCTACTTACCCTGCTAAGAATTCTATGGCTGGAATAAATGGG
 AACAGGTTCCCAACCAAGGACATTAGAGCGTCCCTGAAATCGACGTTCTGTGCGCACAGGCTGTGCGAGAAACTGGCA
 AACTGTCACCCCTTGACTCTTAAGAAACAGTATTGCGGGAAGAAGAAGACTAGGACCATACTCGGCACCAATAACTTCA
 TCGCACTAGCCACCGAGCAGTGTTGAGTGGTGTACCCAGGGCTTCATGAAAAAGGCGTTAACTCGCCCATCGCCCTC
 GGAAAGAACAAGTTTAAGGAGCTACAGACTCCGGTCTTGGGCGAGTGCTTGAAGCTGATCTCGCATCCTGCGATCGATC
 CACGCTGCAATTGTCCGCTGGTTTGCCGCCAACCTTCTTATGAAGTTGCTGTGCTGAAGAGCATCTACCGTCGTACG
 TGCTGAAGTGTGCCAGACTTACTGGTCACGAGTCCGGCGCAGTACTAAGAGAGGTGGCTGTGCTGTGGCGACCG
 ATCAGCTCTGTCTAAGACCACTTATAGTTTGGTGATCTATGCACAGCATATGGTGCTTAGTTACTTCAAAAGTGGTCA
 CCCCCATGGCCTTCTGTTCTTACAAGACCAGCTAAAGTTTGAGGACATGCTCAAGGTTCAACCCCTGATCGTCTATTCCG
 ACGACCTCGTGCTGTATGCCGAGTCTCCACCATGCCAACTATCACTGGTGGGTTGAACATCTGAATTTGATGCTGGG
 TTTGAGACGGACCCAAAGAAGACAGCAATAACAGACTCGCCATCATTTCTAGGCTGTAGAATAATAAATGGGCGCCAGCT
 AGTCCCCAACCGTGACAGGATCCTCGCGGCCCTCGCTATCACATGAAGGCGAGTAATGTTTCTGAATACTATGCCTCAG
 CGGCTGCAATACTCATGGACAGCTGTGCTTGTGGAGTATGATCTGAATGGTTTGAAGAACTTGAGTTGGAATAGCG
 CAGTGCGCCCGCAAGGACGGCTACAGCTTCCCGGCACGCCGTTCTTCATGTCCATGTGGGAAAACTCAGGTCCAATTA
 TGAGGGGAAGAAGTCGAGAGTGTGCGGGTACTGCGGGGCCCGGCCCGTACGCTACTGCTGTGGCCTCGAGCTGTGCA
 TTTACCACACCCACTTCCACCAGCATTGTCCAGTCACAATCTGGTGTGGCCATCCAGCGGGTCTGGTCTTGTAGTGAG
 TGCAAATCCCCTGTAGGGAAAGGCACAAGCCCTTAGACGAGGTGCTGGAACAAGTCCCGTATAAGCCCCACGGACCGT
 TATCATGCATGTGGAGCAGGGTCTACCCCCCTTGATCCAGGTAGATACCAAACCTCGCCGCGGATTAGTCTCTGTAGGC
 GTGGAATTAGGGGAAATGAAGTTGGACTACCAGACGGTGATTATGCTAGCACCGCCTTGCTCCCTACCTGCAAAGAGATC
 AACATGGTGTGCTGTGCTTCCAATGTATTGCGCAGCAGGTTTCATCATCGGCCACCCGGTGCTGGGAAAAACATACTGGCT
 CCTTCAACAGGTCCAGGATGGTGATTTATTACACCAACTCACCAGACCATGCTTGACATGATTAGGGCTTTGGGGA
 CGTGCCGGTTCAACGTCCCGGCAGGCACAACGCTGCAATTTCCCGTCCCTCCCGCACCGGTCCGTGGGTTGCGATCCTA
 GCCGGCGGTTGGTGCTTGGAAGAATTCTTCTAGATGAAGCAGCGTATTGCAATCACCTTGATGTTTTGAGGCTTCT
 TAGTAAACTACCTCACCTGTCTAGGAGACTTCAAGCAACTCCACCCAGTGGGTTTGATTCTCATTGCTATGTTTTG
 ACATCATGCCTCAAACCTCAACTGAAGACCATCTGGAGGTTTGACAGAAATATCTGTGATGCCATTAGCCAGATTACAGG
 GACAACTCATGTCCATGGTCAACACAACCCGTGTGACCTACGTGGAAAACTGTGAGGTATGGGCGAGTCTCACCCC
 CTACCACAGGACCGAGAGGACGACGCCATCACTATTGACTCCAGTCAAGGCGCCACATTGATGTGGTTACATTGCATT
 TGCCCACTAAAGATTCACTCAACAGGCAAGAGCCCTTGTGCTATCACCAGGGCAAGACAGCTATCTTTGTGATGAC
 CCACACAGGACAGTGCAGGGCTTGTGTTGATCTTCTGCAAAAGGCACGCCGTCAACCTCGCAGTGCAGTGCAGCGGCA
 GCTGATCGTGTGGATGAGAAATAACAAAGATGCACGGTTGCTCAGGCTTAGGCAACGGGGATAAAATTAGGGCCACAG
 ACAAGCGTGTGTAGATTCTCTCCGCGCATTGTGCTGATCTAGAAGGGTCGAGCTCTCCGCTCCCCAAGTGCACAC
 AACTTGGGATTTTATTTCTCACCTGATTTAACACAGTTTGCTAAACTCCAGTAGAACTTGACACTACTGGCCCGTGGT
 GTCAACCCAGAACAATGAAAAGTGGCCGATCGGCTGGTTGCCAGCCTTCGCCCTATCCATAAATACAGCCGCGCGTGCA
 TCGGTGCCGCTATATGGTGGGCCCTTCGGTGTCTAGGCACTCCTGGGTCGTGTACTACTATCTCACAATAATTTGTT
 AAGGGCGGGGCTCAAGTGCTTCCGGAGACGGTTTTAGCACCGGCCGAATTGAGGTAGACTGCCGGGAATATCTTGATGA
 TCGGGAGCGAGAAGTTGCTGCGTCCCTCCACACGCTTTCATTGGCGACGTCAAAGGCACTACCGTTGGAGGATGTATC
 ATGTACCTCCAGATACCTCCCGCGCTCCTCCCAAGGAATCAGTTGCGGTAGTGGGGTTTCAAGCCCCGAAAGGCC
 GCGAAAGCATTGTGCACACTGACAGATGTGTACCTCCAGATCTTGAAGCCTATCTCCACCGGAGACCCAGTCCAAGTG
 CTGGAAAATGATGTTGGACTTCAAAGAAGTTCGACTAATGGTCTGGAAAGACAAAACAGCCTATTTCAAACCTGAAGGTC
 GCTATTTACCTGGTATCAGCTTGCAGCTATGCCTCGTACATCCGTGTTCCCGTCAACTCTACGGTGTACTTGGACCCC
 TGCATGGGCCCCGCCCTTGTGAACAGGAGAGTCGTGGGTCCACCCACTGGGGGCTGACCTCGCGGTACCCCTTATGA
 TTACGGCGCTAAAATTATCTGTCTAGCGCTACCATGGTGAAATGCCCCCGGATACAAAATTCTGGCGTGCAGGAGT
 TCTCGTTGGATGACCCAGTTAAGTACAAACATACCTGGGGGTTTGAATCGGATACAGCGTATCTGTATGAGTTCACCGGA
 AACGGTGAGGACTGGGAGGATTACAAATGATGCGTTTCTGCGCGCCAGGAAGGGAAAATTTATAAGGCCACTGCCACCAG
 CTTGAAGTTTTATTTCCCCCGGGCCCTGTCAATTGAACCAACTTTAGGCCTGAATTGAAATGAAATGGGGTCCATGCAA
 GCCTTTTTGACAAAATTGGCCAACTTTTGTGGATGCTTTCACGGAGTCTTGGTGTCCATTGTTGATATCATTATATTT
 TTGGCCATTTTGTGGCTTACCATCGCCGGTGGCTGGTGGTCTTTGTCATCAGATTGGTTTGTCCGCGATACTCCG

Fig. 1G-4

TACGCGCCCTGCCATTACTCTGAGCAATTACAGAAGATCTTATGAGGCCTTTCTTTCCAGTGCCAAGTGGACATTCCC
 ACCTGGGGAACATAACATCCTTTGGGGATGCTTTGGCACCATAAGGTGTCAACCCTGATTGATGAAATGGTGTGCGGTG
 AATGTACCGCATCATGGAAAAAGCAGGGCAGGCTGCCTGGAAACAGGTGGTGAAGCGAGGCTACGCTGTCTCGCATTAGTA
 GTTTGGATGTGGTGGCTCATTTTCAGCATCTAGCCGCCATTGAAGCCGAGACCTGTAAATATTTGGCCTCCCGGCTGCCC
 ATGTACACAACCTGCGCATGACAGGGTCAAATGTAACCATAGTGTATAATAGCACTTTGAATCAGGTGTTTGCTATTTT
 TCCAACCCCTGGTTCCCGCCAAAGCTTCATGATTTTCAGCAATGGTTAATAGCTGTACATTCTCCATATTTCTCTG
 TTGCAGCTTCTGTACTCTTTTGTGTGCTGTGGTTGCGGGTTCCAATACTACGTACTGTTTTGGTTTCCGCTGGTTA
 GGGGCAATTTTCTTTGAACTCACAGTGAATTACACGGTGTGTCCACCTGCCTCACCCGGCAAGCAGCCACAGAGATC
 TACGAACCCGGTAGGTCTCTTTGGTGCAGGATAGGGTATGACCGATGTGGGGAGGACGATCATGACGAGCTAGGGTTAT
 GATACCGCTGGCTCTCCAGCGAAGGCCACTTGACTGGTGTTCACGCTGGTTGGCGTTCTTGCTCTCAGCTACACGG
 CCCAGTTCATCCGAGATATTCGGGATAGGGAATGTGAGTCGAGTTTATGTTGACATCAAACATCAACTCATCTGCGCC
 GAACATGACGGGCAGAACACCACCTTGCTCGTCATGACAACATTTAGCCGTGTTTCAGACCTATTACCAACATCAAGT
 CGACGGCGCAATTGGTTTCACTAGAATGGCTTCCTCTCTCTGTTGGTTTAAATGTCTCTTGGTTTC
 TCAGGCGTTCGCTGCAAAACCATGTTTCAGTTCGAGTCTTGAGATTAAGACCAACACCACCGCAGCGGCAAGCTTTG
 CTGCTCTCAAGACATCAGTTGCCCTAGGCATCGCGACTCGGCCTCTGAGGCGATTGCAAAATCCCTCAGTGCCGTACG
 GCGATAGGGACACCCGTGTATGTTACCATCACAGCAATGTGACAGATGAGAATTATTTACATTCTTCTGATCTCCTCAT
 GCTTTCTTCTGCTTTCTATGCTTCTGAGATGAGTGAAAAGGGATTAAAGTGTTTGGCAATGTGTCAGGCATCG
 TGGCTGTGTGTCAATTTTACCAGCTACGTCCAACATGTCAAGGAGTTTACCCAACGCTCCCTGGTGGTGCACCATGTG
 CGGTGCTCCATTTTATGACACCTGAGACCATGAGGTGGGCAACTGTTTTAGCCTGTCTTTTGCCATTCTGTTGGCAAT
 TTGAATGTTTAAAGTATGTTGGAGAAATGCTTGACCGCGGGCTGTTGCTCGCGATTGCTTTCTTTGTGGTGTATCGTCCG
 TTCTGTTTTGCTGTGCTCGCCAACGCCAGCAACGACAGCAGCTCCCATCTACAGCTGATTTACAACCTGACGCTATGTGA
 GCTGAATGGCACAGATTGGCTAGCTAACAAATTTGATTGGGCAAGTGGAGAGTTTGTGATCTTTCCCGTTTGTACTCACA
 TTGTCTCCTATGGTGCCTCACTACCAGCCATTTCTTGACACAGTCGCTTAGTCACTGTGTCTACCGCCGGGTTTGT
 CACGGGCGGTATGCTTAAGTAGCATCTACGCGGTCTGTGCCCTGGCTGCGTTGACTTGCTTCGTATTAGGTTTGCAA
 GAATTGCATGTCTGGCGCTACGCGTGTACCAGATATACCAACTTTCTTCTGGACACTAAGGGCAGACTCTATCGTTGGC
 GGTGCGCTGTATCATAGAGAAAAGGGGCAAAGTTGAGGTGGAAGGTCTGATCGACCTCAAAAGAGTTGTGCTTGAT
 GGCTCCGTGGCAACCCCTATAACCAGAGTTTCAGCGGAACAATGGGGTCTGCTTAGATGACTTCTGTACGATAGCAG
 GCTCCACAAAAGGTGCTTTTGGCGTTTCTATTACCTACACGCCAGTGATGATATATGCCCTAAAGGTGAGTCGCGGCCG
 ACTGCTAGGGCTTCTGCACCTTTTGATCTTCTGAATGTGCTTTACCTTCGGGTACATGACTTTCGCGCACTTTCAGA
 GTACAAATAAGTTCGCGCTCACTATGGGAGCAGTAGTTGCACTCTTTGGGGGTGTACTCAGCCATAGAAACCTGGAAA
 TTCATCACCTCCAGATGCCGTTTGTGCTTGTAGGCGCAAGTACATTCTGGCCCTGCCACACGTTGAAAGTGCCGC
 AGGCTTTTCATCCGATTGCGGCAAATGATAACCAAGCATTTGTGCTCGGCGTCCCGGCTCCACTACGGTCAACGGCACAT
 TGGTGCCCGGTTAAAAAGCTCGTGTGGGTGGCAGAAAAGCTGTAAACAGGGAGTGGTAAACCTTGCAAAATATGCC
 AAATAACAACGGCAAGCAGCAGAGAGAGAAAGAGGGGGATGGCCAGCCAGTCAATCAGCTGTGCCAGATGCTGGGTAGA
 TCATCGCTCAGCAAAACAGTCCAGAGGCAAGGGACCGGAAAGAAAAATAAGAAGAAAAACCGGAGAAGCCCCATTTT
 CCTAGCGACTGAAGATGATGTACACATCACTTTACCCCTAGTGAGCGGCAATTGTGTCTGCTCAATCCAGACCGC
 CTTTAATCAAGGCGCTGGGACTTGACCCCTGTGAGATTAGGGAGGATAAGTTACACTGTGGAGTTTAGTTTGCCTACGC
 ATCATACTGTGCGCTGATCCGCTCACAGCATCACCTCAGCATGATGGGCTGGCATTCTTGAGGCATCTCAGTGTTTG
 AATTGGAAGAATGTGTGGTGAATGGCACTGATTGACATTGTGCTCTAAGTCACTATTCAATTAGGGCGACCGTGTGGG
 GGTGAGATTTAATTGGCGAGAACCATGCGGCCGAAATTAATAAAAAA

Fig. 1H-1

>V7-Nsp2d324-523.seq

ATGACGTATAGGTGTTGGCTCTATGCCTTGGCATTGTATTGTGAGGAGCTGTGACCATTGGCACAGCCCAAACTTGCT
 GCACAGAAACACCCCTTCTGTGATAGCCTCCTTCAGGGGAGCTTAGGGTTTGTCCCTAGCACCTTGCTTCCGGAGTTGCAC
 TGCTTTACGGTCTCTCCACCCCTTAACCATGTCTGGGATACTTGATCGGTGCACGTGTACCCCAATGCCAGGGTGT
 ATGGCGGAGGGCCAAGTCTACTGCACACGATGCCTCAGTGCACGGTCTCTCCTTCCCCTGAACCTCCAGGTTTCTGAGCT
 CGGGGTGCTAGGCTATTCTACAGGCCCCAAGAGCCACTCCGGTGGACGTTGCCACGTGCATTCCCCACTGTTGAGTGCT
 CCCCCGCCGGGGCTGCTGGCTTTCTGCAATCTTTCCAATCGCACGAATGACCAGTGGAAACCTGAACCTCCAACAAAGA
 ATGGTACGGGTGCGAGCTGAGCTTTACAGAGCCGGCCAGCTCACCCCTGCAGTCTTGAAGGCTCTACAAGTTTATGAACG
 GGGTTGCCGCTGGTACCCCATTTGTTGGACCTGTCCCTGGAGTGGCCGTTTTCCGCAATTCCCTACATGTGAGTGATAAAC
 CTTCCCGGGAGCAACTCACGTGTTGACCAACCTGCCGCTCCCGCAGAGACCCAAGCCTGAAGACTTTTGCCCTTTGAG
 TGTGCTATGGTACTGTCTATGACATTGGTCATGACGCCGTGATGTATGTGGCCGAAAGGAAAGTCTCCTGGGCCCTCG
 TGGCGGGGATGAAGTGAATTTGAAGCTGTCCCGGGGAGTTGAAGTTGATTGCGAACCGGCTCCGCACCTCCTTCCCGC
 CCCACCACACAGTGGACATGTCTAAGTTCGCTTACAGCCCTGGGTGTGGTGTCTATGCGGGTGAACGCCAACAC
 GGCTGCCCTTCCGCTGACACTGCCCTGAAGGCAACTGCTGGTGGAGCTTGTGACTTGTCTTCACTGGAAGTTCAGAA
 CAAAGAAATTCGCCATGCTAACCAATTTGGCTACCAGACCAAGCATGGTGTCTCTGGCAAGTACCTGCAGCGGAGGCTGC
 AAGTTAATGGTCTCCGAGCAGTAACTGACCTAAACGGACCTATCGTCGTACAGTACTTCTCCGTTAAGGAGAGTTGGATC
 CGCCATTTGAACTGGCGGGGAGAACCCAGCTACTCTGGGTTTGAAGACCTCCTCAGAATAAGGGTTGAGCCTAACACGTC
 GCCATTGGCTGACAAGGAAGAAAAATTTCCGGTTTGGCAGTCACAAGTGGTACGGCGCTGGAAGAGAGCAAGAAAAAG
 CACGCTCTTGTGCGACTGCTACAGTCGCTGGCCGCGCTTTGTCGTTCTGTAAACCCGGCAGGCCAAGGAGCAGGAGTT
 GCCGGCGCAACAAGGCTGAGCACCTCAAACACTACTCCCCGCTGCCGAAGGGAATTGTGGTTGGCACTGCATTTCCGC
 CATCGCAACCGGATGGTGAATTCCAAATTTGAAACCAACCTTCCCGAAAGAGTGAGACCTCCAGATGACTGGGCTACTG
 ACGAGGATCTTGTGAATGCCATCCAAATCCTCAGACTCCCTGCGGCCTTAGACAGGAACGGTGCTGTACTAGCGCCAAG
 TACGTACTTAAGCTGGAAGGTGAGCATTGGACTGTCACTGTGACCCCTGGGATGTCCCTTCTTGTCCCTCTTGAATG
 TGTTACGGGCTGTTGTGGGCACAAGGGCGGTCTTGGTTCCTCAGATGCAGTCGAGGTCTCCGGATTTGACCTGCCTGCC
 TTGACCGGCTGGCTGAGGTGATGCACCTGCCTAGCAGTGTATCCAGCCGCTCTGGCCGAAATGTCTGGCGATTCCGAT
 CGTTCCGCTTCTCCGGTCACCACCGTGTGGACTGTTTCCGAGTTCTTGGCCGTACAGCGGAGGGAATCACCTGACCA
 AGTGCCTTAGGGAAAATTTATCAGCCTTGTGAGGTGATTGAGGACTGCTGCTGTTCCAGAACAAACCAACCGGGTCA
 CCCCCGAGGAGGTGCGAGCAAGATTGACCTGTACCTCCGTGGTGAACAAATCTTGAAGAATGCTTGGCCAGGCTTGAG
 AAAGCGCGCCGCCACGCGTAATCGACACCTCCTTTGATTGGGATGTTGTGCTCCCTGGGGTTGAGGCGGCAACCCAGAC
 GATCAAGCTGCCCCAGGTCAACCAAGTGTGCTGCTGGTCCCTGTTGTGACTCAAAAGTCTTGCCAATTCGCGACCTC
 GCGGAATGTGCTCGACCGGTGACACCTTGTAGTGAGCCGATCCCTGTGCCCGCACCGCGCGTAAGTTTCAGCAGGTG
 AAAAGATTGAGTTCGGCGGCGGCAATCCACCGTACCAGGACGAGCCCTGGATTTGTCTGCTTCTCACAGACTGAATA
 TGAGGCTCTCCCCAGCACCGCCGAGAGCGGGGGCGTTCTGGAGTAGAGGGGCATGAAGCTGAGGAAACCTGAGTG
 AAATCTCGGACATGTGGGTAACATTAACCTGCGTCCGTGTCAAGCAGCTCCTTGTCCAGCGTGAGAATCACACGC
 CCAAATACTCAGCTCAAGCCATCATCGACTCGGGCGGGCCCTGCAGTGGGCATCTCAAGAGGTAAAGGAAACATGCCT
 TAGTGTATGCGCGAGGCATGTGATGCGACTAAGCTTGATGACCTGCTACGCAGGAATGGCTTTCTCGCATGTGGGATC
 GGGTGGACATGCTGACTTGGCGCAACAGTCTGTTTACCAGGCGATTGACCTTAGATGGCAGGTTAAAGTTCTCCCA
 AAAATGATACTCGAGACACCGCGCCCTATCCGTGTGAGTTGTGATGATGCCTCACACGCCTGCACCTTCCGTAGGTGC
 GGAGAGCGACCTTACCATTGGCTCAGTTGCTACTGAAGATGTTCCACGCATCCTCGAGAAAATAGAAAATGTGGCGAGA
 TGGCCAACAGGGACCTTGGCCTTCTCCGAGGATAAACCGGTAGATGACCAACTTGTCAACGACCCCGGATATCGTCG
 CGGAGGCTGACGAGAGCACATCAGCTCCGTCCGAGGCACAGGTGGCGCCGGCTTTTACCATTGCGCGCTTCAGA
 TGGCGCGGATGCGGACGGGGGGGGCCGTTTCCGACGGTAAAAAGAAAGCTGAAAGGCTCTTTGACCAACTGAGCCGTC
 AGGTTTTTGACCTCGTCTCCATCTCCCTGTTTTCTTCTCACGCCTTTTCTACCCTGGCGGTGGTTATTCTCCGGGTGAT
 TGGGGTTTTGACGTTTTACTTATTGTGCTCTTTTTATGTTACAGTTACCCAGCCTTGGTATTGCTCCCTCTTGGG
 TGTGTTTTCTGGGCTTCTCGGCGCTTGAATGGGGTTTTTGGCTGCTGGTTGGCTTTGCTGTTGGTCTGTTCAAGC
 CTGTGCTCGACCCAGTCGGCGCTGCTGTGAGTTTGACTGCCAGAGTGTAGAAACATCCTTATTCTTTGAGCTTCTC
 AAACCTTGGGACCTGTTGCGAGCCTTGTGTGGGCCCCGTGGTCTCGGCTTGCATTCTTGGCAGGTTACTGGGCGG
 GGCACGCTGCATCTGGCACTTTTGTAGGCTTGGCATTGTTGCAGACTGTATCTTGGCTGGAGCTTACGTGCTTCTC
 AAGGTAGGTGTAAAAAGTGTGGGATCTTGTATAAGAAGTGTCCAATGAGGTGCTTTTAACTGTTTCTTTTCA
 CGTGCGACCAAGTGTGCACTTATCGACCTGTGCGATCGGTTTTGTGCGCCAAAAGGAATGGACCCATTTTTCTGCCAC

Fig. 1H-2

TGGGTGGCGCGGTGCTGGGCCGGCCGAAGCCCCATTGAGCAACCCTCTGAAAAACCCATCGCGTTTGCCAGTTGGATG
 AAAAGAAGATTACGGCTAGGACTGTGGTCGCCCAGCCTTATGACCCCAACCAGCCGTAAAGTGCTTGCGGGTATTGCAG
 GCGGGTGGGGCAGTGGTGGCTAAGGCGGTCCCAAAAGTGGTCAAGGTTTCCGCTGTTCCATTCGAGCCCCCTTCTTTCC
 CACTGGAGTGAAAGTTGACCCGTATTGCAGGGTCGTGGTTGACCCTGACACTTTCACTGCAGCTCTCCGGTCTGGCTACT
 CCACCACAAACCTCGTCCCTTGGTGTGGGGACTTTGCCAGCTGAATGGATTAAAAATCAGGCAAAATTTCAAGCCTTCA
 GGGGGAGGCCACATCTCATGGCTGCCCTGCATGTTGCCTGCTCGATGGCTCTGCACATGCTTGCTGGGATTTATGTGAC
 TGCGGTGGGTTCTTGCGGACCGGCACCAACGACCCGTGGTGGCGTAACCCGTTTGCCGTCCCTGGCTACGGACCTGGCT
 CTCTCTGCACGTCCAGATTGTGCATTTCCCAACACGGCCTTACCCTGCCCTTGACAGCACTTGTCGGGGATTTCGGTATT
 CAAGAAATTGCCTTGGTCTTTTTGATTTTTGTTTCCATCGGAGGCATGGCTCATAGGTTGAGCTGTAAGGCTGACATGCT
 GTGTGCTTGCTTGCAATTGCCAGCTATGTTGGGTACCTTTACCTGGTTGCTTTGTGTGTTTCTTGCTGGTTGCGCT
 GTTTTCTTTCACCCCTCACCATCCTATGGTTGGTGTTCCTTGATTTCTGTAATATGCCTCAGGAATCTTGCC
 ATGGTGTGTTGGTTTCTTCTGGCTTCTGGTCTTATACTAATGTTGCTGGCCTTGTCACCCCTACGACATTCATCA
 TTACACCAAGTGGCCCCCGGGTGTGCGCCTTGCTACCGCACCAGATGGGACCTACTTGCCGCTGTCCGCCGCGCTG
 CGTTGACTGGCCGACCATGCTGTTACCCCGTCCAGCTGGGTCTTCTTGAGGGTGCTTTCAGAACTCGAAAGCCC
 TCACTGAACACCGTCAATGTGATCGGGTCTCCATGGGCTCTGGCGGGTGTTTACCATCGACGGGAAAGTCAAGTGCGT
 AACTGCCGCACATGCTTACGGGCAATTGAGCTCGGGTTTCCGGGGTCTGGCTCAATCAAATGCTTGACTTTGACGTAA
 AGGGAGATTTTCGTATAGCTGATTGCCGAATTGGCAAGGGGCTGCCCCAAGACCCAATTCTGCACGGATGGATGGACT
 GGCCGTGCCTATTGGCTAACATCCTCTGGCGTCGAACCCGGCGTCATTGGAAAAGGATTCGCTTCTGCTTACCCGATG
 TGGCGATTCCGGGTCCCAAGTGATCACCGAGGCCGGTGAGCTTGTCGGCGTTCACACGGGATCGAATAAACAAGGGGGG
 GCATTGTTACGCGCCCTCAGGCCAGTTTGTAAATGTGGCACCATCAAGCTAAGCGAATTAAGTGAATCTTTGCTGGG
 CCTAAGTCCCGCTCGGTGATGTGAAGGTGGCAGCCACATAATTAAGACATAAGCGAGGTGCCTTCAGATCTTTGTG
 CTTGCTTGCTGCCAAACCTGAACTGGAAGGAGGCTCTCCACCGTCAACTTCTTTGTGTGTTTTCTCTGTGGAGAA
 TGATGGGACATGCCTGGACGCCCTTGGTTGCTGTGAGTTCTTTATTTGAATGAGGTTCTCCAGCCGTCTGGTCCGG
 AGTGTCTTCTCTTGGAAATGTTGTGCTATCTGGCTCACGCCATGGTCTGCGCAAGTTCTGATGATCAGGCTTCTGAC
 AGCAGCTCTAACAGGAACAGATGGTCACTTGCTTTTTGAGCCTCGGTGAGTGACCGGTTTTGTGCGAGATCTTGCGG
 CCACTCAGGGGCATCCGTTGCAGGCAGTGATGAATTTGAGCACCTATGCATTCTGCTCGGATGATGGTTGTGACCTCA
 CCAGTCCAGGTGATCACGTGTGGTGTGCTGCACCTACTTGCCATCATTTGTACTTGTTAAGTACCGTGGCTGCACCA
 TATCCTTGTGCGGATGGAGTGTTCTCTGCGGCTTCTTCTTGAGATACTTTGCCGAGGAAAGTTGAGGGAAGGGGTGT
 CGCAATCCTGCGGAATGAATCATGAGTCTGACTGGTGCCCTCGCTATGAGACTCAATGACGAGGACTTGGATTTCTT
 ATGAAATGGACTGATTTTAAGTGCTTTGTTCTGCGTCCAACATGAGGAATGCAGCGGGTCAATTTATCGAGGCTGCCTA
 TGCTAAAGCACTTAGAGTAGAACTGGCCAGTTGGTGAGGTTGATAAAGTTCGAGGTACTTTGGCCAACTTGAAAGCTT
 TTGCTGATACCGTGGCACCTCAACTCTGCCCCGTGACATTGTTGTGCTCTCGGCCACACGCTGTTGGCAGTATCTTC
 GACCTAAAGGTTGGTAGCAACAGCATACCTCCAAGCCATTGAGACCAGAGTCTTGCTGGGTCCAAAATGACCGTGGC
 GCGCGTCTGACCCGACCCCAACGCCCCACCCGACCCGTGCCATCCCCCTCCACCGAAAGTTCTGGAGAAATGGCC
 CCAACGCTTGGGGGATGAGGACCGTTGAATAAGAAGAAGAGGCGCAGGATGGAAGCCCTCGGCATCTATGTTATGGG
 GGGAAAAAATACCAGAAATTTTGGGACAAGAATCCGGTGATGTGTTTATGAGGAGGTCCATAATAACACAGATGAGTG
 GGAGTGTCTCAGAGTTGGCGACCCTGCCGACTTTGACCTGAGAAGGGAATCTGTGTGGACATGTACCAATTGAAAACA
 AGGCTTACCATGTTTACACCTCCCCATCTGGTAAGAAGTTCTTGGTCCCCGTCAACCCAGAGAAATGGAAGAGTCCAATGG
 GAAGCTGCAAAGCTTCCGTGGAGCAGGCCCTAGGTATGATGAATGTGACGGCGAACTGACTGCCAAAGAACTGGAGAA
 ACTGAAAAGAATAATTGACAACTCCAGGGCTGACTAAGGAGCAGTGTTTAACTGCTAGCCGCCAGCGACTTGACCCG
 CTGTGGTCGCGCGGCTTGGTTGTTACTGAAACAGCGGTAAAAATAGTCAAATTTACAAACCGGACCTTACCCTGGGAC
 CTGTGAATTTAAAGTGCCAGTGAGGTTGAGCTAAAAGACGCGTTGAGCACAAACACCCGGTTGCGAGACCGATC
 GATGGTGGAGTTGTGCTCTGCGTTCGCGGTTCTTCTGCTTATAGACGTCTTGATCTCCGGTGCTGATGCATCTCCAA
 GTTACTTGCCCATCACGGGCCGGGAAACACTGGGATCGATGGCACGCTCTGGGATTTTGAAGTCCGAAGCCACTAAAGAGG
 AAGTCGCACTCAGTGCAGCAATAATACAGGCTTGACATTAGGCGCGGCGACGCTCTGAAATTTGGTCTCCCTTACAAG
 CTGTACCTGTTAGGGGTAACCTGAGCGGGTGAAGGAGTTCTGCAGAATACAAGTTTGGAGACATACCTTACAAAAC
 CCCAGTGACACTGGAAGCCAGTGACGCGGCTGCTGCTTACGCCAACGCCACTCCGGTGACTGATGGGCGCTCCG
 TCTTGGCCACGACCATGCCCCCGGGTTTGAAGTATATGTACCGACCATACAGCGTCTGTCTTGATTACCTTGACTCT
 AGGCTGACTGCCCTAAACAGCTGACAGAGCAGGCTGCGAAGATGCCGCACTGAAAGACCTCTCTAAATATGACTTGT
 CACCAAGGCTTTGTTTACCTGGAGTCTTCGCTTGTGCGGAAATACCTGTTTGCCATGTAGGTAAGTGCCACCCG

Fig. 1H-3

TTCATCGGCTTCTACTTACCCTGCTAAGAATTCTATGGCTGGAATAAATGGGAACAGGTTCCCAACCAAGGACATTGAG
 AGCGTCCCTGAAATCGACGTTCTGTGCGCACAGGCTGTGCGAGAAAACCTGGCAAACCTGTACCCCTTGTACTCTTAAGAA
 ACAGTATTGCGGGAAGAAGAAGACTAGGACCATACTCGGCACCAATAACTTCATCGCACTAGCCACCCAGGAGCTGTTGA
 GTGGTGTACCCAGGCTTCATGAAAAAGGCGTTAACTCGCCCATCGCCCTCGGAAAGAACAAGTTTAAAGGAGCTACAG
 ACTCCGGTCTGGGAGGTGCTTGAAGCTGATCTCGCATCCTGCGATCGATCCACGCTGCAATTGTCCGCTGGTTTGC
 CGCCAACCTTCTTTATGAACCTTGCTGTGCTGAAGAGCATCTACCGTCTGACGTGCTGAACTGCTGCCACGACTTACTGG
 TCACGCAGTCCGGCGCAGTGACTAAGAGAGGTGGCTGTGCTGTGGCGACCCGATCACCTCTGTGCTAACACCATTTAT
 AGTTTGGTGATCTATGCACAGCATATGGTGCTTAGTTACTTCAAAAGTGGTCACCCCATGGCCTTCTGTTCTTACAAGA
 CCAGCTAAAGTTTGAAGGACATGCTCAAGGTTCAACCCCTGATCGTCTATTGCGACGACCTCGTGCTGTATGCCGAGTCTC
 CCACCATGCCAACTATCACTGGTGGGTTGAACATCTGAATTTGATGCTGGGTTTTCAGACGGACCCAAAGAAGACAGCA
 ATAACAGACTCGCCATCATTTCTAGGCTGTAGAATAATAAATGGGCGCCAGCTAGTCCCAACCGTGACAGGATCCTCGC
 GGCCCTCGCTATCACATGAAGGCGAGTAATGTTTCTGAATACTATGCCTCAGCGGCTGCAATACTCATGGACAGCTGTG
 CTTGTTTGGAGTATGATCCTGAATGGTTTGAAGAACTTGAGTTGGAATAGCGCAGTGCGCCGCAAGGACGGCTACAGC
 TTTCCCGGCACGCCGTTCTTCATGTCCATGTGGGAAAACTCAGGTCCAATTATGAGGGGAAGAAGTCGAGAGTGTGCGG
 GTACTGCGGGGCCCCGGCCCCGTACGCTACTGCCTGTGGCTCGACGTCTGCATTTACCACACCCACTTCCACCAGCATT
 GTCCAGTCACAATCTGGTGTGGCCATCCAGCGGGTCTGGTCTTGTAGTGAGTGCAAATCCCCTGTAGGGAAAGGCACA
 AGCCCTTTAGACGAGGTGCTGGAACAAGTCCCGTATAAGCCCCACGACCGTTATCATGCATGTGGAGCAGGGTCTCAC
 CCCCCTTGATCCAGGTAGATACCAAACCTCGCCGCGGATTAGTCTCTGTGAGGCGTGGAATTAGGGGAAATGAAGTTGGAC
 TACCAGACGGTGATTATGCTAGCACCGCCTTGCTCCTACCTGCAAAGAGATCAACATGGTCTGCTGCTTCCAATGTA
 TTGCGCAGCAGGTTCATCATCGGCCACCCGGTGCTGGGAAAACTACTGGCTCCTTCAACAGGTCCAGGATGGTGATGT
 TATTTACACCAACTCACCAGACCATGCTTGACATGATTAGGGCTTTGGGACGTGCCGTTCAACGTCCCGGCAGGCA
 CAACGCTGCAATTCCTCGTCCCTCCCGCACCGGTCCGTGGGTTGCGATCCTAGCCGGCGGTTGGTGCTTGCAAGAAT
 TCCTTCTAGATGAAGCAGCGTATTGCAATCACCTTGATGTTTTGAGGCTTCTAGTAAACTACCTCACCTGTCTAGG
 AGACTTCAAGCACTCCACCCAGTGGGTTTTGATTCTCATTGCTATGTTTTGACATCATGCCTCAAACCTCAACTGAAGA
 CCATCTGGAGTTTGGACAGAAATCTGTGATGCCATTGAGCCAGATTACAGGGACAACTCATGTCCATGGTCAACACA
 ACCCTGTGACCTACGTGGAAAAACCTGTGAGGTATGGGACGGTCTCACCCCTACCACAGGGACCGAGAGGACGACGC
 CATCATATTGACTCCAGTCAAGGCGCCACATTGATGTGGTTACATTGCATTTGCCACTAAAGATTCACTCAACAGGC
 AAAGAGCCCTTGTTGCTATCACCAGGGCAAGACAGCTATCTTTGTGTATGACCCACACAGGCAGCTGCAGGGCTGTTT
 GATCTTCTGCAAAAGGCACGCCCTCAACCTCGCAGTGCACTGCGACGGGAGCTGATCGTGCTGGATAGAAATAACAA
 AGAATGCACGGTTGCTCAGGCTTAGGCAACGGGGATAAATTTAGGGCCACAGACAAGCGTGTGTAGATTCTCTCCGCG
 CCATTTGTGCTGATCTAGAAGGGTCGAGCTCTCCGCTCCCAAGGTGCGACACAACCTGGGATTTTATTTCTCACCTGAT
 TTAACACAGTTTGCTAACTCCAGTAGAACTTGACCTCACTGGCCCGTGGTGTCAACCCAGAACAATGAAAAGTGGCC
 GGATCGGCTGGTGGCAGCCTTCCGCTATCCATAAATACAGCCGCGGTGCTCGGTGCGCGCTATATGGTGGGCCCTT
 CGGTGTTCTAGGCACTCCTGGGGTCTGTGTCATACTATCTACAAAATTTGTTAAGGGCGGGGCTCAAGTGCTTCCGGAG
 ACGGTTTTAGCACCGGCCAATTGAGGTAGACTGCCGGGAATATCTTGATGATCGGGAGCGAGAAGTTGCTGCGTCCCT
 CCCACACGCTTTCATTGGCGACGTCAAAGGCACTACCGTTGGAGGATGTATCATGTACCTCCAGATACCTCCCGCGCG
 TCCTTCCCAAGGAATCAGTTGCGGTAGTCGGGTTTCAAGCCCCGAAAGCCGCGAAAGCATTGTGCACACTGACAGAT
 GTGTACCTCCAGATCTTGAAGCCTATCTCCACCCGAGACCCAGTCCAAGTGCTGGAAAATGATGTTGGACTTCAAAGA
 AGTTCGACTAATGGTCTGGAAAGACAAAACAGCCTATTTCAAACCTGAAGGTGCTATTTACCTGGTATCAGCTTGCCA
 GCTATGCCTCGGTACATCCGTGTTCCGCTCAACTCTACGGTGTACTTGGACCCCTGCATGGGCCCCGCCCTTTGCAACAGG
 AGATGCTCGGTCCACCCACTGGGGGGCTGACCTCGCGGTACCCCTTATGATTACGGCGCTAAAATTATCTGTCTAG
 CGGTACCATGGTGAAATGCCCCCCGGATACAAAATTTGCGGTGCGCGGAGTTCTGTTGGATGACCCAGTTAAGTACA
 AACATACCTGGGGTTTGAATCGGATACAGCGTATCTGTATGAGTTACCGGAAACGGTGAGGACTGGGAGGATTACAAT
 GATGCGTTTCTGCGCGCCAGGAAGGGAAAATTTATAAGGCCACTGCCACCAGCTTGAAGTTTTATTTCCCGGGGCC
 TGTCAATTGAACCAACTTAGGCCCTGAATTGAAATGAAATGGGGTCCATGCAAAGCCTTTTGACAAAATTGGCCAATTT
 TTGTGGATGCTTTCACGGAGTCTTGGTGTCCATTGTTGATATCATTATATTTTGGCCATTTTGTGTTGCTTACCATC
 GCCGGTGGCTGGTGGTCTTTGTCATCAGATTGGTTTGTCCGCGATACTCCGTACGCGCCTGCCATTCACTCTGAGCA
 ATTACAGAAGATCTTATGAGGCCTTCTTTCCAGTGCCAAGTGGACATTTCCACCTGGGGAACATAACATCTTTGGGG
 ATGCTTTGGCACCATAAGGTGTCAACCCTGATTGATGAAATGGTGTGCGCTGCAATGTACCGCATCATGAAAAAGCAGG
 GCAGGCTGCTGGAACAGGTGGTGGAGCGAGGCTACGCTGTCTCGCATTAGTAGTTTGGATGTGGTGGCTCATTTTCAGG

Fig. 1H-4

ATCTAGCCGCCATTGAAGCCGAGACCTGTAATATTTGGCCTCCCGGCTGCCCATGCTACACAACCTGCGCATGACAGGG
 TCAAATGTAACCATAGTGTAATAGCACTTTGAATCAGGTGTTTGCTATTTTTCAACCCCTGGTCCCGGCCAAAGCT
 TCATGATTTTCAGCAATGGTTAATAGCTGTACATTCTCCATATTTCTCTGTTGCAGCTTCTTGACTCTTTTGTTG
 TGCTGTGGTTCGGGTTCCAATACTACGTACTGTTTTGGTTTCCGCTGGTTAGGGGCAATTTTCTTTCGAACCTCACAG
 TGAATTACACGGTGTGTCCACCTTGCCTCACC CGCAAGCAGCCACAGAGATCTACGAACCCGGTAGGTCTCTTGGTGC
 AGGATAGGGTATGACCGATGTGGGAGGACGATCATGACGAGCTAGGGTTTATGATACCGCCTGGCCTCTCCAGCGAAGG
 CCACTTGACTGGTGTGTACGCCTGGTTGGCGTTCTTGTCTTCAGCTACACGGCCAGTTCCATCCCGAGATATTCGGGA
 TAGGGAATGTGAGTCGAGTTTATGTTGACATCAAACTCAACTCATCTGCGCCGAACATGACGGGCAGAACACCACTTG
 CCTCGTCATGACAACATTTAGCCGCTGTTTACAGCCTATTACCAACATCAAGTCGACGGCGGCAATTGGTTTCACTAGA
 ATGGCTTCGTCCTTCTTTCTCTGCTGGTGGTTTTAAATGTCTCTTGGTTTCTCAGGCGTTCGCTGCAAAACCATGTTT
 CAGTTCGAGTCTTGAGATATTAAGACCAACACCACCGCAGCGGCAAGCTTGTCTCTCAAGACATCAGTTGCCTTA
 GGCATCGGACTCGGCTCTGAGGCGATTGCAAAATCCCTCAGTGCCGTACGGCGATAGGGACACCCGTGTATGTTACC
 ATCAGCCCAATGTGACAGATGAGAATTATTACATTCTTCTGATCTCTCATGCTTCTTCTTGCCTTTCTATTGCTTC
 TGAGATGAGTGAAAAGGGATTTAAGGTGGTATTTGGCAATGTGTGAGGCATCGTGGCTGTGTGTCAATTTTACCAGCT
 ACGTCCAACATGTCAAGGAGTTTACCAACGCTCCCTGGTGGTGCACCATGTGCGGTTGCTCCATTTATGACACCTGAG
 ACCATGAGGTGGGCACTGTTTAGCCTGTCTTTTGCATTCTGTTGGCAATTTGAATGTTAAGTATGTTGGAGAAAT
 GCTTGACCGCGGCTGTTGCTCGCGATTGCTTCTTTGTGGTGTATCGTGCCGTTCTGTTTGTGTGCTGCCAACGCC
 AGCAACGACAGCAGCTCCATCTACAGCTGATTTACAACCTGACGCTATGTGAGCTGAATGGCAGAGTTGGCTAGCTAA
 CAAATTTGATTGGGCGAGTGGAGAGTTTGTATCTTTCCCGTTTGACTCACATTGTCTCCTATGGTGCCCTCACTACCA
 GCCATTTCTTGACACAGTCGCTTTAGTCACTGTGTCTACCGCCGGGTTTGTTCACGGGCGGTATGTCCTAAGTAGCATC
 TACGCGGTCTGTGCCCTGGCTGCGTTGACTTGCTTCGTATTAGGTTTGCAAAGAATTGCATGTCCTGGCGCTACGCGTG
 TACCAGATATACCACTTTCTTCTGGACACTAAGGGCAGACTCTATCGTTGGCGGTGCGCTGTCTCATAGAGAAAAGGG
 GCAAAGTTGAGGTGGAAGGTCTATGATCGACCTCAAAAGAGTTGTGCTTGATGGCTCCGTGGCAACCCCTATAACCAGA
 GTTTTCAGCGGAACAATGGGGTCTCTTAGATGACTTCTGTACGATAGCACGGCTCCACAAAAGGTGCTTTTGGCGTTT
 TCTATTACCTACACGCCAGTGATGATATATGCCCTAAAAGGTGAGTCGCGGCCGACTGCTAGGGCTTCTGCACCTTTGAT
 CTTCTGAATTGTGCTTTACCTTCGGGTACATGACTTTCGCGCACTTTCAGAGTACAAATAAGGTGCGGCTCACTATGG
 GAGCAGTAGTTGCACTCCTTTGGGGGGTGTACTCAGCCATAGAAACCTGGAAATTCATCACCTCCAGATGCCGTTTGTGC
 TTGCTAGGCCGCAAGTACATTCTGGCCCTGCCACCACGTTGAAAGTGCCGAGGCTTTCATCCGATTGCGGCAATGA
 TAACCACGCAATTTGTGTCGCGCGTCCCGGCTCCACTACGGTCAACGGCACATTGGTGCCCGGGTTAAAAAGCCTCGTGT
 TGGGTGGCAGAAAAGCTGTTAAACAGGGAGTGGTAAACCTTGTCAAATATGCCAAATAACAACGCAAGCAGCAGAAGAG
 AAAGAAGGGGATGGCCAGCCAGTCAATCAGCTGTGCCAGATGCTGGGTAAGATCATCGCTCAGCAAAACCACTCCAGAG
 GCAAGGGACCGGGAAGAAAAATAAGAAGAAAAACCCGAGAAGCCCCATTTCTCTAGCGACTGAAGATGATGTCAGA
 CATCACTTTACCCCTAGTGAGCGGCAATTGTGTCTGTCTGTCATCCAGACCGCCTTAAATCAAGGCGCTGGGACTTGAC
 CTTGTGAGATTGAGGGAGGATAAGTTACACTGTGGAGTTTAGTTTGCCTACGCATCATACTGTGCGCCTGATCCGCGTCA
 CAGCATCACCTCAGCATGATGGGCTGGCAATCTTGAGGCATCTCAGTGTTTGAATTGGAAGAATGTGTGGTGAATGGCA
 CTGATTGACATTGTGCTCTAAGTCACCTATTCAATTAGGGCGACCGTGTGGGGTGAGATTTAATTGGCGAGAACCATG
 CGGCCGAATTAATAAAAAA

Fig. 11-1

>V7-Nsp2d543-632.seq

ATGACGTATAGGTGTTGGCTCTATGCCTTGGCATTGTATTGTGAGGAGCTGTGACCATTGGCACAGCCCAAACTTGCT
 GCACAGAAACACCTTCTGTGATAGCCTCCTTCAGGGGAGCTTAGGGTTTGCCCTAGCACCTTGCTTCCGGAGTTGCAC
 TGCTTTACGGTCTCTCACCCCTTTAACCATGTCTGGGATACTTGATCGGTGCACGTGTACCCCAATGCCAGGGTGT
 ATGGCGGAGGGCCAAGTCTACTGCACACGATGCCTCAGTGCACGGTCTCTCTTCCCCTGAACCTCCAGGTTCTGAGCT
 CGGGGTGCTAGGCCTATTCTACAGGCCCGAAGAGCCACTCCGGTGGACGTTGCCACGTGCATTCCCCACTGTTGAGTGCT
 CCCCCGCCGGGGCTGCTGGCTTTCTGCAATCTTCCAATCGACGAATGACCAGTGGAACCTGAACCTCCAACAAAGA
 ATGGTACGGTGCAGCTGAGCTTTACAGAGCCGGCCAGCTCACCCCTGCAGTCTGAAGGCTCTACAAGTTTATGAACG
 GGGTTGCCGCTGTTACCCATTGTTGGACCTGTCCCTGGAGTGGCCGTTTTGCCAATCCCTACATGTGAGTGATAAAC
 CTTCCCGGGAGCAACTCAGTGTTGACCAACCTGCCGCTCCCGCAGAGACCAAGCCTGAAGACTTTTGGCCCTTTGAG
 TGTGCTATGGTACTGTCTATGACATTGGTCATGACGCCGTGATGTATGTGGCCGAAAGGAAAGTCTCTGGGCCCCCTCG
 TGGCGGGGATGAAGTGAATTTGAAGCTGTCCCGGGGAGTTGAAGTTGATTGCGAACCGGCTCCGCACCTCTTCCCGC
 CCCACCACACAGTGGACATGTCTAAGTTCGCCTTCACAGCCCTGGGTGTGGTGTCTATGCGGGTGAACGCCAACAC
 GGTGCTTCCCGCTGACACTGTCCCTGAAGGCAACTGCTGGTGGAGCTGTTTGACTTGCTTCCACTGGAAGTTCAGAA
 CAAAGAAATTCGCATGCTAACCAATTTGGCTACCAGACCAAGCATGGTGTCTTGGAAGTACCTGCAGCGGAGGCTGC
 AAGTTAATGGTCTCGAGCAGTAACCTGAACCTAACCGACCTATCGTCGTACAGTACTTCTCCGTTAAGGAGAGTTGGATC
 CGCCATTTGAACTGGCGGGAGAACCCAGCTACTCTGGGTTGAGGACCTCTCAGAATAAGGGTTGAGCCTAACACGTC
 GCCATTGGCTGACAAGGAAGAAAAATTTCCGGTTTGGCAGTACAAGTGGTACGGCGCTGGAAAGAGAGCAAGAAAAG
 CACGCTCTGTGCGACTGCTACAGTCGCTGGCCGCGCTTTGTCCGTTCTGTAACCCGGCAGGCCAAGGAGCAGAGGTT
 GCCGGCGCCAAACAGGCTGAGCACCTCAAACACTACTCCCGCCTGCCGAAGGGAATTGTGGTTGGCACTGCATTTCCGC
 CATCGCCAACCGGATGGTGAATTCAAATTTGAAACACCCCTTCCGAAAGAGTGAGACCTCCAGATGACTGGGCTACTG
 ACGAGGATCTTGTAATGCCATCCAAATCTCAGACTCCCTGCGGCTTAGACAGGAACGGTGTGTACTAGCGCCAAG
 TACGACTTAAGCTGGAAGGTGAGCATTGGACTGTCACTGTGACCCCTGGGATGTCCCTTCTTGTCTCCCTCTTGAATG
 TGTTACGGGCTGTTGTGGGCACAAGGGCGGTCTTGGTTCCCGAGATCGAGTCTCGGATTGACCCCTGCCTGCC
 TTGACCGGCTGGCTGAGGTGATGCACCTGCCTAGCAGTGTATCCCGAGCCGCTTGCCGGAATGTCTGGCGATTCCGAT
 CGTTGGCTTCTCCGGTCACACCGGTGTGGACTGTTTCGAGTCTTTTGGCCGTACAGCGGAGGGAATCACCTGACCA
 AGTGGCTTAGGGAAAAATTTATCAGCCTTTGTGAGGTGATTGAGGACTGCTGCTTCCAGAACAAAACCAACCGGCTCA
 CCCCCGAGGAGGTGCGAGCAAGATTGACCTGTACCTCCGTGGTGCAACAAATCTGAAGAATGCTTGGCAGGCTTGAG
 AAAGCGCGCCCGCCACGCGTAATCGACACCTCTTTGATTGGGATGTTGTGCTCCCTGGGGTTGAGGCGGCAACCCAGAC
 GATCAAGCTGCCCCAGGTCAACAGTGTCTGCTCTGGTCCCTGTTGTGACTCAAAAGTCTTGGACAACAACTCGGTCC
 CCTGACCGCTTTTCACTGGCTAACTACTACTACCTGTGCGCAAGGTGACGAAGTTCGTACCGTGAAAGACTAACCGCC
 GTGCTCTCAAGTTGAAAAAGGTTGTTGAGAAGAATATGGGCTCATGCCAACCGAGCCTGGTCCACGGCCCACTGCC
 ACGCGGGCTGACGAACTCAAAGACAGATGGAGGAGGACTGCTGAACTGGCTAACGCCAGACGACTTCGGACATGA
 TGGCTTGGGAGTGCAGCAGGTTGACCTAAAACTTTGGGTCAAGAACTACCCGCGGTGGACACCACCCCTCCGCCA
 AAAGTTACGCTCGAAAAACGAAGCCTGTCAAGAGCTTGGCGAGAGAAAGCCTGTCCCGCCCCGCGCAGGAAGGTTGG
 GTCCGATTGTGGCAGCCCGGTTTATTAGGCGGCGATGTCCCTAACAGTTGGGAAGATTGGCTGTTAGTAGCCCCCTTG
 ATCTCCGACCCACCTGAGCCGGCAACACCTTCAAGTGAGCTGGTGATTGTGTCTCACCGCAATGCATCTTCAGGCCG
 GCGACACCTTGAGTGAGCCGGCTCAATTCGCGACCTCGCGGAACGTGTCTCGACCGGTGACACCTTGAGTGAGCC
 GATCACACGCCCAAAATACTCAGCTCAAGCCATCATCAGTCCGGCGGGCCCTGCAGTGGGCATCTCAAGAGGTAAGG
 AAACATGCCTTAGTGTATGCGCGAGGCATGTGATGCGACTAAGCTTGATGACCTGCTACGAGGAATGGCTTTCTCGC
 ATGTGGGATCGGGTGGACATGCTGACTTGGCGCAACACGTCTGTTTACCAGGCGATTGACCTTAGATGGCAGGTTAAA
 GTTCTCCAAAAATGATACTCGAGACACCGCCGCTATCCGTGTGAGTTGTGATGATGCTCACACGCTGCACCTT
 CCGTAGGTGCGGAGAGCGACCTTACCATTGGCTCAGTTGCTACTGAAGATGTTCCACGCATCCTCGAGAAAATAGAAAAT
 GTCGGCGAGATGGCAACAGGACCTTGGCCTTCTCCGAGGATAAACCGGTAGATGACCAACTGTCAACGACCCCG
 GATATCGTCGCGGAGGCTGACGAGAGCAGATCAGTCCGTCCGAGGCACAGGTGGCGCGGCTCTTTTACCAGTTTGC
 CGCCTTCAAGTGGCGGATGCGGACGGGGGGGGCCGTTTCCGACGGTAAAAAGAAAGCTGAAAGGCTCTTTGACCAA
 CTGAGCCGTGAGTTTTTACCTCGTCTCCATCTCCCTGTTTTCTCTCACGCTTTTCTACCTGGCGGTGGTTATTC
 TCCGGGTGATTGGGGTTTTGAGCTTTTACTTATTGTGCTCTTTTATGTTACAGTTACCCAGCCTTTGGTATTGCTC
 CCTCTTGGGTGTGTTTTCTGGGTCTTCTCGGCGGTTTGAATGGGGTTTTTGGCTGCTGGTTGGCTTTTCTGTTGGT
 CTGTTCAAGCCTGTGTCGACCCAGTCGGCGCTGCTGTGAGTTGACTCGCCAGAGTGTAGAAACATCCTTATTCTTT

Fig. 11-2

TGAGCTTCTCAAACCTTGGGACCTGTTGCGAGCCTTGTGTGGGCCCGCTCGGTCTCGGTCTTGCCATTCTTGGCAGGT
 TACTGGGCGGGGCACGCTGCATCTGGCACTTTTGTCTAGGCTTGGCATTGTTGCAGACTGTATCTTGGCTGGAGCTTAC
 GTGCTTTCTCAAGGTAGGTGTAAGGAGTGTGGGGATCTGTATAAGAACTGCTCCAATGAGGTGCTTTTAACTGTT
 TCCTTTCACACGTGCGACAGGTGCTCACTATCGACCTGTGCGATCGGTTTTGTGCGCCAAAAGGAATGGACCCCATTT
 TTCTCGCACTGGGTGGCGCGGTGCTGGCCGGCCGAAGCCCCATTGAGCAACCCTCTGAAAAACCCATCGCGTTTGGC
 CAGTTGGATGAAAAGAAGATTACGGCTAGGACTGTGGTCGCCAGCCTTATGACCCCAACCAAGCCGTAAAGTGCTTGGC
 GGTATTGACGGCGGTGGGGCGATGGTGGCTAAGGCGGTCCCAAAAGTGGTCAAGGTTTCCGCTGTTCCATTCCGAGCCC
 CCTTCTTTCCTACTGGAGTGAAAGTTGACCCTGATTGACGGGTGCTGGTTGACCTGACACTTTCCTGACAGCTCTCCGG
 TCTGGCTACTCCACCACAAACCTCGTCTTGGTGTGGGGGACTTTGCCAGCTGAATGGATTAAAAATCAGGCAAATTTT
 CAAGCCTTCAGGGGGAGGCCACATCTCATGGCTGCCCTGCATGTTGCCCTGCTCGATGGCTCTGCACATGCTTGTGGGA
 TTTATGTGACTGCGGTGGGTCTTTCGGGCACCGGCACCAACGACCCGTGGTGGCTAACCCGTTTGCCGTCCCTGGCTAC
 GGACCTGGCTCTCTCTGCACGTCCAGATTGTGCAATTTCCCAACACGGCCTTACCCTGCCCTTGACAGCACTTGTGGCGGG
 ATTCCGTATTCAAGAAATTGCCTTGGTGTGTTTTGATTTTTGTTTCCATCGAGGCATGGCTCATAGGTTGAGCTGTAAGG
 CTGACATGCTGTGTCTTGTCTGCAATTGCCAGCTATGTTGGGTACCTCTTACCTGGTTGCTTGTGTGTTTCTTGC
 TGGTTGCGCTGTTTTCTTTCACCCCTCACCATCCTATGGTGGTGTGTTTTCTTGATTTCTGTGAATATGCCTTCAGG
 AATCTTGGCCATGGTGTGTTGGTTTCTCTTGGCTTCTTGGTGTATATAATGTTGCTGGCCTTGTACCCCTACG
 ACATTATCATTACACAGTGGCCCCCGCGGTGTTGCCGCTTGGCTACCGCACCAGATGGGACCTACTTGGCCGCTGTC
 CGCCGCTGCGTTGACTGGCCGCACCATGCTGTTACCCCGTCCAGCTTGGGTCTCTTCTGAGGGTGCTTTCAGAAC
 TCGAAAGCCCTCACTGAACACCGTCAATGTATCGGGTCTCCATGGGCTTGGCGGGGTGTTTACCATCGACGGGAAAG
 TCAAGTGCGTAACCTGCCGACATGCTTACGGGCAATTGAGCTCGGGTTTCCGGGTGCGCTTCAATCAATGCTTGAC
 TTTGACGTAAAGGGAGATTTGCTATAGCTGATTGCCGAATTGGCAAGGGGCTGCCCAAGACCAATTCTGCACGGA
 TGGATGGACTGGCCGTGCCTATTGGCTAACATCCTTGGCGTCAACCCGGCGCTATTGGAAGGATTGCGCTTCTGCT
 TCACCGCATGTGGCGATTCCGGGTCCCAAGTATCACCGAGGCGGTGAGCTTGTGCGGTTTACACGGGATCGAATAAA
 CAAGGGGGGGGATTGTTACGCGCCCTCAGGCCAGTTTGTAAATGTGGCACCATCAAGCTAAGCGAATTAAGTGAATT
 CTTTGTGGGCTAAGGTCCGCTCGGTGATGTGAAGGTGCGCAGCCACATAAATAAGACATAAGCGAGGTGCCTTCAG
 ATCTTTGTGCTTGTGCTGCTGCCAACCTGAACCTGGAAGGAGGCTCTCCACCGTCCAATCTTTGTGTGTTTTTCTC
 CTGTGGAGAATGATGGACATGCTGGACGCCCTTGGTGTGCTGTGAGTTTCTTATTTTGAATGAGGTTCTCCAGCCGT
 CCTGGTCCGGAGTGTTTTCTCTTGGAAATGTTTGTGCTATCTGGCTCACGCCATGGTCTGCGCAAGTTCTGATGATCA
 GGCTTCTGACAGCAGCTTTAACAGGAACAGATGGTCACTTGCTTTTTGAGCCTCGGTGAGTGACCGGTTTTGTGCGCA
 GATCTTGGCGCACTCAGGGGCATCCGTTGACGGCAGTGATGAATTTGAGCACCTATGCATTCTGCTCGGATGATGGT
 TGTGACCTCACAGTCCCAAGTACAGTGTGCTGCTGCACTACTTGCCATCATTTTGTACTTGTAAAGTACCGTG
 GCCTGCACCATATCCTTGTGGCGATGGAGTGTCTGCGGCTTCTTCTGAGATACTTGGCGAGGGAAAGTTGAGG
 GAAGGGGTGTCGAATCCTGCGGAATGAATCATGAGTCTGACTGGTGCCTCGCTATGAGACTCAATGACGAGGACTT
 GGATTTCTTATGAAATGGAATGATTTAAGTGCTTTGTTTCTGCGTCAACATGAGGAATGCAGCGGGTCAATTTATCG
 AGGCTGCCTATGCTAAAGCACTTAGAGTAGAACTGGCCAGTTGGTGCAGGTTGATAAAGTTGAGGTACTTTGGCCAAA
 CTTGAAGCTTTTGTGATACCGTGGCACCTCAACTCTGCCCCGTGACATTGTTGTGCTCTCGGCCACACGCTGTTGG
 CAGTATCTTGCACCTAAAGGTTGGTAGCACAAGCATACCTCCAAGCCATTGAGACCAGAGTCTTGTGCTGGGTCCAAAA
 TGACCGTGGCGCGCTGCTGACCCGACCCCCACGCCCCACCCGACCCGTGCCATCCCCCTCCACCGAAAGTTCTG
 GAGAATGGCCCCAACGCTTGGGGGATGAGGACGTTTGAATAAGAAGAAGGGCGCAGGATGGAAGCCCTCGGCATCTA
 TGTTATGGGCGGGAAAAAATACCAGAAATTTTGGGACAAGAATTCCGGTGATGTGTTTTATGAGGAGGTCCATAATAACA
 CAGATGAGTGGGAGTGTCTCAGAGTTGGCGACCTGCCGACTTGACCCTGAGAAGGGAACCTGTGTGGACATGTACC
 ATTGAAAACAAGGCTTACCATGTTTACACCTCCCATCTGGTAAGAAGTTCTTGGTCCCGTCAACCCAGAGAATGGAAG
 AGTCCAATGGGAAGCTGCAAGCTTTCCGTGGAGCAGGCCCTAGGTATGATGAATGTGACGGCGAACTGACTGCCAAG
 AACTGGAGAACTGAAAAGAATAATTGACAACTCCAGGGCCTGACTAAGGAGCAGTGTTTAACTGCTAGCCGCCAGCG
 ACTTGACCCGCTGTGGTGGCGCGGCTTGGTTGTTACTGAAACAGCGGTAAAAATAGTCAAAATTTACAACCGGACCTTC
 ACCCTGGGACCTGTGAATTTAAAGTGGCCAGTGAGGTTGAGCTAAAAGACGCGTTGAGCACAACCAACCCGGTTGC
 GAGACCGATCGATGGTGGAGTTGTGCTCCTGCGTTCGCGGTTCTTCTGCTTATAGACGCTTGTATCTCCGGTGCTGATG
 CATCTCCCAAGTTACTTGCCCATCACGGGCCGGGAAACACTGGGATCGATGGCACGCTCTGGGATTTTGAAGTCCGAAGCC
 ACTAAAGAGGAAGTGCACCTCAGTGGCAATAATACAGGCTTGTGACATTAGGCGCGGCGACGCTCTGAAATTTGGTCT
 CCCTTACAAGCTGTACCCTGTAGGGGTAACCTGAGCGGGTGAAAGGAGTTCTGCAGAATACAAGTTTGGAGACATAC

Fig. 11-3

CTTACAAAACCCCACTGACACTGGAAGCCCACTGCACGCGGCTGCCTGCCTTACGCCCAACGCCACTCCGGTGACTGAT
 GGGCGCTCCGTCTTGGCCACGACCATGCCCCCGGGTTGAGTTATATGTACCGACCATACCAGCGTCTGTCTTGATTA
 CTTGACTCTAGGCCTGACTGCCCTAAACAGCTGACAGAGCACGGCTGCGAAGATGCCGCACTGAAAGACCTCTCTAAAT
 ATGACTTGTCCACCCAAGGCTTTGTTTTACCTGGAGTTCTTCGCCTTGTCGGGAAATACCTGTTTGCCCATGTAGGTAAG
 TGCCACCCGTTTCATCGGCCTTCTACTTACCCTGCTAAGAATTCTATGGCTGGAATAAATGGGAACAGGTTCCCAACCAA
 GGACATTGAGAGCGTCCCTGAAATCGACGTTCTGTGCGCACAGGCTGTGCGAGAAAACCTGGCAAACCTGTACCCCTTGTA
 CTCTTAAGAAACAGTATTGCGGGAAGAAGAAGACTAGGACCATACTCGGCACCAATAAATTCATCGCACTAGCCACCGA
 GCAGTGTGAGTGGTGTACCCAGGGCTTCATGAAAAAGGCTTTAACTCGCCCATCGCCCTCGGAAAGAACAAGTTTAA
 GGAGCTACAGACTCCGGTCTGGGCAGGTGCCCTGAAGCTGATCTCGCATCTGCGATCGATCCACGCCTGCAATTGTCC
 GCTGGTTTGCCGCCAACCTTCTTTATGAACCTGCTGTGCTGAAGAGCATCTACCGTCGTACGTGCTGAACGCTGCCAC
 GACTTACTGGTCACGAGTCCGGCGCAGTGACTAAGAGAGGTGGCTGTCTGCTGGCGACCCGATCACCTCTGTGTCTAA
 CACCATTATAGTTTGGTGATCTATGCACAGCATATGGTGCTTAGTTACTTCAAAAGTGGTCACCCCATGGCCTTCTGT
 TCTTACAAGACCAGCTAAAGTTTGAGGACATGCTCAAGGTTCAACCCCTGATCGTCTATTGCGACGACCTCGTGCTGTAT
 GCCGAGTCTCCACCATGCCAACTATCACTGGTGGGTGAACATCTGAATTTGATGCTGGGGTTTCAGACGGACCCAAA
 GAAGACAGCAATAACAGACTCGCATCATTTCTAGGCTGTAGAATAAATAATGGGCGCCAGCTAGTCCCAACCGTGACA
 GGATCTCGCGGCCCTCGCTATCACATGAAGGCGAGTAATGTTTCTGAATACTATGCCTCAGCGGCTGCAATACTCATG
 GACAGCTGTGCTTGTGGAGTATGATCCTGAATGGTTTGAAGAACTGTAGTTGGAATAGCGCAGTGCGCCCGCAAGGA
 CGGCTACAGCTTTCCCGGCACGCGCTTCTCATGTCCATGTGGGAAAACTCAGGTCCAATTATGAGGGGAAGAAGTCGA
 GAGTGTGCGGGTACTGCGGGGCCCCGGCCCCGTACGCTACTGCTGTGGCTCGACGTCTGCATTTACCACACCCACTTC
 CACCAGCATTTGCCAGTACAATCTGGTGTGGCCATCCAGCGGGTCTGGTTCTGTAGTGAGTGCAAATCCCTGTAGG
 GAAAGGCACAAGCCCTTTAGACGAGGTGCTGGAACAAGTCCCGTATAAGCCCCACGGACCGTTATCATGCATGTGGAGC
 AGGGTCTCACCCTTGTATCCAGGTAGATACCAACTCGCCGCGGATTAGTCTCTGTGAGGCGTGGAAATAGGGGAAAT
 GAAGTTGGAATACCAGACGGTGATTATGCTAGCACCGCTTGTCTTACCTGCAAAGAGATCAACATGGTCGCTGTGCG
 TTCCAATGTATTGCGCAGCAGTTTCATCATCGGCCACCCGGTGTGGGAAACATACTGGCTCCTTCAACAGGTCCAGG
 ATGGTGATGTTATTTACACACCAACTCACCAGACCATGCTTGACATGATTAGGGCTTTGGGGACGTGCCGGTTCAACGTC
 CCGGCAGGCACAACGCTGCAATTTCCCGTCCCTCCCGCACCGGTCCGTGGGTTGCGATCCTAGCCGGCGGTTGGTGCTC
 TGGCAAGAATTCTTCTAGATGAAGCAGCGTATTGCAATCACCTTGATGTTTTGAGGCTTCTTAGTAAACTACCCTCA
 CCTGTCTAGGAGACTTCAAGCAACTCCACCCAGTGGGTTTTGATTCTCATTGCTATGTTTTGACATCATGCCTCAAAC
 CAATGAAGACCATCTGGAGGTTTGACAGAAATATCTGTGATGCCATTGAGCCAGATTACAGGGACAAACTCATGTCCAT
 GGTCAACACAACCCGTGTGACCTACGTGGAACAACTGTGAGGTATGGGCAGGTCTCACCCTTACCACAGGGACCGAG
 AGGACGACGCCATCACTATTGACTCCAGTCAAGGCGCCACATTCGATGTGGTTACATTGCATTTGCCCACTAAAGATTCA
 CTCAACAGGCAAGAGCCCTTGTGCTATCACCAGGGCAAGACAGCTATCTTGTGTATGACCCACACAGGCAGCTGCA
 GGGCTTGTGTTGATCTTCTGAAAAGGCACGCCGTCAACCTCGCAGTGCAGTGCAGCGGCGAGCTGATCGTGCTGGATA
 GAAATAACAAAGATGACCGGTTGCTCAGGCTCTAGGCAACGGGATAAATTTAGGGCCACAGACAAGCGTGTGTAGAT
 TCTCTCCGCGCATTTGTGCTGATCTAGAAGGGTCGAGCTCTCCGCTCCCAAGGTGCGACACAACCTTGGGATTTATTT
 CTCACCTGATTTAACACAGTTTGCTAAACTCCAGTAGAATTGCACCTCACTGGCCCGTGGTGTCAACCCAGAACAATG
 AAAAGTGGCCGGATCGGCTGGTTGCCAGCTTCGCCCTATCCATAAATACAGCCGCGCGTGCATCGGTGCCGGCTATATG
 GTGGGCCCTTCGGTGTCTAGGCACTCTGGGTCGTGTACTATCTCACAAAATTTGTTAAGGGCGGGGCTCAAGT
 GCTTCGGGAGACGGTTTTAGCACCGGCCGAATTGAGGTAGACTGCCGGAATATCTTGATGATCGGGAGCGAGAAGTTG
 CTGCGTCCCTCCACACGCTTTCATTGGCGACGTCAAAGGCACTACCGTTGGAGGATGTATCATGTACCTCCAGATAC
 CTCCCGCGCTCCTTCCCAAGGAATCAGTTGCGGTAGTCGGGTTTCAAGCCCCGAAAAGCCGGAAGCATTGTGCAC
 ACTGACAGATGTGACCTCCAGATCTTGAAGCCTATCTCACCCGAGACCCAGTCCAAGTGTGGAAAATGATGTTGG
 ACTTCAAGAAAGTTCGACTAATGGTCTGGAAGACAAAACAGCCTATTTCCAATTGAAGGTGCTATTTACCTGGTAT
 CAGCTTGCCAGCTATGCCTCGTACATCCGTGTTCCCGTCAACTCTACGGTGTACTTGGACCCCTGATGGGCCCGCCCT
 TTGCAACAGGAGAGTCGTGGGTCCACCACTGGGGGCTGACCTCGCGGTACCCCTTATGATTACGGCGCTAAATTA
 TCCTGTCTAGCGGTACCATGGTGAAATGCCCCCGGATACAAAATCTGGCGTGCGCGGAGTTCTGTTGGATGACCCA
 GTTAAGTACAAACATACCTGGGGTTTGAATCGGATACAGCGTATCTGTATGAGTTCACCGGAAACGGTGAGGACTGGGA
 GGATTACAATGATGCGTTTCTGCGCGCCAGGAAGGAAAATTTATAAGGCACTGCCACAGCTTGAAGTTTATTTTC
 CCCCCGGCCCTGTCTTGAACCACTTAGGCTGAATTGAAATGAAATGGGGTCCATGCAAAGCCTTTTGAACAAAT
 GGCCAACTTTTGTGGATGCTTTCACGGAGTCTTGGTGTCCATTGTTGATATCATTATTTTTGGCCATTTGTTTGG

Fig. 11-4

CTTCAACCATCGCCGGTTGGCTGGTGGTCTTTTGCATCAGATTGGTTTGCTCCGCGATACTCCGTACGCGCCCTGCCATTC
 ACTCTGAGCAATTACAGAAGATCTTATGAGGCCCTTTCTTTCCAGTGCCAAGTGGACATTTCCACCTGGGGAACATAACA
 TCCTTTGGGGATGCTTTGGCACCATAAGGTGTCAACCCTGATTGATGAAATGGTGTGCGCTCGAATGTACCGCATCATGG
 AAAAAGCAGGGCAGGCTGCCTGGAAACAGGTGGTGAGCGAGGCTACGCTGTCTCGCATTAGTAGTTTGGATGTGGTGGCT
 CATTTTCAGCATCTAGCCGCCATTGAAGCCGAGACCTGTAATATTTTGGCCTCCCGCTGCCCATGTACACAACCTGCG
 CATGACAGGGTCAAATGTAACCATAGTGTATAATAGCACTTTGAATCAGGTGTTTGCTATTTTCCAACCCCTGGTTCCC
 GGCCAAAGCTTCATGATTTTCAGCAATGGTTAATAGCTGTACATTCCTCCATATTTTCTCTGTTGCAGCTTCTTGACT
 CTTTTGTTGTGCTGTGGTTGCGGGTTCCAATACTACGTACTGTTTTTGGTTTCCGCTGGTTAGGGGCAATTTTCTTTC
 GAACTCACAGTGAATTACACGGTGTGTCCACCTTGCTCACCCGGAAGCAGCCACAGAGATCTACGAACCCGGTAGGTC
 TCTTTGGTGCAGGATAGGGTATGACCGATGTGGGAGGACGATCATGACGAGCTAGGGTTATGATACCGCCTGGCCTCT
 CCAGCGAAGGCCACTTGACTGGTGTACGCTGGTTGGCGTTCTTGCTTTCAGCTACACGGCCAGTTCATCCCGAG
 ATATTCGGGATAGGGAATGTGAGTCGAGTTTATGTTGACATCAACATCAACTCATCTGCGCCGAACATGACGGGCAGAA
 CACCACCTTGCTCTGTCATGACAACATTTAGCCGTGTTTACAGCTTATTACCAACATCAAGTCGACGGCGCAATTGGT
 TTCACCTAGAATGGCTTCGTCCCTCTTTTCTCTGTTGGTTTAAATGTCTTGGTTTCTCAGGCGTTCCGCTGCA
 AACCATGTTTCAGTTCGAGTCTTGACAGATATTAAGACCAACACCACCGCAGCGCAAGCTTGTGTCTCCAGACATC
 AGTTGCCCTTAGGCATCGGACTCGGCTCTGAGGCGATTGCAAAATCCCTCAGTGCCGTACGGCGATAGGGACACCCGT
 GTATGTTACCATCACAGCAATGTGACAGATGAGAATTATTTACATTCCTCTGATCTCCTCATGCTTCTTCTTGCTTT
 TCTATGCTTCTGAGATGAGTGAAAAGGGATTAAAGGTGGTATTGGCAATGTGTGAGGCATCGTGGCTGTGTGTGAAT
 TTTACCAGCTACGTCCAACATGTCAAGGAGTTTACCAACGCTCCCTGGTGGTGCAGCATGTGCGGTTGCTCCATTTCT
 GACACCTGAGACCATGAGGTGGGCAACTGTTTACGCTGTCTTTTGCCATTCTGTTGGCAATTTGAATGTTAAGTATG
 TTGGAGAAATGCTTGACCGCGGGTGTGCTCGCGATTGCTTTCTTGTGGTGTATCGTGCCGTTCTGTTTTGCTGTGCT
 CGCCAACGCCAGCAACGACAGCAGCTCCATCTACAGCTGATTTACAACCTGACGCTATGTGAGCTGAATGGCACAGATT
 GGTAGCTAAACAAATTTGATTGGGAGTGGAGAGTTTGTGATCTTTCCGTTTTGACTCACATTGTCTCTATGGTGCC
 CTCACTACCAGCCATTTCTTGACACAGTCGTTTGTGCTACTGTGTCTACCGCCGGGTTTGTTCACGGGCGGTATGCTCT
 AAGTAGCATCTACGCGTCTGTGCCCTGGCTGCGTTGACTTGCTTCGTCATTAGGTTTGCAAAGAATTGCATGCTCTGGC
 GCTACGCGTGTACCAGATATACCAACTTTCTTCTGGACACTAAGGGCAGACTCTATCGTTGGCGGTGCGCTGTATCATA
 GAGAAAAGGGGCAAAGTTGAGGTGGAAGTCATCTGATCGACCTCAAAGAGTTGTGCTTGATGGCTCCGTGGCAACCC
 TATAACCAGAGTTTCAGCGGAACAATGGGGTCGTCTTAGATGACTTCTGTACGATAGCACGGCTCCACAAAAGGTGCT
 TTTGGCGTTTTCTATTACCTACACGCCAGTGATGATATATGCCCTAAAGGTGAGTCGCGGCCGACTGCTAGGGCTTCTGC
 ACCTTTTGATCTTCTGAATTGTGCTTTCACCTTGGGTACATGACTTTCGCGCACTTTCAGAGTACAAATAAGGTGCGG
 CTCATATGGGAGCAGTAGTTGCACTCCTTTGGGGGTGTACTCAGCCATAGAAACCTGGAAATTCATCACCTCCAGATG
 CCGTTTGTGCTTGCTAGGCCGCAAGTACATTCTGGCCCTGCCACCACGTTGAAAGTGCCGAGGCTTTCATCCGATTG
 CGGCAATGATAACACGCATTTGTGCTCGGCGTCCCGCTCCACTACGGTCAACGGCACATTGGTGCCCGGGTTAAAA
 AGCCTCGTGTGGGTGGCAGAAAAGCTGTTAAACAGGGAGTGGTAAACCTTGTCAAATATGCCAAATAACAACGGCAAGC
 AGCAGAAGAGAAAGAGGGGGATGGCCAGCCAGTCAATCAGCTGTGCCAGATGCTGGGTAAGATCATCGCTCAGCAAAAC
 CAGTCCAGAGGCAAGGGACCGGGAAAGAAAATAAGAAGAAAAACCCGGAGAAGCCCCATTTTCTCTAGCGACTGAAGA
 TGATGTGAGACATCACTTTACCCCTAGTGAGCGGCAATTGTGTCTGTGCTCAATCCAGACCGCCTTAAATCAAGGCGCTG
 GGACTTGACCCCTGTGAGATTAGGGAGGATAAGTTACACTGTGGAGTTTAGTTTGCCTACGCATCATACTGTGCGCCTG
 ATCCGCGTCACAGCATCACCTCAGCATGATGGGCTGGCATTCTGAGGCATCTCAGTGTTTGAATTGGAAGAATGTGTG
 GTGAATGGCACTGATTGACATTGTGCTCTAAGTCACCTATTCAATTAGGGCGACCGTGTGGGGTGAGATTTAATTGGC
 GAGAACCATGCGGCCGAAATTAATAAAAAA

Fig. 1J-1

>V7-Nsp2d633-726.seq

ATGACGTATAGGTGTTGGCTCTATGCCTTGGCATTGTATTGTACAGGAGCTGTGACCATTGGCACAGCCAAAACCTTGCT
 GCACAGAAACACCCCTTCTGTGATAGCCTCCTTCAGGGGAGCTTAGGGTTTGTCCCTAGCACCTTGCTTCCGGAGTTGCAC
 TGCTTTACGGTCTCTCCACCCCTTTAACCATGTCTGGGATACCTGATCGGTGCACGTGTACCCCAATGCCAGGGTGT
 ATGGCGGAGGGCCAAGTCTACTGCACACGATGCCTCAGTGCACGGTCTCTCCTTCCCCTGAACCTCCAGGTTTCTGAGCT
 CGGGGTGCTAGGCTATTCTACAGGCCGAAGAGCCACTCCGGTGGACGTTGCCACGTGCATTCCCACTGTTGAGTGCT
 CCCCAGCGGGGCTGCTGGCTTCTGCAATCTTCCAATCGCACGAATGACCAGTGGAACCTGAACCTCCAACAAAGA
 ATGGTACGGGTGCGAGCTGAGCTTTACAGAGCCGGCCAGCTACCCCTGCAGTCTTGAAGGCTCTACAAGTTATGAACG
 GGGTTGCCGCTGGTACCCATTGTTGGACCTGTCCCTGGAGTGGCCGTTTTCGCCAATTCCTACATGTGAGTGATAAAC
 CCTTCCCGGGAGCAACTCAGTGTGTACCAACCTGCCGCTCCCGCAGAGACCAAGCCTGAAGACTTTTGGCCCTTTGAG
 TGTGCTATGGCTACTGTCTATGACATTGGTCATGACGCCGTATGTATGTGGCCGAAAGGAAAGTCTCCTGGGCCCTCG
 TGGCGGGATGAAGTGAAATTTGAAGCTGTCCCGGGGAGTTGAAGTTGATTGCGAACCGGCTCCGCACCTCCTTCCCGC
 CCCACCACACAGTGGACATGTCTAAGTTCGCCCTCACAGCCCTGGGTGTGGTGTCTATGCGGGTGAACGCCAACAC
 GGCTGCCTTCCCGCTGACACTGTCCCTGAAGGCAACTGCTGGTGGAGCTTGTGACTTGCTTCCACTGGAAGTTCAGAA
 CAAAGAAATTCGCGATGCTAACCAATTTGGCTACCAGACCAAGCATGGTGTCTCTGGCAAGTACCTGCAGCGGAGGCTGC
 AAGTTAATGGTCTCGAGCAGTAAGTACCTAACCGGACCTATCGTCGTACAGTACTTCTCCGTTAAGGAGAGTTGGATC
 CGCCATTTGAACTGGCGGGAGAACCCAGTACTCTGGGTTTGAGGACCTCTCAGAATAAGGGTTGAGCCTAACACGTC
 GCCATTGGCTGACAAGGAAGAAAAATTTCCGGTTTGGCAGTCACAAGTGGTACGGCGCTGGAAAGAGAGCAAGAAAAAG
 CACGCTTGTGCGACTGTACAGTCGCTGGCGCGCTTTGTCCGTTCTGTGAACCCCGCAGGCCAAGGAGCACGAGGTT
 GCGGGCGCAACAAGGCTGAGCACCTCAAACTACTCTCCCGCTGCCGAAGGGAATTGTGGTTGGCACTGCATTTCCGC
 CATCGCAACCGGATGGTGAATTCAAATTTGAAACACCTTCCGAAAGAGTGAGACCTCCAGATGACTGGGCTACTG
 ACGAGGATCTGTGAATGCCATCCAAATCCTCAGACTCCTCGCGCCTTAGACAGGAACGGTGTGTACTAGCGCCAAG
 TACGTACTTAAGCTGGAAGGTGAGCATTGGAAGTGTACTGTGACCCCTGGGATGTCCCTTCTTGTCTCCTCTTGAATG
 TGTTACGGGCTGTTGTGGGCACAAGGGCGGTCTTGGTTCGCCAGATGCAGTCGAGGTCTCCGGATTGACCCTGCCTGCC
 TTGACCGGCTGGCTGAGGTGATGCACCTGCCTAGCAGTGTATCCAGCCGCTCTGGCCGAAATGTCTGGCGATTCCGAT
 CGTTCGGCTTCTCCGGTACCACCGTGTGGACTGTTTCGAGTTCCTTGGCCGTACAGCGGAGGGAATCACCTGACCA
 AGTGCGCTTAGGGAATATATCAGCCTTTGTGAGGTGATTGAGGACTGCTGCTGTTCCAGAACAAAAACCAACCGGTCA
 CCCCAGGAGGCTCGCAGCAAGATTGACCTGTACCTCCGTGGTGCAACAAATCTGAAGAATGCTTGGCCAGGCTTGAG
 AAAGCGCGCCGCCACCGTAATCGACACCTCCTTGTATTGGGATGTTGTGCTCCCTGGGGTTGAGGCGGCAACCCAGAC
 GATCAAGCTGCCCCAGGTCAACCAAGTGTGCTGCTGGTCCCTGTTGTGACTCAAAAGTCTTGGCAACAACCTCGGTCC
 CCTGACCGCCTTTTCACTGGCTAACTACTACTACCGTGCGCAAGGTGACGAAGTTCGTACCGTGAAAGACTAACCGCC
 GTGCTCTCAAGTTGGAAGGTTGTTGAGAGAATATGGGCTCATGCCAACCGAGCCTGGTCCAGGCCACACTGCC
 ACGCGGCTCGACGAACCAAGACCAAGATGGAGGAGGACTTGTGAACTGGCTAACGCCAGACACTTGGACATGA
 TGGCCTGGGAGTGCAGCAGGTTGACCTAAAACTTGGGTCAAGAACTACCCGCGGTGGACACCACCAACCCCTCCGCCA
 AAAGTTCAGCCTCGAAAAACGAAGCCTGTCAAGAGCTTCCGGAGAGAAAGCCTGTCCCGCCCGCGCAGGAAGGTTGG
 GTCCGATTGTGGCAGCCCGGTTTCAATAGGCGCGATGTCCCTAACAGTTGGGAAGATTGGGTGTTAGTAGCCCTTTG
 ATCTCCCGACCCACCTGAGCCGGCAACACCTTCAAGTGAGCTGGTATTGTGCTCCTACCGCAATGCATCTCAGGCCG
 GCGACACCTTGAGTGAGCCGGCTCAATTCGCGACCTCGCGGAAGTGTGCTCGACCGGTGACACCTTGAGTGAGCC
 GATCCCTGTGCCCGCACCGCGGCTAAGTTTCAGCAGGTGAAAAGATTGAGTTCGGCGGCGGAATCCACCGTACCAGG
 ACGAGCCCTGGATTGTCTGCTTCTCAGAGCTGAATATGAGGCCTCTCCCGCAGCACCGCGCAGAGCGGGGGCGTT
 CTGGGAGTAGAGGGGATGAAGCTGAGGAAACCTGAGTGAAATCTCGGACATGTGCGGTAAACATTAACCTGCGTCCGT
 GTCATCAAGCAGCTCCTTGTCCAGCGTGAGATGTGAGTTGTGATGATGCCTCACACGCTGCACCTTCCGTAGGTGCGG
 AGAGCGACCTTACCATTTGGCTCAGTTGCTACTGAAGATGTTCCACGCATCCTCGAGAAAATAGAAAATGTGCGCGAGATG
 GCCAACAGGGACCTTGGCCTTCTCCGAGGATAAACCGGTAGATGACCAACTGTCAACGACCCCGGATATCGTCGCG
 GAGGCTGACGAGAGCACATCAGTCCGTCCGAGGCACAGGTGGCGCCGGCTCTTTACCGATTGCGGCCCTCAGATG
 GCGCGGATGCGGACGGGGGGGGCGTTTCGGACGGTAAAAAGAAAAGCTGAAAGGCTCTTTGACCAACTGAGCCGTGAG
 GTTTTGTACCTGTCTCCATCTCCCTGTTTCTTCTCACGCCTTTTACCTGGCGGTGTTATTCTCCGGGTGATTG
 GGGTTTGCAGCTTTACTCTATTGTGCCTCTTTTATGTTACAGTTACCCAGCCTTTGGTATTGCTCCCTCTTGGGTG

Fig. 1J-2

TGTTCCTGGGCTCTCTCGGCGCTTCGAATGGGGGTTTTGGCTGCTGGTTGGCTTTTGCTGTTGGTCTGTTCAAGCCT
 GTGTCCGACCCAGTCGGCGCTGCTTGTGAGTTTGACTCGCCAGAGTGTAGAAACATCCTTCATTCTTTGAGCTTCTCAA
 ACCTTGGGACCTGTTCGAGCCTTGTGTGGGCCCCGTCGGTCTCGGTCTTGCCATTCTTGGCAGGTTACTGGGCGGG
 CACGCTGCATCTGGCACTTTTTGCTTAGGCTTGGCATTGTTGCAGACTGTATCTTGGCTGGAGCTTACGTGCTTTCTCAA
 GGTAGGTGTAAAAAGTGTGGGGATCTTGATATAAGAACTGCTCCCAATGAGGTCGCTTTAACGTGTTTCCTTTCACACG
 TGCAGCCAGGTCGTCACTTATCGACCTGTGCGATCGGTTTTGTGCGCCAAAAGGAATGGACCCCATTTTTCTCGCCACTG
 GGTGGCGCGGGTGTGGGCCGGCCGAAGCCCCATTGAGCAACCCTCTGAAAAACCCATCGCGTTTGGCCAGTTGGATGAA
 AAGAAGATTACGGCTAGGACTGTGGTCGCCAGCCTTATGACCCCAACCAAGCCGTAAAGTGCTTGGGGTATTGCAGGC
 GGTGGGGCGATGGTGGCTAAGGCGGTCCCAAAGGTCAAGGTTTCCGCTGTTCCATTCCGAGCCCCCTTCTTCCCA
 CTGGAGTGAAAGTTGACCTGATTGCAGGGTCGTGGTTGACCTGACACTTTACTGCAGCTCTCCGGTCTGGCTACTCC
 ACCACAAACCTCGTCTTGGTGTGGGGGACTTTGCCAGCTGAATGGATTAATAATCAGGCAAAATTTCCAAGCCTTCAGG
 GGGAGGCCACATCTCATGGCTGCCCTGCATGTTGCTGCTCGATGGCTCTGCACATGCTTGTGGGATTATGTGACTG
 CGGTGGGTTCTTGGCGACCGGCCAACGACCCGTGGTGGCTAACCCGTTTGGCGTCCCTGGCTACGGACCTGGCTCT
 CTCTGCACGTCAGATTGTGATTTCCCAACAGGCCTTACCCTGCCCTGACAGCACTTGTGGCGGATTGGTATTCA
 AGAAATTGCTTGGTGTGTTTGTATTTTGTTCATCGGAGGCATGGCTCATAGGTTGAGCTGTAAAGGCTGACATGCTGT
 GTGCTTGTGCAATTGCCAGCTATGTTGGGTACCTTTACCTGGTTGCTTGTGTGTTTCTTGTGTTTGGCTGT
 TTTCTTTCGACCCCTCACCATCCTATGGTGGTGTGTTTCTTGATTCTGTGAATATGCCTTCAGGAATCTTGGCCAT
 GGTGTTGTTGGTTTCTCTTGGCTTCTTGGTCGTTATACTAATGTTGCTGGCCTTGTACCCCTACGACATTATCATT
 ACACCACTGGCCCCCGGCTGTTGCCGCTTGGCTACCGCACCAGATGGGACCTACTTGGCCGCTGTCCGCCGCTGCG
 TTGACTGGCCGACCATGCTGTTACCCCGTCCAGCTTGGGTCTCTTCTTGAGGGTGCTTTCAGAACTCGAAAGCCCTC
 ACTGAACACCGTCAATGTGATCGGGTCTCCATGGGCTCTGGCGGGGTGTTTACCATCGACGGGAAAGTCAAGTGCGTAA
 CTGCCGCATGTCCTTACGGGCAATTGAGCTCGGGTTTCCGGGGTGGCTTCAATCAATGCTTGACTTTGACGTAAAG
 GGAGATTTCGCTATAGCTGATTGCCGAATTGGCAAGGGGCTGCCCAAGACCAATTCTGCACGGATGGATGGACTGG
 CCGTGCCCTATTGGCTAAGATCCTCTGGCGTCGAACCCGGCGTCATTGGAAAAGGATTGCTTCTGCTTACCCGATGTG
 GCGATTCCGGGTCCCACTGATCACCAGGCGGCTGAGCTTGTGCGGCTTACACGGGATCGAATAAACAAGGGGGGGG
 ATTGTTACGCGCCCTCAGGCCAGTTTTGTAATGTGGCACCATCAAGCTAAGCGAATTAAGTGAATTTCTTGTGGGCC
 TAAGGTCCCGCTCGGTGATGTGAAGGTGGCAGCCACATAATTAAGACATAAGCGAGGTGCCTTCAGATCTTGTGCT
 TGCTTGTGCCAAACCTGAACTGGAAGGAGGCTCTCCACCGTCCAACTTCTTGTGTGTTTTTCTCTGTGGAGAATG
 ATGGGACATGCTGGAGCCCTTGGTGTGTGAGTTCTTATTTTGAATGAGGTTCTCCAGCCGTCTGGTCCGGAG
 TGTTTTCTCTTGAATGTTGTGCTATCTGGCTCACGCCATGGTCTGCGCAAGTTCTGATGATCAGGCTTCTGACAG
 CAGCTCTTAACAGGAACAGATGGTCACTTGCCTTTTTCAGCCTCGGTGCAAGTACCGGTTTTGTGCGAGATCTTGGGCC
 ACTCAGGGGATCCGTTGACGGCAGTGATGAATTTGAGCACCTATGCATTCTGCCTCGGATGATGGTTGTGACCTCACC
 AGTCCCAGTGATCAGTGTGGTGTGTCGACCTACTTGGCATATTTGTACTTGTAAAGTACCGTGGCCTGCACCATA
 TCCTTGTGGGATGGAGTGTTCTGCGGCTTCTTCTTGGATACTTTGCCGAGGGAAGTTGAGGGAAGGGGTGTCG
 CAATCTGCGGAATGAATCATGAGTCTGACTGGTGCCCTCGCTATGAGACTCAATGACGAGGACTTGGATTTCTTAT
 GAAATGGAATGATTTTAAAGTCTTGTCTGCGTCCAACATGAGGAATGCAGCGGTCAATTTATCGAGGCTGCCTATG
 CTAAGCACTTAGAGTAGAACTGGCCAGTTGGTGAGGTTGATAAAGTTGAGGTACTTTGGCCAACTTGAAGCTTTT
 GCTGATACCGTGGCACCTCAACTCTCGCCGGTGACATTGTTGTGCTCTCGGCCACACGCTGTTGGCAGTATCTTCGA
 CCTAAAGGTTGGTAGCACAAGCATACCTCCAAGCCATTGAGACCAGAGTCTTGTGGGTCCAAAATGACCGTGGCGC
 GCGTCTGACCCGACCCACGCCCCACCCGACCCGTGCCATCCCCCTCCACCGAAAGTTCTGGAGAATGGCCCC
 AACGCTTGGGGGATGAGGACCGTTTGAATAAGAAGAAGAGGCGCAGGATGGAAGCCCTCGGCATCTATGTTATGGGCGG
 GAAAAATACAGAAATTTTGGGACAAGAATTCGGTGATGTGTTTATGAGGAGGTCCATAATAACACAGATGAGTGGG
 AGTGTCTCAGAGTTGGCGACCTGCCGACTTGAACCTGAGAAGGGAATCTGTGTGGACATGTACCATTGAAAACAAG
 GCTTACCATGTTTACACCTCCCATCTGGTAAGAAGTTCTTGGTCCCGTCAACCCAGAGAATGGAAGAGTCCAATGGGA
 AGCTGCAAAGCTTTCCGTGGAGCAGGCCCTAGGTATGATGAATGTGACGGCGAACTGACTGCCAAAGAACTGGAGAAAC
 TGAAAAGAAATTTGACAACTCAGGGCCTGACTAAGGAGCAGTGTTTAACTGCTAGCCGCCAGCGACTTGACCCGCT
 GTGGTGGCGGGCTTGGTTGTTACTGAAACAGCGGTAATAAGTCAAATTTACAACCGGACCTTACCCTGGGACCT
 GTGAATTTAAAGTGGCCAGTGAGGTTGAGCTAAAAGACGCGGTTGAGCACAAACACCCGGTTGCGAGACCGATCGA

Fig. 1J-3

TGGTGGAGTTGTGCTCCTGCGTTCCGCGGTTCTTCGCTTATAGACGTCTTGATCTCCGGTGCTGATGCATCTCCCAAGT
 TACTTGCCCATCACGGGCCGGGAAACACTGGGATCGATGGCACGCTCTGGGATTTTGAGTCCGAAGCCACTAAAGAGGAA
 GTCGCACTCAGTGCCAAATAATACAGGCTTGACATTAGCGCGGCGACGCTCCTGAAATTGGTCTCCCTTACAAGCT
 GTACCTGTAGGGGTAACCTGAGCGGGTGAAAGGAGTTCTGCAGAATACAAGGTTTGAGACATACCTTACAAAACCC
 CCAGTGACACTGGAAGCCAGTGACGCGGCTGCCTGCCTTACGCCAACGCCACTCCGGTGACTGATGGCGCTCCGTC
 TTGGCCACGACCATGCCCCCGGTTTGAAGTTATATGTACCGACCATACCAGCGTCTGTCTTGATTACCTTGACTCTAG
 GCCTGACTGCCCTAAACAGCTGACAGAGCACGGCTGCGAAGATGCCGCACTGAAAGACCTCTCTAAATATGACTTGCCA
 CCCAAGGCTTTGTTTTACCTGGAGTTCTTCGCTTGTGCGGAAATACCTGTTTGCCCATGTAGGTAAGTGCCCAACCGTT
 CATCGGCCTTCTACTTACCCTGCTAAGAAATCTATGGCTGGAATAAATGGGAACAGGTTCCCAACCAAGGACATTAGAG
 CGTCCCTGAAATCGACGTTCTGTGCGCACAGGCTGTGCGAGAAAACCTGGCAAACTGTACCCCTTGTAATCTTAAGAAAC
 AGTATTGCGGGAAGAAGAAGACTAGGACCATACCTCGGCACCAATAACTTCATCGCACTAGCCACCGAGCAGTGTTGAGT
 GGTGTTACCCAGGGCTTCATGAAAAAGGCGTTTAACTCGCCCATCGCCCTCGGAAAGAAAGTTTAAAGGAGCTACAGAC
 TCCGGTCTCGGGCAGGTGCCTTGAAGCTGATCTCGCATCCTGCGATCGATCCACGCTGCAATTGTCCGCTGGTTTGCCG
 CCAACCTCTTTATGAACCTGCTGTGCTGAAGAGCATCTACCGTCTGACGTGCTGAACCTGCTGCCACGACTTACTGGTC
 ACGCAGTCCGGCGCAGTGACTAAGAGAGGTGGCCTGTGCTGCGGACCCGATCACCTCTGTGTCTAACACCATTTATAG
 TTTGGTGATCTATGCACAGCATATGGTGCTTAGTTACTTCAAAAGTGGTCACCCCATGGCCTTCTGTTCTTACAAGACC
 AGCTAAAGTTTGAGGACATGCTCAAGGTTCAACCCCTGATCGTCTATTCGGACGACCTGTGCTGTATGCCGAGTCTCCC
 ACCATGCCAACTATCACTGGTGGGTTGAACATCTGAATTTGATGCTGGGTTTCAGACGGACCCAAAGAAGACAGCAAT
 AACAGACTCGCCATCATTTCTAGGCTGTAGAATAATAAATGGGCGCCAGCTAGTCCCAACCGTGACAGGATCCTCGCGG
 CCTCGCTATCACATGAAGGCGAGTAATGTTTCTGAATACTATGCCTCAGCGGCTGCAATACTCATGGACAGCTGTGCT
 TGTTTGGAGTATGATCCTGAATGGTTTGAAGAACTTGAGTTGGAATAGCGCAGTGCGCCCGCAAGGACGGCTACAGCTT
 TCCCGGCACGCCGTTCTTCATGTCCATGTGGGAAAACTCAGGTCCAATTATGAGGGGAAGAAGTCGAGAGTGTGCGGGT
 ACTGCGGGGCCCCGGCCCCGTACGCTACTGCTGTGGCCTCGACGTCTGCATTTACCACACCCACTTCCACCAGCATTGT
 CCAGTCACAATCTGGTGTGGCCATCCAGCGGGTCTGGTCTTGTAGTGAGTGCAATCCCTGTAGGGAAAGGCACAAG
 CCTTTAGACGAGGTGCTGGAACAAGTCCCGTATAAGCCCCACGGACCGTTATCATGCAATGTGGAGCAGGGTCTACCC
 CCTTGATCCAGGTAGATACCAACTCGCCGCGGATTAGTCTCTGTAGGCGTGGAATTAGGGGAAATGAAGTTGGACTA
 CCAGACGGTGATTATGTAGCACCGCCTTGCTCCCTACTGCAAGAGATCAACATGGTCGCTGTGCTTCCAATGTATT
 GCGCAGCAGGTTTCATCATCGGCCACCCGGTGTGGGAAACATACTGGCTCCTTCAACAGGTCCAGGATGGTGATGTTA
 TTACACACCAACTCACCAGACCATGCTTGACATGATTAGGGCTTTGGGACGTGCCGGTTCAACGTCCCGGCAGGCACA
 ACGCTGCAATTCCTCGTCCCCCTCCCGCACCGGTCCGTGGGTTTCGATCCTAGCCGGCGGTTGGTGTCTGGCAAGAATT
 CTTCTAGATGAAGCAGCGTATTGCAATCACCTTGATGTTTTGAGGCTTCTAGTAAACTACCTCACCTGTCTAGGAG
 ACTTCAAGCAACTCCACCCAGTGGGTTTTGATTCTCATTGCTATGTTTTGACATCATGCCTCAAACTCAACTGAAGACC
 ATCTGGAGGTTTGGACAGAATATCTGTGATGCCATTAGCCAGATTACAGGGACAAAACCTCATGTCCATGGTCAACACAAC
 CCGTGTGACCTACGTGGAAAAACCTGTAGGTATGGGCAGGTCTCACCCCTACCAAGGGACCGAGAGGACGACGCCA
 TCACTATTGACTCCAGTCAAGGCGCCACATTGATGTGGTTACATTGCATTTGCCCACTAAAGATTCACTCAACAGGCAA
 AGAGCCCTTGTTGCTATCACCAGGGCAAGACAGCTATCTTTGTGTATGACCCACACAGGCAGCTGCAGGGCTTGTTGA
 TCTTCCTGCAAAAGGCACGCCCGTCAACCTCGCAGTGCACTGCGACGGGCAGCTGATCGTGCTGGATAGAAATAACAAAG
 AATGCACGGTTGCTCAGGCTTAGGCAACGGGGATAAATTTAGGGCCACAGACAAGCGTGTGTAGATTCTCTCCGCGCC
 ATTTGTGCTGATCTAGAAGGGTCGAGCTCTCCGCTCCCCAAGGTGCGACACAACCTGGGATTTTATTTCTACCTGATTT
 AACACAGTTTGCTAAACTCCAGTAGAACTTGACCTCACTGGCCCGTGGTGTCAACCCAGAACAAATGAAAAGTGGCCGG
 ATCGGCTGGTTGCCAGCCTTCGCCCTATCCATAAATACAGCCGCGCGTGCATCGGTGCCGGCTATATGGTGGGCCCTTCG
 GTGTTTCTAGGCACTCCTGGGTCGTGTACTATCTCACAAAATTTGTTAAGGGCGGGCTCAAGTGCTTCCGGAGAC
 GGTTTTAGCACCAGGCGCAATTGAGGTAGACTGCCGGGAATATCTTGATGATCGGGAGCGAGAAGTTGCTGCGTCCCTCC
 CACACGCTTTTATTGGCGACGTCAAAGGCACTACCGTTGGAGGATGTCATCATGTACCTCCAGATACCTCCCGCGCTC
 CTTCCCAAGGAATCAGTTGCGGTAGTCGGGGTTTCAAGCCCCGAAAAAGCCGGAAGCATTGTGCACACTGACAGATGT
 GTACCTCCAGATCTTGAAGCCTATCTCCACCCGGAGACCCAGTCCAAGTGCTGGAAAATGATGTTGGACTTCAAAGAAG
 TTCGACTAATGGTCTGGAAGACAAAACAGCCTATTTCCAACCTGAAGGTGCTATTTACCTGGTATCAGCTTGCCAGC
 TATGCCTCGTACATCCGTGTTCCCGTCAACTCTACGGTGTACTTGACCCCTGCATGGGCCCCGCCCTTGCAACAGGAG

Fig. 1J-4

AGTCGTCGGGTCCACCCACTGGGGGGCTGACCTCGCGGTACCCCTTATGATTACGGCGCTAAAATTATCTGTCTAGCG
 CGTACCATGGTGAAATGCCCCCGGATACAAAATCTGGCGTGC GCGGAGTTCTCGTTGGATGACCCAGTTAAGTACAAA
 CATACCTGGGGGTTTGAATCGGATACAGCGTATCTGTATGAGTTCACCGGAAACGGTGAGGACTGGGAGGATTACAATGA
 TGC GTTCGTGCGCGCCAGGAAGGAAAAATTTATAAGGCCACTGCCACCAGCTTGAAGTTTTATTTTCCCCGGGCCCTG
 TCATTGAACCAACTTTAGGCCTGAATTGAAATGAAATGGGGTCCATGCAAAGCCTTTTGACAAAATTGGCCAACTTTTT
 GTGGATGCTTTACGGAGTTCTTGGTGTCCATTGTTGATATCATTATATTTTTGGCCATTTTGTTTGGCTTCACCATCGC
 CGTTGGCTGGTGGTCTTTTGCATCAGATTGGTTTGTCTCCGCGATACTCCGTACGCGCCCTGCCATTCACTCTGAGCAAT
 TACAGAAGATCTTATGAGGCCTTTCTTTCCAGTGCCAAGTGGACATCCACCTGGGGAATAACATCCTTTGGGAT
 GCTTTGGCACCATAAGGTGTCAACCCTGATTGATGAAATGGTGTGCGTCGAATGTACCGCATCATGGAAAAAGCAGGCG
 AGGCTGCCTGGAAACAGGTGGTGAAGCGAGGCTACGCTGTCTCGCATTAGTAGTTTGGATGTGGTGGCTCATTTTCAGCAT
 CTAGCCGCCATTGAAGCCGAGACCTGTAATATTTGGCCTCCCGGCTGCCCATGTACACAACCTGCGCATGACAGGGTC
 AAATGTAACCACTAGTGATATAATAGCACTTTGAATCAGGTGTTTGTATTTTTCCAACCCCTGGTTCCCGGCCAAAGCTTC
 ATGATTTTCAGCAATGGTTAATAGCTGTACATTCTCCATATTTCTCTGTTGAGCTTCTGTACTCTTTTGTGTG
 CTGTGGTTGCGGGTTCCAATACTACGTACTGTTTTGGTTTCCGCTGGTTAGGGGCAATTTTTCTTTCGAATCAGAGT
 AATTACACGGTGTGTCCACCTTGCTCACC CGCAAGCAGCCAGAGATCTACGAACCCGGTAGGTCTCTTTGGTGCAG
 GATAGGGTATGACCGATGTGGGAGGACGATCATGACGAGCTAGGGTTTATGATACCGCTGGCTCTCCAGCGAAGGCC
 ACTTGACTGGTGTTCACCGCTGGTTGGCGTTCTTGCTTTCAGCTACACGGCCAGTTCCATCCCGAGATATTCGGGATA
 GGGAAATGTGAGTCGAGTTTATGTTGACATCAACATCAACTCATCTGCGCCGAACATGACGGGCAGAACACCACTTGCC
 TCGTCATGACAACATTTACGCCGTGTTTCAGACCTATTACCAACATCAAGTCGACGGCGGCAATTGGTTTCACTAGAAT
 GGCTTCGTCCCTCTTTTCTCGTGGTTGGTTTTAAATGTCTCTTGGTTTCTCAGGCGTTGCGCTGCAAACCATGTTTCA
 GTTCGAGTCTTGAGATATTAAGACCAACACCACCGCAGCGGCAAGCTTTGCTGTCTCAGACATCAGTTGCCTTAGG
 CATCGCGACTCGGCCTCTGAGGCGATTGCGAAAATCCCTCAGTGCCGTACGGCGATAGGGACACCCGTGTATGTTACCAT
 CACAGCCAATGTGACAGATGAGAATTATTTACATCTCTGATCTCCTCATGCTTCTTCTTGCCTTTCTATGCTTCTG
 AGATGAGTGAAAAGGGATTAAAGGTGGTATTTGGCAATGTGTGAGGCATCGTGGCTGTGTGTCAATTTTACCAGTAC
 GTCCAACATGTCAAGGAGTTTACCAACGCTCCCTGGTGGTGCACCATGTGCGGTTGCTCCATTTTCATGACACTGAGAC
 CATGAGGTGGGCACTGTTTGGCTGTCTTTTGGCAATCTGTTGGCAATTTGAATGTTTAAAGTATGTTGGAGAAATGC
 TTGACCGCGGGCTGTTGCTCGCGATTGCTTTCTTGTGGTGTATCGTGCCGTTCTGTTTGTGTGCTCGCCAACGCCAG
 CAACGACAGCAGCTCCCATCTACAGCTGATTTACAACCTGACGCTATGTGAGCTGAATGGCAGAGATTGGCTAGCTAACA
 AATTTGATTGGGCACTGGAGAGTTTGTGCTTTCCCGTTTTGACTCACATTGTCTCTATGGTGCCTCACTACCAGC
 CATTTCTTGACACAGTCGCTTATGCTACTGTGTCTACCGCCGGTTTGTTCACGGGCGGTATGTCTAAGTAGCATCTA
 CGCGGTCTGTGCCCTGGCTGCGTTGACTTGCTTCGTATTAGGTTTGCAAAGAATTGCATGTCTGGCGCTACGCGTGTA
 CCAGATATACCAACTTTCTTCTGGACACTAAGGGCAGACTCTATCGTTGGCGGTGCGCTGTATCATAGAGAAAAGGGG
 AAAGTTGAGGTGGAAGGTATCTGATCGACCTCAAAGAGTTGTGCTTGATGGCTCCGTGGCAACCCCTATAACCAGAGT
 TTCAGCGGAACAATGGGGTCGCTTAGATGACTTCTGTACGATAGCAGGCTCCACAAAAGGTGCTTTTGGCGTTTTT
 TATTACCTACACGCCAGTGATGATATATGCCCTAAAGGTGAGTCGCGGCGACTGCTAGGGCTTCTGCACCTTTTGATCT
 TCCTGAATTTGCTTTTACCTTCGGGTACATGACTTTCGCGCACTTTCAGAGTACAAATAAGGTGCGCTCACTATGGGA
 GCAGTAGTTGCACTCCTTTGGGGGTGTA CTAGCCATAGAACTGGAAATTCATCACCTCCAGATGCCGTTTGTGCTT
 GCTAGGCCGCAAGTACATTCTGGCCCTGCCACCACGTTGAAAGTGCCGAGGCTTTCATCCGATTGCGGCAATGATA
 ACCACGCATTTGTGTCGCGCGTCCCGGCTCCACTACGGTCAACGGCACATTGGTGCCCGGGTTAAAAAGCCTCGTGTG
 GGTGGCAGAAAAGCTGTTAAACAGGAGTGGTAAACCTGTCAAATATGCCAAATAACAACGGCAAGCAGCAGAGAGAA
 AGAAGGGGGATGGCCAGCCAGTCAATCAGCTGTGCCAGATGCTGGGTAAGATCATCGCTCAGCAAAACAGTCCAGAGGC
 AAGGGACCGGGAAGAAAAATAAGAAGAAAAACCCGGAGAAGGCCATTTTCTCTAGCGACTGAAGATGATGTGAGACA
 TCACCTTACCCCTAGTGAGCGCAATTGTGTCTGTCGTAATCCAGACCGCTTAAATCAAGGCGCTGGGACTTGACCCC
 TGTGAGATTGAGGGAGGATAAGTTACACTGTGGAGTTAGTTTGCCTACGCATCATACTGTGCGCTGATCCGCGTCACA
 GCATCACCTCAGCATGATGGCTGGCATTCTTGAGGCATCTCAGTGTGTTGAATTGGAAGAATGTGTGGTGAATGGCACT
 GATTGACATTGTGCTCTAAGTCACCTATTCAATTAGGGCGACCGTGTGGGGTGAGATTTAATTGGCGAGAACCATGCG
 GCCGAAATTAATAAAAAA

Fig. 1K-1

>V7-Nsp2d543-726.seq

ATGACGTATAGGTGTTGGCTCTATGCTTGGCATTGTGATTGTGACGAGCTGTGACCATTTGGCAGACGCCAAACTTGCT
GCACAGAAACACCTTCTGTGATAGCCTCCTTCAGGGGAGCTTAGGGTTTGTCCCTAGCACCTTGCTTCCGGAGTTGCAC
TGCTTTACGGTCTCTCCACCCCTTTAACCATGTCTGGGATACTTGATCGGTGCACGTGTACCCCCAATGCCAGGGTGT
ATGGCGGAGGGCCAAGTCTACTGCACACGATGCTCAGTGCACGGTCTCTCCTTCCCCTGAACCTCCAGGTTTCTGAGCT
CGGGGTGCTAGGCCTATTCTACAGGCCCGAAGAGCCACTCCGGTGGACGTTGCCACGTGCATTCCCCACTGTTGAGTGTCT
CCCCCGCCGGGGCCTGCTGGCTTTCTGCAATCTTCCAATCGCACGAATGACCAGTGGAAACCTGAACCTCCAACAAAGA
ATGGTACGGGTGCGAGCTGAGCTTTACAGAGCCGGCCAGCTCACCCCTGCAGTCTTGAAGGCTCTACAAGTTTATGAACG
GGGTTGCCGCTGGTACCCATTGTTGGACCTGTCCCTGGAGTGGCCGTTTTCGCCAATTCCCTACATGTGAGTGATAAAC
CCTTCCCGGAGCAACTCACGTGTTGACCAACCTGCCGCTCCCGCAGAGACCCAAGCCTGAAGACTTTTGCCCTTTGAG
TGTGCTATGGCTACTGTCTATGACATTGGTCATGACGCCGTGATGATGTGGCCGAAAGGAAAGTCTCCTGGGCCCTCG
TGCGGGGATGAAGTGAATTTGAAGCTGTCCCGGGGAGTTGAAGTTGATTGCGAACCGGCTCCGCACCTCCTTCCCGC
CCACCACACAGTGGACATGTCTAAGTTCGCTTACAGCCCTGGGTGTGGTGTCTATGCGGGTGAACGCCAACAC
GGCTGCCTTCCCGCTGACACTGTCCCTGAAGGCAACTGCTGGTGGAGCTGTTTGACTTGCTTCCACTGGAAGTTCAGAA
CAAAGAAATTCGCCATGCTAACCAATTTGGCTACCAGACCAAGCATGGTGTCTCTGGCAAGTACCTGCAGCGGAGGCTGC
AAGTTAATGGTCTCCGAGCAGTAACGACCTAAACGGACCTATCTGCTACAGTACTTCTCCGTTAAGGAGAGTTGGATC
CGCCATTTGAAACTGCGGGAGAACCCAGCTACTCTGGGTTTGAGGACCTCCTCAGAATAAGGGTTGAGCCTAACACGTC
GCCATTGGCTGACAAGGAAGAAAAAATTTCCGGTTTGCGAGTACAAGTGGTACGGCGCTGGAAAGAGAGCAAGAAAAAG
CACGCTCTTGTCGACTGCTACAGTCTGCTGGCCGCGCTTTGTCCGTTCTGTGAAACCCGGCAGGCCAAGGAGCACGAGGTT
GCCGGCGCCAACAAGGCTGAGCACCTCAAACTACTCCCCGCTGCCGAAGGGAATTGTGGTTGGCACTGCATTTCCGC
CATCGCCAACCGGATGGTGAATTCAAATTTGAAACACCCCTTCCCGAAAGAGTGAGACCTCCAGATGACTGGGCTACTG
ACGAGGATCTTGTAATGCCATCCAAATCTCAGACTCCCTGCGGCTTAGACAGGAACGGTGCTTGTACTAGCGCCAAG
TACGTACTTAAGCTGGAAGGTGAGCATTGGACTGTCACTGTGACCCCTGGGATGTCCCTTCTTGCTCCCTCTGAATG
TGTTCAGGGCTGTTGTGGGCAAGGGCGGTCTTGGTCCCAGATGCAGTCGAGGTCTCCGGATTTGACCTGCCTGCC
TTGACCGGCTGGCTGAGGTGATGCACCTGCCTAGCAGTGCTATCCAGCCGCTCTGGCCGAAATGTCTGGCGATTCCGAT
CGTTCCGCTTCTCCGGTCACCACCGTGTGGACTGTTTCGAGTCTTTGCCCGTCACAGCGGAGGGAATCACCTGACCA
AGTGCCTTAGGGAAAATTATCAGCCTTTGTGAGGTGATTGAGGACTGCTGCTGTTCCAGAACAAAACCAACGGGTCA
CCCCGGAGGAGGTGCGAGCAAAGATTGACCTGTACCTCCGTGGTGCAACAAATCTGAAGAATGCTTGGCCAGGCTTGAG
AAAGCGCGCCGCCACGCGTAATCGACACCTCCTTTGATTGGGATGTTGTGCTCCCTGGGGTTGAGGCGGCAACCCAGAC
GATCAAGCTGCCCCAGGTCAACCAGTGTCTGTCTGGTCCCTGTTGTGACTCAAAAGTCTTGGACAACAACCTCGGTCC
CCCTGACCGCTTTTACTGGCTAACTACTACTACCTGCGCAAGGTGACGAAGTTCGTACCGTGAAAGACTAACCGCC
GTGCTCTCAAGTTGGAAAAGTTGTTGAGAGAAGATATGGGCTCATGCCAACCGAGCCTGGTCCACGGCCCACTGCC
ACGCGGGCTCGACGAACCAAGACCAGATGGAGGAGGACTTGCTGAAACTGGCTAACGCCACAGACGACTTCGGACATGA
TGCCCTGGGAGTCGAGCAGGTTGACCTAAAACTTGGTCAAGAACTACCCGCGGTGGACACCACCAACCCCTCCGCCA
AAAGTTCAGCCTCGAAAAACGAAGCCTGTCAAGAGCTTGCCGGAGAGAAAGCCTGTCCCCGCCCGCGCAGGAAGTTGG
GTCCGATTGTGGCAGCCCGGTTTCATTAGCGCGCGATGTCCCTAACAGTTGGGAAGATTGCTGTTAGTAGCCCCTTTG
ATCTCCCGACCCACCTGAGCCGGCAACACCTTCAAGTGAGCTGGTGATTGTGCTTACCGCCAATGCATCTTCAGGCCG
GTGACACCTTGTAGTGAGCCGGCTCAATTTCCCGCACCTCGCGGAACTGTGCTCGACCGGTGACACCTTGAGTGAGCC
GCTGTAGTTTGTGATGATGCCTCACACGCTGCACCTTCCGTAGTGTGCGGAGAGCGACCTTACCATTGGCTCAGTTGCTA
CTGAAGATGTTCCACGCATCCTCGAGAAAAATAGAAAATGTGCGCGAGATGGCCAACCAAGGACCTTGCCCTTCTCCGAG
GATAAACCGGTAGATGACCAACTTGTCAACGACCCCGGATATCGTCGCGGAGGCTGACGAGAGACATCAGTCCGTC
CGCAGGCACAGGTGGCGCCGGCTTTTTACCGATTTGCCGCTTTCAGATGGCGCGGATGCGGACGGGGGGGGCCGTTTC
GGACGGTAAAAAGAAAAGCTGAAAGGCTCTTTGACCAACTGAGCCGTGAGGTTTTTGACCTCGTCTCCCATCTCCCTGTT
TTCTTCTACGCTTTTTTACCCTGGCGGTGGTTATTCTCGGGTGATTGGGGTTTTGCAGCTTTTACTCTATTGTGCCT
CTTTTTATGTTACAGTTACCCAGCCTTGGTATTGCTCCCTCTTGGGTGTGTTTTCTGGGTCTTCTCGGCGGTTTCGAA
TGGGGGTTTTTGGCTGCTGGTTGGCTTTTGTGTTGGTCTGTTCAAGCCTGTGTCCGACCCAGTCGGCGCTGCTTGTGAG
TTTGACTCGCCAGAGTGAGAAACATCTTCATTCTTTGAGCTTCTCAAACCTTGGGACCTGTTGCGAGCCTTGTGTG
GGGCCCCGTGGTCTCGGTCTTGCCATTCTTGGCAGGTTACTGGGCGGGGCACGCTGCATCTGGCACTTTTGTCTAGGC

Fig. 1K-2

TTGGCATTTGTTGCAGACTGTATCTTTGGCTGGAGCTTACGTGCTTTCTCAAGTAGAGTGTAAAAAGTCTGGGATCTTG
ATAAGAACTGCTCCCAATGAGGTGCGTTTTAACGTGTTTCTTTTACACGTGCGACCAAGGTGCTCACTTATCGACCTGTG
CGATCGGTTTTGTGCGCCAAAGGAATGGACCCATTTTTCTGCCACTGGTGGCGGGGTGCTGGGCGGCCGAAGCC
CCATTGAGCAACCTCTGAAAAACCATCGCGTTTGCCAGTTGGATGAAAAGAAGATTACGGCTAGGACTGTGGTCGCG
CAGCCTTATGACCCCAACCAAGCCGTAAGTGCTTGCGGGATTGCAAGCGGGTGGGGCGATGGTGGCTAAGGCGGTCCC
AAAAGTGGTCAAGGTTTTCCGCTGTTCAATTCGAGCCCCCTTCTTTCCCACTGGAGTGAAAGTTGACCTGATTGCAAGG
TCGTGGTTGACCTGACACTTTCACTGCAGCTCTCCGGTCTGGCTACTCCACCACAAACCTCGTCTTGGTGTGGGGGAC
TTTGCCAGCTGAATGGATTAAAAATCAGGCAAAATTTCAAGCCTTCAGGGGAGGCCACATCTCATGGCTGCCCTGCA
TGTTGCTGCTCGATGGCTCTGCACATGCTTGCTGGGATTTATGTACTGCGGTGGGTTCTTGCGGCACCGCACCAAGC
ACCCGTGGTGCCTAACCCGTTTGGCTCCCTGGCTACGGACCTGGCTCTCTGCACGTCCAGGATGTGCATTTCCCAA
CACGGCTTACCCTGCCCTGACAGCACTTGTCGGGGATTCCGGATTCAAGAAATTCCTTGGTGTGTTTTGATTTTTGT
TTCCATCGGAGGACATGGCTCATAGGTTGAGCTGTAAAGCTGACATGCTGTGTGCTTGTGCAATTGCCAGCTATGTTT
GGTGACCTCTTACCTGTTGTTTTGTGTTTTCTTGCTGGTGGCTGTTTTCTTGTGACCCCTCACCATCTATGG
TTGGTGTTTTTCTGATTTCTGTGAATATGCCTTCAGGAATCTGGCCATGGTGTGTTGGTTTTCTTGTGGCTCTTG
TCGTTATACTAATGTTGCTGGCCTTGACCCCTACGACATTCATCATTACACCAGTGGCCCCCGGGTGTGCGCCT
TGGCTACCGCACCAGATGGGACCTACTGGCCGCTGTCCGCCGCTGCGTTGACTGGCCGCACCATGCTGTTACCCG
TCCCAGCTTGGGTCTCTTCTGAGGGTCTTTAGAACTCGAAAGCCCTCACTGAACACCGTCAATGTGATCGGGTCTC
CATGGGCTCTGGCGGGGTGTTTACCATCGACGGGAAAGTCAAGTGCCTAACTGCCGCACATGCTTACGGGCAATTCAG
CTCGGGTTTCCGGGGTGGCTTCAATCAAATGCTGACTTTGACGTAAGGGAGATTTCGCTATAGCTGATTGCCCGAAT
TGGAAGGGGTGCCCCAAAGACCAATTCGCACGGATGGATGGAATGGCCGTGCTTATTGGTCAACATCTCTGGCT
CGAACCCGGCGTCAATTGAAAAAGGATTGCGCTTCTGCTTACCGCATGTGGCGATTCCGGGTCCCGAGTGATCACCAGG
CCGGTGAGCTTGTGCGGTTTACACGGGATCGAATAAACAGGGGGGGCTGTTTACGCGCCCTCAGGCCAGTTTTGT
AATGTGACCCATCTCAAGCTCAGCGAATTAAGTGAATCTTTGCTGGGCTAAGGTCCCGCTCGGTGATGTGAAGGTGG
CAGCCACATAATTAAGACATAAGCGAGGTGCTTCAGATCTTGTGCTTGTGCTGCCAACCTGAACTGGAAGGAG
GCCTCTCCACCGTCCAACCTCTTGTGTGTTTTTCTCTGTGGAGAATGATGGGACATGCTGGACGCCCTTGGTGTCT
GTGAGTTTCTTTATTTGAATGAGGTTCTCCAGCGCTCTGGTCCGGAGTGTCTCTCTTGAATGTTTGTGCTATC
CTGGCTCAGCCATGGTCTGCGCAAGTTCTGATGATCAGGCTCTGACAGCAGCTCTTAACAGGAACAGATGGTCACTG
CCTTTTTCAGCCTCGGTGCACTGACCGGTTTTGTGCGAGATCTTGGGCCACTCAGGGGATCCGTTGCAGGCACTGATG
AATTTGAGCACCTATGCATTCCTGCTCGGATGATGGTTGTGACCTCACCAGTCCAGTGATCAGTGTGGTGTGTCGCA
CTACTTGGCATCATTTTGTACTTGTTTAAGTACCGTGGCTGCACCATATCTTGTGGCGATGGAGTGTCTCTGCGG
CTTTCTTCTTGAGATCTTTGCCAGGGAAGGTTGAGGGAAGGGGTGTCGAATCTCGGAATGAATCATGAGTCTCTG
ACTGGTGCCTCGTATGAGACTCAATGACGAGGACTTGATTTCTTGAATGGAGTGAATTTAAGTGTGTTGTTTCT
TGCGTCCAATGAGGAATGACGGGGTCAATTTATCGAGGCTCGCTATGCTAAGCACTTAGAGTAGAACTGGCCAGT
TGGTGCAAGTTGATAAAGTTGAGGTACTTTGGCCAACTTGAAGCTTTTGTGATACCGTGGCACCTCACTCTCGCC
GGTGACATGTTGTGCTCTCGGCCACGCGCTGTTGGCAGTATCTCGACCTAAGGTTGGTAGCACAAGCATACCCT
CCAAGCCATTGAGACCAGAGTCTTGTGGGTCCAAATGACCGTGGCGCGCTGTCGACCCGACCCACGCCCCAC
CCGACCCGTGCCATCCCCCTCCACCGAAAGTTCTGGAGAATGGCCCAACGCTTGGGGGATGAGGACCGTTTGAAT
AAGAAGAAGAGGCGCAGGATGGAAGCCCTCGGCATCTATGTTATGGGCGGGAAGAAATACAGAAATTTGGGACAAGAA
TTCCGGTGATGTGTTTTATGAGGAGGTCCATAATAACACAGATGAGTGGGAGTGTCTCAGAGTTGGCGACCTGCCGACT
TTGACCTGAGAAGGGAACCTCTGTGTGGACATGTACATTGAAAAAGGCTTACCATGTTTACACCTCCCATCTGGT
AAGAAGTTCTTGGTCCCGTCAACCCAGAGAATGGAAGAGTCCAATGGGAAGCTGCAAAGCTTCCGTGGAGCAGGCCCT
AGGTATGATGAATGTGACGGCGAACTGACTGCCAAAGAACTGGAGAACTGAAAAGAATATGACAACTCAGGGCC
TGACTAAGGAGCAGTGTTTAACTGCTAGCCGCCAGCAGTATGACCCGCTGTGGTTCGCGGGCTTGGTTGTACTGAAA
CAGCGGTAAAAATAGTCAAATTTACAACCGGACCTTACCTGGGACCTGTGAATTTAAAGATGGCCAGTGAGGTTGAG
CTAAAGACCGGGTTGAGCACAACCAACCGGTTGCGAGACCGATGATGGTGGAGTTGTGCTCCTGCGTTCGCGGT
TCCTTCGCTTATAGACGTCTGATCTCCGGTGTGATGATCTCCCAAGTTACTTGCCCATCAGGGCCGGGAAACACTG
GGATCGATGGCAGCTCTGGGATTTTGTGTCGAAGCCACTAAAGAGGAAGTGCAGTCACTCAGTGCAGAAATAATACAGGCT
TGTGACATTAGGCGCGGCGACGCTCTGAAATTTGGTCTCCCTTACAAGCTGACCTGTTAGGGGTAACCTGAGCGGGT

Fig. 1K-3

GAAAGGAGTTCTGCAGAATACAAGGTTTGGAGACATACTTACAAAACCCCCAGTGACACTGGAAGCCCAGTGACACGCGG
 CTGCCTGCCTTACGCCCAACGCCACTCCGGTGACTGATGGGCGCTCCGTCTTGGCCACGACCATGCCCCCGGTTTGGAG
 TTATATGTACCGACCATACCAGCGTCTGTCTTGATTACCTTGACTCTAGGCCTGACTGCCCTAAACAGCTGACAGAGCA
 CGGCTGCGAAGATGCCGACTGAAAGACCTCTCTAAATATGACTTGTCCACCAAGGCTTGTGTTTACCTGGAGTTCTTC
 GCCTTGTGCGGAAATACCTGTTTGGCCATGTAGGTAAGTGCCACCCGTTTCATCGGCCTTCTACTTACCCTGCTAAGAAAT
 TCTATGGCTGGAATAAATGGGAACAGGTTCCCAACCAAGGACATTAGAGCGTCCCTGAAATCGACGTTCTGTGCGCACA
 GGCTGTGCGAGAAAACCTGGCAACTGTCACCCCTTGACTCTTAAGAAACAGTATTGCGGGAAGAAGAAGACTAGGACCA
 TACTCGGCACCAATAACTTCATCGCACTAGCCACCGAGCAGTGTGAGTGGTGTACCCAGGGCTTCATGAAAAAGGCG
 TTTAACCTGCCCATCGCCCTCGGAAAGAACAAAGTTAAGGAGCTACAGACTCCGGTCTGGGCAGGTGCCTTGAAGCTGA
 TCTCGCATCCTGCGATCGATCCACGCTGCAATTGTCCGCTGGTTTGGCGCAACCTTCTTATGAACTTGCCTGTGCTG
 AAGAGCATCTACCGTCGTACGTGCTGAACTGCTGCCACGACTTACTGGTCACGCACTCCGGCGCAGTGACTAAGAGAGGT
 GGCCTGTGCTGTGCGGACCCGATCACCTCTGTGCTAACACCATTTATAGTTTGGTGATCTATGCACAGCATATGGTGCT
 TAGTTACTTCAAAGTGGTCACCCCATGGCCTTCTGTTCTTACAAGACCAGCTAAAGTTTGAGGACATGCTCAAGGTTT
 AACCCCTGATCGTCTATTCGGACGACCTCGTGCTGTATGCCGAGTCTCCACCATGCCAACTATCACTGGTGGGTTGAA
 CATCTGAATTTGATGCTGGGGTTTACAGCGGACCCAAAGAAGACAGCAATAACAGACTCGCCATCATTTCTAGGCTGTAG
 AATAATAAATGGGCGCGAGCTAGTCCCAACCGTGACAGGATCCTCGCGGCCCTCGCCTATCACATGAAGGCGAGTAATG
 TTTCTGAATACTATGCCTCAGCGGCTGCAATACTCATGGACAGCTGTGCTTGTGTTGGAGTATGATCCTGAATGGTTTGA
 GAACCTGTAGTTGGAATAGCGCAGTGCGCCCGCAAGGACGGCTACAGCTTTCGGGCACGCCGTTCTTCATGTCATGTG
 GGAAAACTCAGGTCCAATTATGAGGGGAAGAAGTCGAGAGTGTGCGGGTACTGCGGGGCCCGGCCGCTACGCTACTG
 CCTGTGGCCTCGACGTCTGCATTTACACACCCACTTCCACAGCATTGTCCAGTCACAATCTGGTGTGGCCATCCAGCG
 GGTTCTGGTTCTTGTAGTGAGTGCAAAATCCCTGTAGGGAAAGGCACAAGCCCTTAGACGAGGTGCTGGAACAAGTCCC
 GTATAAGCCCCACGGACCGTTATCATGCATGTGGAGCAGGGTCTCACCCCTTGTATCCAGGTAGATAACAACTCGCC
 GCGGATTAGTCTCTGTGAGGCGTGAATTAGGGGAAATGAAGTTGGACTACCAGACGGTGATTATGCTAGCACCGCCTTG
 CTCCCTACCTGCAAAGAGATCAACATGGTGCCTGTGCTTCCAATGTATTGCGCAGCAGGTTTCATCATCGGCCACCCGG
 TGCTGGGAAACATACTGGCTCCTTCAACAGGTCCAGGATGGTGATGTTATTACACACCACTCACAGACCATGCTTG
 ACATGATTAGGGCTTTGGGGACGTGCCGGTTCAACGTCCCGGCAGGCACAACGCTGCAATCCCCGTCCCCTCCCGCACC
 GGTCCGTGGGTTGCGATCCTAGCCGGCGGTTGGTGTCTGGCAAGAATCCTTCTAGATGAAGCAGCGTATTGCAATCA
 CCTTGATGTTTTGAGGCTTCTTAGTAAACTACCTCACCTGTCTAGGAGACTTCAAGCACTCCACCCAGTGGGTTTTG
 ATTCTCATTGCTATGTTTTGACATCATGCCTCAAACCTCACTGAAGACCATCTGGAGGTTTGGACAGAATATCTGTAT
 GCCATTACGCCAGATTACAGGGACAACTCATGTCCATGGTCAACACAACCCGTGTGACCTACGTGGAAAAACCTGTGAG
 GTATGGGCAGGTCTCACCCCTACACAGGGACCGAGAGGACGACGCCATCACTATTGACTCCAGTCAAGGCGCCACAT
 TCGATGTGGTTACATTGCATTTGCCACTAAAGATTCACTCAACAGGCAAGAGCCCTTGTGCTATCACAGGGCAAGA
 CACGCTATCTTTGTGTATGACCACACAGGCAGCTGACGGCTTGTGATCTTCTGCAAAAGGCACGCCGTCACCT
 CGCAGTGCATGCGACGGCGAGCTGATCGTGTGGATAGAAATAACAAAGATGCACGGTGTCTCAGGCTTAGGCAACG
 GGGATAAATTTAGGGCCACAGACAAGCGTGTGTAGATTCTCCGCGCCATTGTGCTGATCTAGAAGGGTCGAGCTCT
 CCGCTCCCAAGGTGCGACACAACCTGGGATTTTATTTCTCACCTGATTTAACACAGTTTGCTAACTCCAGTAGAACT
 TGACCTCACTGGCCCGTGGTGTCAACCCAGAACAAATGAAAAGTGGCCGGATCGGCTGGTTGCCAGCCTTCGCCCTATCC
 ATAAATACAGCCGCGCTGCATCGGTGCCGCTATATGGTGGGCCCTTCGGTGTCTTAGGCACTCCTGGGGTCTGTGCA
 TACTATCTCAAAAATTTGTTAAGGGCGGGCTCAAGTGCTTCCGGAGACGGTTTTGAGCACCAGGCGCAATTGAGGTAGA
 CTGCCGGGAATATCTTGATGATCGGGAGCGAGAAGTTGCTGCGTCCCTCCACACGCTTTTATTGGCGACGTCAAGGCA
 CTACCGTTGGAGGATGTATCATGTCACTCCAGATACCTCCCGCGCTCCTCCCAAGGAATCAGTTGCGGTAGTCGGG
 GTTTCAAGCCCCGGAAAAGCCGCGAAAGCATTGTGCACACTGACAGATGTGTACCTCCAGATCTTGAAGCCTATCTCCA
 CCCGGAGACCCAGTCCAAGTGCTGGAAAATGATGTTGGACTTCAAGAAAGTTGACTAATGGTCTGGAAAGACAAAACAG
 CCTATTTCAAACCTGAAGGTCGCTATTTCACTGGTATCAGCTTCCAGCTATGCCTCGTACATCCGTGTTCCCGTCAAC
 TCTACGGTGTACTTGGACCCCTGCATGGGCCCGCCCTTTGCAACAGGAGAGTCGTGCGGTCCACCACTGGGGGCTGA
 CCTCGGGTCACCCCTTATGATTACGGCGCTAAAAATATCTGTCTAGCGCTACCATGGTGAAATGCCCCCGGATACA
 AAATCTGGCGTGCAGGAGTTCTCGTTGGATGACCCAGTTAAGTACAAACATACCTGGGGGTTTGAATCGGATACAGCG
 TATCTGTATGAGTTCACCGGAAACGGTGAGGACTGGGAGGATTACAATGATGCGTTTCTGTGCGGCCAGGAAGGGAAAAAT

Fig. 1K-4

TTATAAGGCCACTGCCACCAGCTTGAAGTTTTATTTTCCCCGGGCCCTGTCATTGAACCAACTTTAGGCCTGAATTGAA
 ATGAAATGGGGTCCATGCAAAGCCTTTTGGCAAAATGGCCAACTTTTGTGGATGCTTTCACGGAGTCTTGGTGTCC
 ATTGTTGATATCATTATATTTTGGCCATTTTGTGGCTTACCATCGCCGGTGGCTGGTGGTCTTTGCATCAGATT
 GGTGCTCCGCGATACTCCGTACGCGCCCTGCCATTCACTCTGAGCAATTACAGAAGATCTTATGAGGCCTTTCTTTCC
 CAGTGCCAAGTGGACATTCCACCTGGGGAATAACATCCTTTGGGGATGCTTTGGCACCATAAGGTGTCAACCCTGAT
 TGATGAAATGGTGTGCGTCGAATGTACCGCATCATGGAAAAAGCAGGGCAGGCTGCCTGGAAACAGGTGGTGAAGCGAGG
 CTACGCTGTCTCGCATTAGTAGTTGGATGTGGTGGCTATTTTTCAGCATCTAGCCGCCATTGAAGCCGAGACCTGTAAA
 TATTTGGCCTCCCGGTGCCATGCTACACAACCTGCGCATGACAGGGTCAAATGTAACCATAGTGTATAATAGCACTTT
 GAATCAGGTGTTTGTATTTTTCCAACCCCTGGTTCGGGCCAAAGCTTCATGATTTTCAGCAATGGTTAATAGCTGTAC
 ATTCCTCCATATTTTCTCTGTGTCAGCTTCTTGTACTCTTTTGTGTGTGTGGTTGCGGGTCCAATACTACGTACT
 GTTTTGGTTTCCGCTGGTTAGGGGCAATTTTCTTTCGAACTCACAGTGAATTACAGGTGTGTCCACCTTGCCCTCACC
 CGGCAAGCAGCCACAGAGATCTACGAACCCGGTAGGTCTCTTGGTGCAGGATAGGGTATGACCGATGTGGGGAGGACGA
 TCATGACGAGCTAGGGTTTATGATACCGCTGGCTCTCCAGCGAAGGCCACTTGACTGGTGTACGCTGGTGGCGT
 TCTTGTCTTCAGCTACCGCCAGTTCATCCCGAGATATTCCGGATAGGGAATGTGAGTCGAGTTTATGTTGACATC
 AAACATCAACTCATCTGCGCCGAACATGACGGGCAGAACACCACCTTGCCTCGTCATGACAACATTTAGCCGTGTTTCA
 GACCTATTACCAACATCAAGTCGACGGCGCAATTGGTTTTCACCTAGAATGGCTTCGCTCCTTCTTTCTCGTGGTTGG
 TTTTAAATGTCTTGGTTTCTCAGGCGTTCGCTGCAACCATGTTTCAGTTCGAGTCTTGAGATATTAAGACCAACA
 CCACCGCAGCGGCAAGCTTGTCTCTCAAGACATCAGTTGCCTTAGGCATCGCGACTCGGCCTCTGAGGCGATTGCG
 AAAATCCCTCAGTGCCGTACGGCGATAGGGACACCGGTGTATGTTACCATCACAGCCAATGTGACAGATGAGAATTTT
 ACATTCTCTGATCTCTCATGCTTCTTCTTGCCTTTCTATGCTTCTGAGATGAGTGAAAAGGGATTTAAGGTGGTAT
 TTGGCAATGTGTGAGGCATCGTGGCTGTGTGTCAATTTACCAGTACGTCCAACATGTCAAGGAGTTTACCAACGC
 TCCTCGTGGTGTGACCATGTGCGGTTGCTCCATTTATGACACCTGAGACCATGAGGTGGGCAACTGTTTAGCCTGTCT
 TTTTGCATTCTGTTGGCAATTTGAATGTTAAGTATGTTGGAGAAATGCTTGACCGGGGCTGTGCTCGGATTGCTT
 TCTTGTGGTGTATCGTGCGGTTCTGTTTGTGTGCTCGCAACGCCAGCAACGACAGCAGCTCCCATCTACAGCTGAT
 TTACAACCTGACGCTATGTGAGCTGAATGGCAGATTTGGCTAGCTAACAAATTTGATTGGGCGATGGAGAGTTTGTCA
 TCTTCCCGTTTGTACTCACATTGTCTCCTATGGTGCCTCACTACCAGCCATTTCTTGACACAGTCGCTTTAGTCACT
 GTGTCTACCGCGGGTTTGTTCACGGCGGTATGTCTAAGTAGCATCTACGCGGTCTGTGCCCTGGCTGCGTTGACTTG
 CTTCTGCTATTAGGTTTGAAGAATTGCATGTCTGGCGCTACGCGTGTACCAGATATACCACTTTCTTCTGGACACTA
 AGGGCAGACTCTATCGTTGGCGGTGCGCTGTATCATAGAGAAAAGGGGCAAGTTGAGGTGGAAGGTCTCTGATCGAC
 CTCAAAAGAGTTGTGCTTGATGGCTCCGTGGCAACCCCTATAACCAGAGTTTCAGCGGAACAATGGGGTCTGCTTATAGT
 GACTTCTGTACGATAGCAGGCTCCACAAAAGGTGCTTTTGGCGTTTCTATTACCTACAGCCAGTGATGATATATGC
 CCTAAAGGTGAGTCGCGGCCGACTGCTAGGGCTTCTGACCTTTTGTCTTCTGAAATGTGCTTTCACCTTCGGGTACA
 TGACTTTCGCGCACTTTCAGAGTACAAATAAGGTGCGGCTCACTATGGGAGCAGTAGTTGCACTCCTTTGGGGGTGTAC
 TCAGCCATAGAAACCTGGAAATTATCACCTCCAGATGCGGTTTGTGCTTGTAGGCCGCAAGTACATTCTGGCCCTGCT
 CCACCACGTTGAAAGTGCCGCAAGGCTTTCATCCGATTGCGGCAATGATAACCACGCATTGTGCTCGGCGTCCCGGCT
 CCACTACGTTCAACGGCAATTGGTGCCCGGTTAAAAAGCCTCGTGTGGGTGGCAGAAAAGCTGTTAAACAGGGAGTG
 GTAAACCTTGTCAAATATGCCAAATAACAACGGCAAGCAGAGAAGAGAAAGAGGGGGATGGCCAGCCAGTCAATCAGC
 TGTGCCAGATGCTGGGTAAGATCATGCTCAGCAAAACAGTCCAGAGGCAAGGGACCGGAAAGAAAAATAAGAAGAAA
 AACCCGGAGAAGCCCATTTTCTCTAGCGACTGAAGATGATGTGACACATCACTTACCCCTAGTGAGCGGCAATTGTG
 TCTGTCGTCAATCCAGACCGCTTTAATCAAGGCGCTGGGACTTGACCCCTGTGAGATTAGGGAGGATAAGTTACTGT
 TGGAGTTAGTTTGCCTACGCATCACTGTGCGCTGATCCGCGTCACAGCATCACCTCAGCATGATGGGCTGGCATT
 CTTGAGGCATCTCAGTGTGTTGAATTGGAAGAATGTGTGGTGAATGGCACTGATTGACATTGTGCTCTAAGTCACCTATT
 CAATTAGGGCGACCGTGTGGGGTGAGATTTAATTGGCGAGAACCATGCGGCCGAAATTAACAAAAA

Fig. 1L-1

>V7-Nsp2d727-813.seq

ATGACGTATAGGTGTTGGCTCTATGCCTTGGCATTGTATTGTGACGAGCTGTGACCATTTGGCACAGCCCAAACTTGCT
 GCACAGAAACACCCCTTCTGTGATAGCCTCCTTCAGGGGAGCTTAGGGTTTGTCCCTAGCACCTTGCTTCCGGAGTTGCAC
 TGCTTTACGGTCTCTCCACCCCTTAACCATGTCTGGGATACTTGATCGGTGCACGTGTACCCCAATGCCAGGGTGT
 ATGGCGGAGGGCCAAGTCTACTGCACACGATGCCTCAGTGCACGGTCTCTCCTTCCCCTGAACCTCCAGGTTCTGAGCT
 CGGGGTGCTAGGCCTATTCTACAGGCCCGAAGAGCCACTCCGGTGGACGTGGCCACGTGCATTCCCCACTGTTGAGTGCT
 CCCCCGCCGGGGCTGCTGGCTTTCTGCAATCTTCCAATCGCACGAATGACCAGTGGAAACCTGAACCTTCCAAACAAGA
 ATGGTACGGGTCGCAGCTGAGCTTTACAGAGCCGGCCAGCTCACCCTGCAGTCTTGAAGGCTCTACAAGTTTATGAACG
 GGGTTGCCGTGGTACCCCATTTGTTGACCTGTCCCTGGAGTGGCCGTTTTTCGCCAATTCCTACATGTGAGTGATAAAC
 CCTTCCCGGAGCAACTCACGTGTTGACCAACCTGCCGCTCCCGCAGAGACCCAAGCCTGAAGACTTTTGCCCTTTGAG
 TGTGCTATGGCTACTGTCTATGACATTGGTCATGACGCCGTGATGTGTTGGCCGAAAGGAAAGTCTCCTGGGCCCTCG
 TGGCGGGGATGAAGTGAAATTTGAAGCTGTCCCGGGGAGTTGAAGTTGATTGCGAACCGGCTCCGCACCTCCTTCCCGC
 CCCACACACAGTGGACATGTCTAAGTTCGCCCTTACAGCCCTGGGTGGTGGTGTCTATGCGGGTCGAACGCCAACAC
 GGTGCTTCCCGCTGACACTGTCCCTGAAGGCAACTGCTGGTGGAGCTTGTGTTGACTTGCTTCACTGGAAGTTCAAGAA
 CAAAGAAATTCGCCATGCTAACCAATTTGGCTACCAGACCAAGCATGGTGTCTTGGCAAGTACCTGCAGCGGAGGCTGC
 AAGTTAATGGTCTCCGAGCAGTAAGTACCTAACCGACCTATCGTCGTACAGTACTTCTCCGTTAAGGAGAGTTGGATC
 CGCCATTTGAAACTGGCGGGAGAACCCAGCTACTCTGGGTTTGGAGCTCCTCAGAATAAGGGTTGAGCCTAACACGTC
 GGCATTGGCTGACAAGGAAAGAAAAATTTTCCGGTTTGGCAGTCACAAGTGGTACGGCGCTGGAAAGAGAGCAAGAAAAAG
 CACGCTCTTGTGCGACTGCTACAGTCGCTGGCCGCGCTTGTCCGTTCTGAAACCCGGCAGGCCAAGGAGCACAGGTT
 GCCCGCGCAACAAGGCTGAGCACCTCAAACTACTCCCCGCTGCCGAAGGGAATTTGGTGGCACTGCATTTCCGC
 CATCGCAACCGGATGGTGAATCCAAATTTGAAACACCCCTTCCCGAAAGAGTGAACCTCCAGATGACTGGGCTACTG
 ACGAGGATCTTGTGAATGCCATCCAAATCCTCAGACTCCTGCGGCTTAGACAGGAACGGTGCTTGTACTAGCGCCAAG
 TACGTAATTAAGCTGGAAGGTGAGCATTGGACTGTCACTGTGACCCCTGGGATGTCCCTTCTTGTCTTCTTGAATG
 TGTTACGGGCTGTTGTGGGCACAAGGGCGGTCTTGGTCCCCAGATGCAGTCGAGGTCTCCGGATTGACCTGCCTGCC
 TTGACCGGCTGGCTGAGGTGATGCACCTGCCTAGCAGTGCATCCAGCCGCTCTGCGCAATGTCTGGCGATTCCGAT
 CGTTCCGCTTCTCCGGTCAACCCGTGTGGACTGTTTCGAGTCTTTCGCGTCTTTCGCCGTACAGCGGAGGGAATCACCTGACCA
 AGTGGCTTAGGGAAAATTATCAGCCTTTGTGAGGTGATTGAGGACTGCTGCTGTTCCAGAACAAAACCAACCGGGTCA
 CCCCCGAGGAGGTGCGAGCAAGATTGACCTGTACTCCGTGGTGCACAAATCTTGAAGAATGCTTGGCCAGGCTTGA
 AAAGCGCCGCCCGCACGCGTAATCGACACCTCCTTGATTGGGATGTTGTGCTCCCTGGGTTGAGGCGGCAACCCAGAC
 GATCAAGCTGCCCCAGGTCAACAGTGTCTGCTGCTGGTCCCTGTTGTGACTCAAAAGTCTTGGACAACACTCGGTCC
 CCTGACCGCTTTTCACTGGCTAATACTACTACCGTGCAGGAGTGAAGATTGTCACCGTGAAAGACTAACCGCC
 GTGCTCTCAAGTTGGAAGGTTGTTGAGAAGAATATGGCTCATGCCAACCGAGCTGGTCCACGGCCCACTGCC
 ACGCGGGCTCGACGAACCAAGACAGATGGAGGAGGACTTGTGAAACTGGCTAACGCCAGACGACTTGGACATGA
 TGGCTGGGAGTCGAGCAGGTTGACCTAAAACTTGGGTCAAGAACTACCGCGGTGGACACCAACCCCTCCGCCA
 AAAGTTGAGCTCGAAAAACGAAGCCTGTCAAGAGCTTGGCGAGAGAAAGCCTGTCCCGCCCCGCGCAGGAAGGTTGG
 GTCCGATTGTGGCAGCCCGGTTTCTTAGGGCGGATGTCCCTAACAGTTGGGAAGATTGGCTGTTAGTACCCCTTG
 ATCTCCCGACCCACCTGAGCCGGCAACACCTTCAAGTGAAGTGGTATTGTGCTCCTACCGCAATGCATCTTACGGCCG
 GCGACACCTTGAGTGAGCCGGCTCCAATTCGCGACCTCGCGGAACTGTGTCTGACCGGTGACACCTTGAGTGAGCC
 GATCCCTGTGCCCGCACCGCGGCTAAGTTTCAGCAGGTGAAAAGATTGAGTTCGGCGGGGCAATCCACCGTACCAGG
 ACGAGCCCTGGATTTGTCTGCTTCTCACAGACTGAATATGAGGCTCTCCCCAGCACCGCGCAGAGCGGGGCGTT
 CTGGGAGTAGAGGGGATGAAGCTGAGGAAACCTGAGTGAAATCTCGGACATGTGGGTAACATTAACCTGCTCCGT
 GTCATCAAGCAGCTCCTTGTCCAGCTGAGAATCACACGCCCAAAATACTCAGCTCAAGCCATCATCGACTCGGGCGGGC
 CCTGCAGTGGGCATCTCAAGAGGTAAGGAAACATGCCTTAGTGTGATGCGCGAGGCATGTGATGCGACTAAGCTTGAT
 GACCCTGTACGAGGAATGGCTTTCTGCATGTGGGATCGGGTGGACATGCTGACTTGGCGCAACAGCTGTGTTTACCA
 GCGGATTTGCACCTTAGATGGCAGGTTAAAGTTCTCCCAAAATGATACTCGAGACACCGCGCCCTATCCGTCTTTTA
 CCGATTTGCCGCTTTCAGATGGCGCGGATGCGGACGGGGGGGGCGTTTCGGACGGTAAAAAGAAAGCTGAAAGGCTC
 TTTGACCAACTGAGCCGTGAGGTTTTGACCTGTCTCCATCTCCCTGTTTTCTTCTCACGCTTTTCTACCTTGGCGG
 TGGTTATTCTCCGGGTGATTGGGGTTTTGAGCTTTTACTCTATTGTGCTCTTTTTATGTTACAGTTACCAAGCCTTG
 GTATTGCTCCCTCTTGGGTGTGTTTTCTGGGTCTTCTCGGCGGTTTGAATGGGGTTTTTGGCTGCTGGTGGCTTTT

Fig. 1L-2

GCTGTTGGTCTGTTCAAGCCTGTGTCCGACCCAGTCGGCGCTGCTTGTGAGTTTGACTCGCCAGAGTGTAGAAACATCCT
 TCATTCTTTTGTGCTTCTCAACCTTGGGACCTGTTGCGAGCCTTGTGTGGGCCCGTCGGTCTCGGTCTTGCCATTC
 TTGGCAGGTTACTGGGCGGGGACGCTGCATCTGGCACTTTTGTCTAGGCTTGGCATTGTTGCAGACTGTATCTTGGCT
 GGAGCTTACGTGCTTTCTCAAGGTAGGTGTAAAAAGTGTGGGGATCTGTATAAGAACTGCTCCCAATGAGGTCGCTTT
 TAACGTGTTTCTTTTACACGTGCGACCCAGGTCGTCACTTATCGACCTGTGCGATCGGTTTTGTGCGCCAAAAGGAATGG
 ACCCATTTTTCTCGCCACTGGGTGGCGCGGTGCTGGGCGGCCGAAGCCCCATTGAGCAACCTCTGAAAAACCCATC
 GCGTTTGGCCAGTTGGATGAAAAGAAGATTACGGCTAGGACTGTGGTCGCCAGCCTTATGACCCCAACCAAGCCGTAAA
 GTGCTTGGCGGTATTGACAGCGGGTGGGGCGATGGTGCTAAGGCGGTCCCAAAAGTGGTCAAGGTTTCCGCTGTTCCAT
 TCCGAGCCCCCTTCTTCCCACTGGAGTGAAAGTTGACCCTGATTGCGAGGTCGTGGTTGACCCTGACACTTTCACTGCA
 GCTCTCCGGTCTGGCTACTCCACCACAAACCTCGTCTTGGTGTGGGGACTTTGCCAGCTGAATGGATTAAAAATCAG
 GCAAAATTTCAAGCCTTACAGGGGAGGCCACATCTCATGGCTGCCCTGCATGTTGCTGCTCGATGGCTCTGCACATGC
 TTGCTGGGATTATGTGACTGCGGTGGGTTCTTGGCGCACCGGCACCAACGACCCGTGGTGGCTAACCCTTTGCCGTC
 CCTGGCTACGGACCTGGCTCTCTGTCACGTCCAGATTGTGCATTTCCCAACACGGCCTTACCCTGCCCTTGACAGCACT
 TGTGGCGGGATTGCGTATTCAAGAAATTGCCTTGGTGGTTTGTATTTTGTTCATCGGAGGCATGGCTCATAGGTTGA
 GCTGTAAGGCTGACATGCTGTGTCTTGTGCAATTGCCAGCTATGTTGGGTACCTTTACCTGGTGTCTTGTGTG
 TTTCTTGTGTTGGCTGTTTTCTTTGCACCCCCTCACCATCTATGGTGGTGTCTTGTGATTCTGTGAATAT
 GCCTTACGAATCTTGGCATGGTGTGTGGTTTCTTTGGCTTCTGGTGGTATACTAATGTTGCTGGCCTTGTCA
 CCCCCTACGACATTCATATTACACAGTGGCCCCGCGGTGTTGCCGCTTGGCTACCGCACCAGATGGGACCTACTTG
 GCGCTGTCCGCCGCGTGGTGGTGGCTGACCTGGCGCACCATGCTGTTTACCCCGTCCAGCTTGGGTCTTCTTGTAGGGTGC
 TTTCAGAACTCGAAAGCCCTCACTGAACACCGTCAATGTGATCGGGTCTCCATGGGCTTGGCGGGGTGTTTACCATCG
 ACGGGAAGTCAAGTGCCTAAGTGGCGACATGTCTTACGGGCAATTGAGCTCGGGTTTCCGGGGTGGCTTCAATCAA
 ATGCTTGACTTTGACGTAAGGGAGATTTCGTATAGCTGATTGCCCGAATTGGCAAGGGGCTGCCCCAAGACCCAATT
 CTGCACGGATGGATGGACTGGCCGTGCCTATTGGCTAACATCTCTGGCTCGAACCCTGGCTCATTGGAAGGATTTCG
 CTTCTGCTTACCGCATGTGGCGATTCCGGGTCCCCAGTGATCACCGAGGCGGTGAGCTTGTGGCGTTACACGGGA
 TCGAATAAACAAGGGGGGGCATTGTACGCGCCCTCAGGCCAGTTTGTAAATGTGGCACCCATCAAGCTAAGCGAATT
 AAGTGAATTTCTTGTGGGCTAAGGTCCCGCTCGGTGATGTGAAGGTGGCGAGCCACATAATTAAAGACATAAGCGAGG
 TGCCCTCAGATCTTGTGCTTGTGCTGCTGCAAACTGAACTGGAAGGAGGCTCTCCACCGTCCAACCTCTTGTGTG
 TTTTTCTCCTGTGGAGAATGATGGGACATGCCTGGACGCCCTGGTGTGCTGAGTTTCTTATTTGAATGAGGTTCT
 CCCAGCGTCTGGTCCGAGTGTTTTCTCTTTGGAATGTTTGTGCTATCTGGCTCACGCCATGGTCTGCGCAAGTTC
 TGATGATCAGGCTCTGACAGCAGCTCTAACAGGAACAGATGGTCACTTGCCTTTTTCAGCCTCGGTGAGTGACCGGT
 TTTGTGCGAGATCTTGGGCCACTCAGGGGCATCCGTTGACGGCAGTGATGAATTTGAGCACCTATGCATTCCTGCCTCG
 GATGATGGTGTGACCTCACCAGTCCAGTGATCACGTGGTGTGCTGACCTACTTGCCATCATTTTGTACTTGTGTA
 AGTACCGTGGCCTGCACCATATCTTGTGGCGATGGAGTGTCTCTGCGGCTTCTTCTTGAGATACTTGGCGAGGGA
 AAGTTGAGGGAAGGGGTGTCGAATCTGCGGAATGAATCATGAGTCTGACTGGTGCCTCGCTATGAGACTCAATGA
 CGAGGACTTGGATTTCTTATGAAATGGACTGATTTAAGTGCTTTGTTTCTGCGTCCAACATGAGGAATGCAGCGGGTC
 AATTTATCGAGGCTGCCTATGCTAAAGCACTTAGAGTAGAACTGGCCAGTTGGTGCAGGTTGATAAAGTTGAGGTA
 TTGGCCAAACTTGAAGCTTTTGTGATACCGTGGCACCTCAACTCTCGCCCGTGACATTGTTGTGCTCTCGGCCACAC
 GCCTGTTGGCAGTATCTTGCACCTAAAGGTTGGTAGCACAAGCATACCTCCAAGCCATTGAGACCAGAGTCCTTGCTG
 GGTCCAAATGACCGTGGCGCGCTGCTGACCCGACCCCCACGCCACCCGACCCGTGCCATCCCCCTCCACCG
 AAAGTTCTGGAGAATGGCCCCAACGCTTGGGGGATGAGGACCGTTTGAATAAGAAGAAGAGGCGCAGGATGGAAGCCCT
 CGGCATCTATGTTATGGGCGGGAATAACAGAAATTTGGGACAAGAATTCCGGTGATGTGTTTATGAGGAGGTCC
 ATAATAACACAGATGAGTGGGAGTGTCTCAGAGTTGGCGACCTGCCGACTTGGACCTGAGAAGGGAATCTGTGTGGA
 CATGTCACCATTTGAAAAAAGGCTTACCATGTTTACACCTCCCCATCTGGTAAGAAGTTCTTGGTCCCCGTCAACCCAGA
 GAATGGAAGAGTCCAATGGGAAGCTGCAAAGCTTCCGTGGAGCAGGCCCTAGGTATGATGAATGTCGACGGCGAATGA
 CTGCCAAAGAACTGGAGAACTGAAAAGAATAATTGACAACTCCAGGGCCTGACTAAGGAGCAGTGTAAACTGCTAG
 CCGCCAGCGACTTGACCCGCTGTGGTGGCGCGCTTGGTTGTTACTGAAACAGCGGTAAAAATAGTCAAATTTACAAC
 CGGACCTTCAACCTGGGACCTGTGAATTTAAAGTGGCCAGTGAGGTTGAGCTAAAGACGCGGTGAGCACAAACCA
 CCGGTTGCGAGACCGATCGATGGTGGAGTTGTGCTCCTGCTTCCGCGGTTCTTCTGCTTATAGACGTCTTGATCTCCG
 GTGCTGATGCATCTCCAAGTTACTTGCCCATCACGGGCGGGAAACACTGGGATCGATGGCACGCTCTGGGATTTTGA

Fig. 1L-3

TCCGAAGCCACTAAAGAGGAAGTCGCACTCAGTCGCGCAAATAATACAGGCTTGTGACATTAGGCGCGGCGACGCTCCTGA
 AATTGGTCTCCCTTACAAGCTGTACCCTGTTAGGGGTAACCTGAGCGGGTGAAAGGAGTTCTGCAGAATACAAGTTTG
 GAGACATACCTTACAAAACCCCACTGACACTGGAAGCCCACTGACGCGGCTGCCTGCCTTACGCCCAACGCCACTCCG
 GTGACTGATGGGCGCTCCGTCTTGGCCACGACCATGCCCCCGGGTTTGAAGTTATATGTACCGACCATACCAGCGTCTGT
 CCTTGATTACCTTGACTCTAGGCTGACTGCCCTAAACAGCTGACAGAGCACGGCTGCGAAGATGCCGCACTGAAAGACC
 TCTCTAAATATGACTTGTCCACCCAAGGCTTTGTTTTACCTGGAGTTCTTCGCTTGTGCGGAAATACCTGTTTGGCCAT
 GTAGGTAAGTGCCACCCGTTTATCGGCTTCTACTTACCCTGCTAAGAATTCTATGGCTGGAATAAATGGGAACAGGTT
 CCCAACCAAGGACATTAGAGCGTCCCTGAAATCGAGCTTCTGTGCGCACAGGCTGTGCGAGAAAACTGGCAAACTGTCA
 CCCCTTGACTCTTAAGAAACAGTATTGCGGAAGAAGAAGACTAGGACCATACTCGGCACCAATAACTTCATCGCACTA
 GCGCCACGAGCAGTGTGAGTGGTGTACCCAGGGCTTCATGAAAAAGGCGTTAACTCGCCCATCGCCCTCGGAAAGAA
 CAAGTTTAAGGAGCTACAGACTCCGGTCTGGGCAAGTGCCTGAAGCTGATCTCGCATCTGCGATCGATCCACGCTG
 CAATTGTCCGCTGGTTTCCGCGCAACCTTCTTTATGAACCTGCTGTGCTGAAGAGCATCTACCGTCGTACGTGCTGAAC
 TGCTGCCACGACTTACTGGTCACGCACTCCGGCGCAGTGACTAAGAGAGGTGGCTGTGCTGTGCGACCCGATCACTC
 TGTGTCTAACACCATTTATAGTTTGGTGATCTATGCACAGCATATGGTGCTTAGTTACTTCAAAGTGGTCACCCCATG
 GCCTTCTGTTCTTACAAGACCAGCTAAAGTTTGAAGACATGCTCAAGGTTCAACCCCTGATCGTCTATTCGGACGACCTC
 GTGCTGTATGCCGAGTCTCCACCATGCCAACTATCACTGGTGGGTTGAACATCTGAATTTGATGCTGGGGTTTCAGAC
 GGACCCAAAGAAGACAGCAATAACAGACTCGCCATCATTTCTAGGCTGTAGAATAATAAATGGGCGCCAGTAGTCCCA
 ACCGTGACAGGATCCTCGCGGCCCTCGCCTATCAGATGAAGGCGAGTAATGTTTCTGAATACTATGCTCAGCGGCTGCA
 ATACTCATGGACAGCTGTGCTTGTGGAGTATGATCCTGAATGGTTGAAGAACTTGAAGTTGAAGTACGCGAGTGGCG
 CCGCAAGGACGGCTACAGCTTCCCGGCACGCCGTTCTTCATGTCCATGTGGGAAAACTCAGGTTCAATTATGAGGGGA
 AGAGTCTGAGAGTGTGCGGGTACTGCGGGGCCCGGCCCGTACGCTACTGCCTGTGGCTCGAGCTGTGATTTACCAC
 ACCCACTTCCACAGCATTTGTCAGTCACAATCTGGTGTGGCCATCCAGCGGTTCTGGTTCTGTAGTGAGTGCAATC
 CCCTGTAGGGAAGGCACAAGCCCTTTAGACGAGGTGCTGGAACAAGTCCCGTATAAGCCCCACGGACCGTTATCATGC
 ATGTGGAGCAGGGTCTCACCCCTTGATCCAGGTAGATACCAACTCGCGCGGATTAGTCTGTGTCAGGCGTGAAT
 AGGGGAAATGAAGTTGGACTACCAGACGGTGATTATGCTAGCACCGCTTGTCTCTACCTGCAAGAGATCAACATGGT
 CGCTGTGCTTCAATGTATTGCGCAGCAGGTCATCATCGCCACCCGCTGCTGGGAAAACATACTGGCTCCTTCAAC
 AGGTCCAGGATGGTGATGTTATTTACACACCAACTCACCAGACCATGCTTGACATGATTAGGGCTTGGGGACGTGCCGG
 TTCAACGTCCCGGCAGGCACAACGCTGCAATTCCTCGTCCCTCCCGCACCGGTCCGTGGGTTGCGATCCTAGCCGGCGG
 TTGGTGTCTGGCAAGAATTCCTTCTAGATGAAGCAGCGTATTGCAATCACCTTGATGTTTTGAGGCTTCTTAGTAAAA
 CTACCTCACCTGTCTAGGAGACTTCAAGCAACTCCACCCAGTGCGGTTTTGATTCTCATTGCTATGTTTTGACATCATG
 CCTCAAACCTCACTGAAGACCATCTGGAGGTTTGGACAGAATATCTGTGATGCCATTAGCCAGATTACAGGGACAACT
 CATGTCCATGGTCAACACAACCCGTGTGACCTACGTGGAAAACTGTGAGGTATGGGAGGTCTCACCCCTACCACA
 GGGACCGAGAGGACGACGCCATCACTATTGACTCCAGTCAAGGCGCCACATTCGATGTGGTTACATTGCATTTGCCACT
 AAAGATTCACTCAACAGGCAAGAGCCCTTGTGCTATCACCAGGGCAAGACACGCTATCTTTGTATGACCCACACAG
 GCAGCTGCAGGGCTTGTGATCTTCTGCAAAAGGCACGCCCTCAACCTCGCAGTGCACTGCGACGGGAGCTGATCG
 TGCTGGATAGAAATAACAAAGAATGCACGGTTGCTCAGGCTTAGGCAACGGGGATAAATTTAGGGCCACAGACAAGCGT
 GTTGTAGATTCTCCGCGCCATTTGTGCTGATCTAGAAGGGTCGAGCTCTCCGCTCCCAAGGTGCGACACAACCTGGG
 ATTTTATTTCTCACCTGATTTAACACAGTTTGCTAACTCCAGTAGAACTGCACCTCACTGGCCCGTGGTGTCAACCC
 AGAACAATGAAAAGTGGCCGGATCGGCTGGTGGCAGCTTCCGCTATCCATAAATACAGCCGCGCTGCATCGGTGCC
 GGCTATATGGTGGGCCCTTCGGTGTCTAGGCACTCTGGGGTGTGTCATACTATCTCACAAAATTTGTTAAGGGCGG
 GGCTCAAGTGCTTCGGAGACGGTTTTCAGCACCGGCCGAATTGAGGTAGACTGCCGGGAATATCTTGATGATCGGGAGC
 GAGAAGTTGCTGCGTCCCTCCACACGCTTTCATTGGCGACGTCAAAGGCACTACCGTTGGAGGATGTCATCATGTACC
 TCCAGATACCTCCCGCGCTCCTTCCCAAGGAATCAGTTGCGGTAGTCCGGGTTTCAAGCCCCGGAAGCCGCGAAAGC
 ATTGTGCACACTGACAGATGTGTACCTCCAGATCTTGAAGCTATCTCACCCGGAGACCCAGTCCAAGTGTGAAAA
 TGATGTTGGACTTCAAAGAAGTTGACTAATGGTCTGGAAGACAAAACAGCCTATTTCAAACCTGAAGGTGCTATTTTC
 ACCTGGTATCAGCTTGCCAGCTATGCTCGTACATCCGTGTTCCCGTCAACTCTACGGTGTACTTGAGCCCTGCATGGG
 CCCCCTCTTGAACAGGAGAGTGTGCGGTCCACCACTGGGGGGTGCCTCGCGGTACCCCTTATGATTACGGCG
 CTAAATATCTGTCTAGCGGTACCATGGTGAATGCCCCCGGATACAAATTTCTGGCGTGGCGGAGTTCTCGTTG
 GATGACCCAGTTAAGTACAAACATACCTGGGGTTTGAATCGGATACAGCGTATCTGTATGAGTTACCCGGAACGGTGA

Fig. 1L-4

GGACTGGGAGGATTACAATGATGCGTTTCGTGCGCGCCAGGAAGGGAAAAATTTATAAGGCCACTGCCACCAGCTTGAAGT
TTTTTTTTCCCCGGGCCCTGTCATTGAACCAACTTTAGGCCTGAATTGAAATGAAATGGGGTCCATGCAAAGCCTTTTT
GACAAAATTGGCCAACCTTTTGTGGATGCTTTCACGGAGTTCTTGGTGTCCATTGTTGATATCATTATATTTTTGGCCAT
TTTGTTTGGCTTACCATCGCCGGTTGGCTGGTGGTCTTTTGCATCAGATTGGTTTGCTCCGCGATACTCCGTACGCGCC
CTGCCATTCACTCTGAGCAATTACAGAAGATCTTATGAGGCCTTTCTTCCAGTGCCAAGTGGACATTCCACCTGGGG
AACTAAACATCCTTTGGGGATGCTTTGGCACCATAAGGTGTCAACCCTGATTGATGAAATGGTGTGCGCTGAATGTACC
GCATCATGGAAAAAGCAGGGCAGGCTGCTGGAAACAGGTGGTGAAGCGAGGCTACGCTGTCTCGCATTAGTAGTTTGAT
GTGGTGGCTCATTTTCAAGCATCTAGCCGCCATTGAAGCCGAGACCTGTAATATTTGGCCTCCCGGTGCCATGTCTACA
CAACTGCGCATGACAGGGTCAAATGTAACCATAGTGATAATAGCACTTTGAATCAGGTGTTTGCTATTTTTCCAACCC
CTGGTTCCCGGCCAAAGCTTCATGATTTTCAAGCAATGGTTAATAGCTGTACATTCTCCATATTTTCTCTGTTGCAGCT
TCTTGTACTCTTTTGTGTGCTGTGGTTGCGGGTCCAATACTACGTACTGTTTTGTTTTCCGCTGGTTAGGGGCAAT
TTTTCTTTCGAACTCACAGTGAATTACACGGTGTGTCCACCTTGCTCACC CGCAAGCAGCCACAGAGATCTACGAACC
CGGTAGGTCTCTTTGGTGCAGGATAGGGTATGACCGATGTGGGAGGACGATCATGACGAGCTAGGGTTTATGATACCGC
CTGGCTCTCCAGCGAAGGCCACTTGACTGGTGTGTACGCTGGTTGGCGTTCTTGTCTTCACTACAGGCCCACTTC
CATCCCGAGATATTCGGGATAGGGAATGTGAGTCGAGTTTATGTTGACATCAAACTCAACTCATCTGCGCGCAACATGA
CGGGCAGAACACCACCTTGCTCGTCTGATGACAACATTTCAAGCGTGTTCAGACCTATTACCAACATCAAGTCGACGGCG
GCAATTGGTTTCACTAGAAATGGCTTCGTCCTTTCTTCTCGTGGTTGGTTTTAAATGTCTCTTGGTTTCTCAGGCGT
TCGCTGCAAAACCATGTTTCAGTTCGAGTCTTGAGATATTAAGACCAACACCACCGCAGCGCAAGCTTTGCTGTCTC
CAAGACATCAGTTGCTTAGGCATCGCGACTCGGCCTCTGAGGCGATTGCAAAATCCCTCAGTGCCGTACGGCGATAGG
GACACCCGTGTATGTTACCATCACAGCAATGTGACAGATGAGAAATTTTACATTCTTCTGATCTCTCATGCTTTCTT
CTTGCCTTTCTATGCTTCTGAGATGAGTGAAAAGGGATTTAAGGTGGTATTTGGCAATGTGTGAGGCATCGTGGCTGTG
TGTGTCAATTTTACCAGCTACGTCCAACATGTCAAGGAGTTTACCACCGCTCCCTGGTGGTCGACCATGTGCGGTTGCT
CCATTTATGACACCTGAGACCATGAGGTGGGCAACTGTTTAGCCTGTCTTTTGCCATTCTGTTGGCAATTTGAATGT
TTAAGTATGTTGGAGAAATGCTTGACCGCGGCTGTTGCTCGCGATTGCTTTCTTTGGTGTATCGTGCCGTTCTGTT
TGCTGTGCTCGCAACGCCAGCAACGACAGCAGCTCCATCTACAGCTGATTTACAACCTGACGCTATGTGAGCTGAATG
GCACAGATGCTCACTAACAAATTTGATTGGGCACTGGAGAGTTTTGTATCTTTCCGTTTTGACTCACATTGTCTCC
TATGGTGCCCTCACTACCAGCCATTTCTTGACACAGTCGCTTGTAGTCACTGTGTCTACCGCGGGTTTGTTCACGGGCG
GTATGTCCTAAGTAGCATCTACGCGTCTGTGCCCTGGCTGCGTTGACTTGCTTCGTATTAGGTTTGCAAAGAATTGCA
TGTCTGGCGCTACGCGTGTACCAGATATAACCACTTTCTTCTGGACACTAAGGGCAGACTCTATCGTTGGCGGTGCGCT
GTATCATAGAGAAAAGGGGCAAAGTTGAGGTGCAAGGTCTGATGACCTCAAAAGAGTTGTGCTTGATGGCTCCGT
GGCAACCCCTATAACAGAGTTTCAAGCGAACAATGGGGTCTGCTTAGATGACTTCTGTACGATAGCACGGCTCCACA
AAAGGTGCTTTTGGCGTTTTCTATTACCTACACGCCAGTGATGATATATGCCCTAAAGGTGAGTCGCGGCCGACTGCTAG
GGCTTCTGCACCTTTTATCTTCTGAATTTGCTTTTACCTTCGGGTACATGACTTTTCGCGCACTTTCAAGGTACAAAT
AAGGTGCGCTCACTATGGGAGCAGTAGTTGCACTCTTTGGGGGTGTACTCAGCCATAGAAACCTGGAAATTCATCAG
CTCCAGATGCGGTTTGTGCTTGTAGGCCGCAAGTACATTCTGGCCCTGCCACCACGTTGAAAGTGCCGAGGCTTTC
ATCCGATTGCGGCAATGATAACACGCATTTGTCCGCGGTCCCGGCTCCACTACGGTCAACGGCACATTGGTGCC
GGGTTAAAAAGCCTCGTGTGGTGGCAGAAAAGCTGTTAAACAGGGAGTGGTAAACCTTGTCAAATATGCCAAATAACA
ACGGCAAGCAGCAGAAGAGAAAGAGGGGGATGGCCAGCCAGTCAATCAGCTGTGCCAGATGCTGGGTAAGATCATCGCT
CAGCAAAACAGTCCAGAGGCAAGGGACCGGGAAGAAAAATAAGAAGAAAAACCCGAGAGGCCCATTTTCTCTAGC
GACTGAAGATGATGTGACACATCACTTTACCCCTAGTGAGCGGCAATTGTGTCTGTCTGCAATCCAGACCGCCTTAATC
AAGGCGCTGGGACTTGACCCCTGTGAGATTCAAGGAGGATAAGTTACACTGTGGAGTTTGTGTTGCTACGCATCACT
GTGCGCTGATCCGCGTCAAGCATCACCTCAGCATGATGGGCTGGCATTCTTGAGGCATCTAGTGTTGAATTGGAA
GAATGTGTGGTGAATGGCACTGATTGACATTGTGCTCTAAGTCACCTATTCAATTAGGGCGACCGTGTGGGGTGAGAT
TTAATTGGCGAGAACCATGCGGCCGAAATTAATAAAAAA

Fig. 1M-1

>V7-Nsp2d324-726.seq

ATGACGTATAGGTGTTGGCTCTATGCCTTGGCATTGTATTGTGAGGAGCTGTGACCATTTGGCACAGCCCAAACTTGCT
 GCACAGAAACACCTTCTGTGATAGCCTCCTTCAGGGGAGCTTAGGGTTTGTCCCTAGCACCTTGCTCCGGAGTTGCAC
 TGCTTTACGGTCTCTCACCCCTTTAACCATTGTCTGGGATACTTGATCGGTGCACGTGTACCCCAATGCCAGGGTGTTT
 ATGGCGGAGGGCCAAGTCTACTGCACACGATGCCTCAGTGCACGGTCTCTCCTCCCTGAACCTCCAGGTTTCTGAGCT
 CGGGGTGCTAGGCTATTCTACAGGCCGAAGAGCCACTCCGGTGGACGTTGCCACGTGCATTCCCACTGTTGAGTGCT
 CCCCCGCCGGGGCTGCTGGCTTTCTGCAATCTTCCAATCGCACGAATGACCAAGTGGAAACCTGAACCTCCAACAAAGA
 ATGGTACGGGTGCGAGCTGAGCTTTACAGAGCCGGCCAGCTCACCCCTGCAGTCTTGAAGGCTCTACAAGTTTATGAACG
 GGGTTGCCGCTGGTACCCATTGTTGGACCTGTCCCTGGAGTGGCCGTTTTCCCAATTCCCTACATGTGAGTGATAAAC
 CCTTCCCGGGAGCAACTCACGTGTTGACCAACCTGCCGCTCCCGCAGAGACCAAGCCTGAAGACTTTTGCCCTTTGAG
 TGTGCTATGGCTACTGTCTATGACATTGGTCATGACGCCGTGATGATGTGGCCGAAAGGAAAGTCTCCTGGGCCCTCG
 TGGCGGGGATGAAGTGAATTTGAAGCTGTCCCGGGGAGTTGAAGTTGATTGCGAACCGGCTCCGCACCTCCTTCCCGC
 CCCACCACACAGTGGACATGTCTAAGTTCGCTTACAGCCCTGGGTGTGGTGTCTATGCGGGTCGAACGCCAACAC
 GGCTGCCTTCCCGTGACACTGTCCCTGAAGGCAACTGCTGGTGGAGCTTGTGACTTGTCTCCACTGGAAGTTTCAGAA
 CAAAGAAATTCGCCATGCTAACCAATTTGGCTACCAGACCAAGCATGGTGTCTCTGGCAAGTACCTGCAGCGGAGGCTGC
 AAGTTAATGGTCTCCGAGCAGTAACTGACCTAACCGGACCTATCGTCGTACAGTACTTCTCCGTTAAGGAGAGTTGGATC
 CGCCATTTGAACTGGCGGGAGAAACCCAGCTACTCTGGGTTTGAGGACCTCTCAGAATAAGGGTTGAGCCTAACACGTC
 GCCATTGGCTGACAAGGAAGAAAAATTTCCGGTTTGGCAGTCACAAGTGGTACGGCGCTGGAAAGAGAGCAAGAAAAAG
 CACGCTCTTGTGCGACTGCTACAGTGGCTGGCCGCGCTTTGTCCGTTCTGAAACCCGGCAGGCCAAGGAGCACGAGGTT
 GCCGGCGCAACCAAGGCTGAGCACCTCAAACTACTTCCCGCTGCCGAAGGGAATTGTGGTTGGCACTGCATTTCCGC
 CATCGCCAACCGGATGGTGAATTCAAATTTGAAACACCTTCCCGAAAGAGTGAGACCTCCAGATGACTGGGCTACTG
 ACGAGGATCTTGTGAATGCCATCCAATCTCAGACTCCTGCCGCTTAGACAGGAACGGTCTTGTACTAGCGCCAAG
 TACGTACTTAAGCTGGAAGGTGAGCATTGGAAGTGTCACTGTGACCCCTGGGATGTCCCTTCTTGTCCCTCTTGAATG
 TGTTACGGGCTGTTGTGGGCACAAGGGCGGTCTTGGTTCCCGAGATGCAGTGCAGGTCTCCGGATTTGACCTGCCTGCC
 TTGACCGGCTGGCTGAGGTGATGCACCTGCCTAGCAGTGTATCCAGCCGCTCTGGCCGAAATGTCTGGCGATTCCGAT
 CGTTCCGCTTCTCCGGTCAACACCGTGTGGACTGTTTGCAGTCTTTGCCGTCACAGCGGAGGGAATCACCTGACCA
 AGTGGCTTAGGGAAAATTATCAGCCTTTGTGAGGTGATTGAGGACTGCTGCTGTCCAGAACAAAACCAACCGGGTCA
 CCCCCGAGGAGGTGCGAGCAAGATTGACCTGTACCTCCGTGGTGCAACAAATCTTGAAGAATGCTTGGCCAGGCTTGA
 AAAGCGCGCCGCCACGCGTAATCGACACCTCCTTGTATTGGGATGTTGTGCTCCCTGGGGTTGAGGCGGCAACCCAGAC
 GATCAAGCTGCCCCAGGTCAACCAAGTGTGCTGCTGGTCCCTGTTGTGACTCAAAAGTCTTGTGTGAGTTTGTGATGA
 TGCCTCACACGCTGCACCTTCCGTAGGTGCGGAGAGCGACCTTACCATTGGCTCAGTTGCTACTGAAGATGTTCCACGC
 ATCCTCGAGAAAATAGAAAATGTGCGGAGATGGCCAACCAAGGACCTTGGCCTTCTCCGAGGATAAACCGGTAGATGA
 CCAACTTGTCAACGACCCCCGGATATCGTCGCGGAGGCTGACGAGAGCACATCAGTCCGTCCGAGGCACAGGTGGCG
 CCGGCTCTTTTACCGATTGCGCCTTCAAGTGGCGGATGCGGACGGGGGGGGCGTTTCGGACGGTAAAAAGAAAA
 GCTGAAAGGCTCTTGAACCACTGAGCGTCAAGTTTGTGACCTCGTCTCCATCTCCCTGTTTTCTTCTACGCTTTT
 CTACCTGGCGGTGGTTATTCTCCGGGTGATTGGGGTTTGCAGCTTTTACTCTATTGTGCCTTTTTATGTTACAGTT
 ACCCAGCCTTTGGTATTGCTCCCTCTTGGGTGTGTTTTCTGGGCTTCTCGGCGGTTTGAATGGGGGTTTTTGGCTGC
 TGGTTGGCTTTTGTGTTGGTCTGTTCAAGCCTGTGTCGACCCAGTGGCGCTGCTTGTGAGTTTGAATGCCAGAGTG
 TAGAAACATCCTTCATTCTTTGAGCTTCTAAACCTTGGGACCTGTTGCGAGCCTTGTGTTGGGCCCCGTGGTCTCG
 GTCTTGGCATTCTTGGCAGGTTACTGGGCGGGGACGCTGCATCTGGCACTTTTGTGTTAGGCTTGGCATTGTTGCAGAC
 TGTATCTTGGCTGGAGCTTACGTGCTTTCTCAAGGTAGGTGTAAGAAAGTGTGGGATCTTGTATAAGAACTGCTCCCAA
 TGAGGTGCTTTTAAAGTGTGCTTTCTCAACGTCGACGACCAAGTGTCACTTATCGACCTGTGCGATCGGTTTTGTGCGC
 CAAAAGGAATGGACCCCATTTTCTCGCCACTGGGTGGCGGGGTGCTGGGCGGGCGAAGCCCCATTGAGCAACCTCT
 GAAAAACCATCGCGTTTGGCCAGTTGGATGAAAAGAAAGATTACGGCTAGGACTGTGGTGCAGCCGCTTATGACCCCAA
 CCAAGCCGTAAAGTGTGCTGCGGGTATTGACGGCGGGTGGGGCGATGGTGGCTAAGGCGGTCCCAAAGTGGTCAAGTTT
 CCGCTGTTCCATTCCGAGCCCCCTTCTTCCCACTGGAGTGAAAGTTGACCCTGATTGCAGGGTCTGGTTGACCTGAC
 ACTTCACTGCAGCTCTCCGGTCTGGCTACTCCACCACAAACCTCGTCTTGGTGTGGGGGACTTTGCCAGCTGAATGG
 ATTAATAATCAGGCAATTTCAAGCCTTCAAGGGGAGGGCCACATCTCATGGCTGCCCTGCATGTTGCTGCTCGATGG
 CTCTGCACATGCTTGGTGGGATTATGTGACTGCGGTGGGTTCTTGGGCACCGGCACCAACGACCCGTGGTGGCTAAC

Fig. 1M-2

CCGTTTGCCGTCCTGGCTACGGACCTGGCTCTCTGACAGTCCAGATTGTGCATTTCCCAACACGGCCTTACCCTGCC
 CTTGACAGCACTTGTGGCGGGATTCCGGTATTCAAGAAATTGCCTTGGTCGTTTGTATTTTGTTCATCGGAGGCATGG
 CTCATAGGTTGAGCTGTAAGGCTGACATGCTGTGTCTTGTGCAATTGCCAGCTATGTTTGGGTACCTCTTACCTGG
 TTGCTTTGTGTGTTTCTTGTGCTGGTTGCGCTGTTTTCTTGCACCCCTCACCATCCTATGGTTGGTGTCTTCTTGAT
 TTCTGTGAATATGCCTTCAGGAATCTTGGCCATGGTGTGTTGGTTCTCTTGGCTTCTTGGTCGTTATACTAATGTTG
 CTGGCCTTGTACCCCTACGACATTCATCATTACACCAGTGGCCCCCGGGTGTGCCGCTTGGCTACCGCACCAGAT
 GGGACCTACTTGGCCGCTGTCCGCCGCGTGGCTGACTGGCCGACCATGCTGTTTACCCCGTCCAGCTTGGGTCTCT
 TCTTGAGGGTGTCTTCAAGACTCGAAAGCCCTCACTGAACACCGTCAATGTGATCGGGTCTCCATGGGCTCTGGCGGG
 TGTTTACCATCGACGGGAAAGTCAAGTGCCTAAGTCCCGACATGTCCTTACGGGCAATTCAGCTCGGGTTTCCGGGGT
 GGCTTCAATCAATGCTTGACTTTGACGTAAAGGGAGATTTCGCTATAGCTGATTGCCGAATTTGGCAAGGGGCTGCC
 CAAGACCAATTCGACGGATGGATGGACTGGCCGTGCCTATTGGCTAACATCCTCTGGCGTCAACCCGGCGTCATTG
 GAAAGGATTGCTTCTGCTTACCGCATGTGGCGATTCCGGGTCCCACTGATCACCAGGGCCGGTGAAGCTTGTGCG
 GTTACACGGGATCGAATAAACAAGGGGGGGGCTTGTACGCGCCCTCAGGCCAGTTTGTAAATGTGGCACCCTCAA
 GCTAAGCGAATTAAGTGAATTTCTTGTGGCCTAAGGTCCGCTCGGTGATGTGAAGGTGGCAGCCACATAATTAAAG
 ACATAAGCGAGGTGCCTTCAGATCTTGTGCTTGTCTGTCGCAAACTGAACTGGAAGGAGGCTCTCCACCGTCCAA
 CTCTTTGTGTGTTTTCTCTGTGGAGAAATGATGGACATGCCTGGACGCCCTTGGTTGCTGTGAGTTTCTTATTTT
 GAATGAGGTTCTCCAGCCGTCTGGTCCGGAGTGTCTCTCTTGGAAATGTTGTGCTATCTGGCTCACGCCATGGT
 CTGCGCAAGTTCTGATGATCAGGCTTCTGACAGCAGCTCTTAACAGGAACAGATGGTCACTTGCCTTTTACGCTCGGT
 GCAGTGACCGGTTTTGTGCGAGATCTTGGGCCACTCAGGGGCATCCGTTGACGGCAGTGATGAATTTGAGCACCTATGC
 ATCTCTGCCTCGGATGATGGTTGTGACCTCACCAGTCCAGTGATCAGGTGGTGTGCTGCACCTACTTGCATCATTT
 TGTACTTGTAAAGTACCGTGGCCTGCACCATATCTTGTGGCGATGGAGTGTCTCTGCGGCTTCTCTTGTAGATAC
 TTTGCCGAGGGAAAGTTGAGGGAAGGGGTGTGCAATCCTGCGGAATGAATCATGAGTCTCTGACTGGTGCCTCGCTAT
 GAGACTCAATGACGAGGACTTGGATTTCTTATGAAATGGACTGATTTAAGTGCTTTGTTTCTGCGTCCAACATGAGGA
 ATGACGCGGGTCAATTTATCGAGGCTGCCTATGCTAAAGCACTTAGAGTAGAAGTGGCCAGTTGGTGCAGGTTGATAAA
 GTTCGAGGTACTTTGGCCAACTTGAAGCTTTTGTGATACCGTGGCACCTCAACTCTCGCCCGGTGACATTGTTGTGCG
 TCTCGGCCACACGCTGTTGGCAGTATCTTCGACCTAAAGGTTGGTAGCACAAGCATACCTCCAAGCCATTGAGACCA
 GAGTCTTGTCTGGTCCAAATGACCGTGGCGCGCTGTGACCCGACCCCAAGCCCAAGCCCAAGCCCAAGCCCAAGCCCAAG
 CCCCTCCCAAGGATTTCTGGAGAATGGCCCAAGCTTGGGGGATGAGGACCGTTTGAATAAGAGAAGAGGCGCAG
 GATGGAAGCCCTCGGCATCTATGTTATGGCGGGGAAAAATACAGAAATTTTGGGACAAGAATTCGGGTGATGTGTTTT
 ATGAGGAGGTCCATAATAACACAGATGAGTGGGAGTGTCTCAGAGTTGGCGACCCTGCCGACTTGAACCTGAGAAGGGA
 ACTCTGTGTGGACATGTCAACATTGAAACAAGGCTTACCATTGTTACACCTCCCATCTGGTAAGAAGTTCTTGGTCCC
 CGTCAACCCAGAGAAATGGAAGAGTCCAATGGGAAGCTGCAAGCTTTCCGTGGAGCAGGCCCTAGGTATGATGAATGTG
 ACGGCGAACTGACTGCCAAAGAACTGGAGAACTGAAAGAATAATTGACAACTCCAGGGCTGACTAAGGAGCAGTGT
 TTAAGTCTAGCGCCAGCGACTTGACCCGCTGTGGTGCAGCGGCTTGGTTGTTACTGAAACAGCGGTAAGAAATAGTC
 AAATTTACAACCGACCTTCAACCTGGGACCTGTGAATTTAAAGTGGCCAGTGAGGTTGAGCTAAAAGACGCGGTTGA
 GCACAACCAACCCGGTTGCGAGACCGATCGATGGTGGAGTTGTGCTCCTGCGTTCCGCGGTTCTTCTGCTTATAGACG
 TCTTGATCTCCGGTGTGATGCATCTCCCAAGTTACTTGGCCATCAGGGCCGGGAAACACTGGGATCGATGGCACGCTC
 TGGGATTTTGAAGTCCGAAGCACTAAAGAGGAAGTGCACCTCAGTGCAGCAATAATACAGGCTTGTGACATTAGGCGCG
 CGACGCTCCTGAAATTTGGTCTCCCTTACAAGCTGTACCTGTTAGGGGTAACCTGAGCGGGTGAAGGAGTTCTGCAGA
 ATACAAGGTTTGGAGACATACCTTACAAAACCCCAAGTGAAGTGGAGCCAGTGCACGCGGCTGCCTTACGCCC
 AACGCCACTCCGGTGAAGTGTGGGCTCCGTCTTGGCCACGACCATGCCCCCGGGTTTGAAGTTATATGTACCGACCAT
 ACCAGCTGTGCTTGTATTACCTTGAAGTGTAGGCTGACTGCCCTAAACAGCTGACAGAGCAGGCTGCGAAGATGCCG
 CACTGAAAGACCTCTCTAAATATGACTTGTCCACCAAGGCTTGTGTTTACCTGGAGTTCTTCTGCTTGTGCGGAAATAC
 CTGTTTGGCCATGTAGGTAAGTGGCCACCGTTTCTGCGCTTCTACTTACCCTGCTAAGAAATCTATGGCTGGAATAAA
 TGGGAACAGGTTCCCAACCAAGGACATTGAGAGCTCCCTGAAATCGACGTTCTGTGCGCACAGGCTGTGCGGAAAACT
 GGCAAACTGTACCCCTTGTACTCTTAAGAAACAGTATTGCGGGAAGAAGAGTACTAGGACATACTCGGCACCAATAAC
 TTCATCGCACTAGCCACCGAGCAGTGTGAGTGGTGTACCCAGGGCTTCTGAAAAAGGCGTTAACTCGCCCATCGC
 CCTCGGAAAGAAAGTTTAAAGGAGTACAGACTCCGGTCTGGGCGAGTGCCTTGAAGCTGATCTCGATCCTGCGATC
 GATCCACGCTGCAATTGTCCGCTGGTTTGGCGCAACCTTCTTATGAACTTGCTGTGCTGAAGAGCATCTACCGTCG

Fig. 1M-3

TACGTGCTGAAGTCTGCCACGACTTACTGGTCACGCAGTCCGGCGCAGTGACTAAGAGAGGTGGCCTGTCGTCTGGCGA
 CCCGATCACCTCTGTGTCTAACACCATTATAGTTTGGTGATCTATGCACAGCATATGGTGCTTAGTTACTTCAAAAGTG
 GTCACCCCATGGCCTTCTGTTCTTACAAGACCAGCTAAAGTTTGAGGACATGCTCAAGGTTCAACCCCTGATCGTCTAT
 TCGGACGACCTCGTGCTGTATGCCGAGTCTCCACCATGCCAACTATCACTGGTGGGTGAACATCTGAATTTGATGCT
 GGGGTTTCAGACGGACCCAAAGAAGACAGCAATAACAGACTCGCCATCATTTCTAGGCTGTAGAATAATAATGGGCGCC
 AGCTAGTCCCCAACCGTGACAGGATCCTCGCGGCCCTCGCCTATCACATGAAGGCGAGTAATGTTTCTGAATACTATGCC
 TCAGCGGCTGCAATACTCATGGACAGCTGTGCTTGTGGAGTATGATCCTGAATGGTTTGAAGAAGTTGAGTTGGAAT
 AGCGCAGTGCGCCCGCAAGGACGGCTACAGCTTCCCGGCACGCCGTTCTTCATGTCCATGTGGGAAAACTCAGGTCCA
 ATTATGAGGGGAAGAAGTCGAGAGTGTGCGGTAAGTGCAGGGCCCGGCCCGTACGCTACTGCCTGTGGCCTCGACGTC
 TGCATTTACCACACCCACTTCCACCAGCATTGTCCAGTCACAATCTGGTGTGGCCATCCAGCGGGTCTGGTTCTTGAG
 TGAGTGCAAAATCCCTGTAGGGAAAGGCACAAGCCCTTTAGACGAGGTGCTGGAACAAGTCCCGTATAAGCCCCACGGA
 CCGTTATCATGCATGTGGAGCAGGGTCTACCCCTTGTATCCAGGTAGATACCAACTCGCCGCGGATTAGTCTCTGTC
 AGGCGTGGAATTAGGGGAAATGAAGTTGGACTACCAGACGGTGATTATGCTAGCACCGCCTTCTCCCTACCTGCAAGA
 GATCAACATGGTCGTGTGCTTCCAATGTATTGCGCAGCAGGTTTCATCATCGGCCACCCGGTGTGGGAAACATACT
 GGCTCCTTCAACAGGTCCAGGATGGTGTATTATTTACACCAACTCACCAGACCATGCTTGACATGATTAGGGCTTTG
 GGGACGTGCCGGTCAACGTCCCGCAGGCACAACGCTGCAATTCCTCGTCCCTCCCGCACCAGGTCCGTGGGTTGCGAT
 CCTAGCCGGCGGTTGGTGTCTGGCAAGAATCCTTCTAGATGAAGCAGCGTATTGCAATCACCTTGATGTTTTGAGGC
 TTCTTAGTAAACTACCTCACCTGTCTAGGAGACTCAAGCACTCCACCCAGTGGGTTTTGATTCTCATTGCTATGTT
 TTTGACATCATGCCTCAAACTCACTGAAGACCATCTGGAGGTTTGACAGAAATATCTGTGATGCCATTAGCCAGATTA
 CAGGGACAACTCATGTCCATGGTCAACACAACCCGTGTGACCTACGTGGAAAACTGTGAGGTATGGGAGGTCTCA
 CCCCCTACCACAGGGACCGAGAGGACGACGCCATCACTATTGACTCCAGTCAAGGCGCCACATTGATGTGGTTACATTG
 CATTTGCCCACTAAAGATTCACTCAACAGGCAAGAGCCCTTGTGCTATCACCAGGGCAAGACGCTATCTTTGTGTA
 TGACCCACACAGGCAGCTGCAGGGCTTGTGATCTTCTGCAAAAGGCACGCCGTCAACCTCGCAGTGCACTGCGACG
 GGCAGCTGATCGTGCTGGATAGAAATAACAAAGATGCACGGTTGCTCAGGCTTAGGCAACGGGGATAAATTTAGGGCC
 ACAGACAAGCGTGTGTAGATTCTCCTCGCGCATTTGTGCTGATCTAGAAGGGTCGAGCTCTCCGCTCCCAAGGTGCG
 ACACAACCTGGGATTTTATTTCTCACCTGATTTAACACAGTTTGCTAAACTCCAGTAGAACTTGCACCTCACTGGCCG
 TGGTGTCAACCCAGAACAATGAAAAGTGGCCGGATCGGCTGGTTGCCAGCCTTCGCCCTATCCATAAATACAGCCGCGG
 TGCATCGGTGCCGCTATATGGTGGGCCCTTGGTGTTCTAGGCACTCCTGGGTCGTGTACTATCTCACAAATTT
 TGTTAAGGGCGGGCTCAAGTGCTCCGGAGACGGTTTTAGCAGCCGGCGAATTGAGGTAGACTGCCGGGAATATCTTG
 ATGATCGGGAGCGAGAAGTTGCTGCGTCCCTCCACACGCTTTCATTGGCGACGTCAAAGGCACTACCGTTGGAGGATGT
 CATCATGTCACTCCAGATACCTCCCGCGCTCTTCCCAAGGAATCAGTTGCGGTAGTCGGGTTTCAAGCCCCGAAA
 AGCCGCGAAAGCATTGTGCACACTGACAGATGTGACCTCCAGATCTTGAAGCCTATCTCCACCCGAGACCCAGTCCA
 AGTGCTGAAAAATGATGTTGGACTTCAAAGAAGTTCGACTAATGGTCTGGAAGACAAAACAGCCTATTTCCAACCTGAA
 GGTGCTATTTACCTGGTATCAGCTTGCCAGCTATGCCTCGTACATCCGTGTTCCCGTCAACTCTACGGTGACTTGGA
 CCCCTGCATGGGCCCCGCCCTTTGCAACAGGAGAGTCGTGCGGTCCACCCACTGGGGGGCTGACCTCGCGGTACCCCTT
 ATGATTACGGCGCTAAAATTATCCTGTCTAGCGCGTACCATGGTGAAATGCCCCCGGATACAAAATCTGGCGTGCGCG
 GAGTTCTCGTTGGATGACCCAGTTAAGTACAAACATACCTGGGGGTTTGAATCGGATACAGCGTATCTGTATGAGTTCAC
 CGGAAACGGTGAGGACTGGGAGGATTACAATGATGCGTTTCTGCGCGCCAGGAAGGGAAATTTATAAGGCCACTGCCA
 CCAGCTTGAAGTTTTATTTTCCCCGGGCCCTGTATTGAACCAACTTTAGGCCTGAATTGAAATGAAATGGGGTCCATG
 CAAAGCCTTTTGACAAAATTTGGCCAACTTTTGTGGATGCTTTCACGGAGTCTTGGTGTCCATTGTTGATATCATTAT
 ATTTTGGCCATTTTGTGTTGGCTTACCATCGCCGTTGGCTGGTGGTCTTTGTCATCAGATTGGTTTGTCTCCGCGATAC
 TCCGTACGCGCCCTGCCATTCACTCTGAGCAATTACAGAAGATCTTATGAGGCCTTTCTTCCAGTGCCAAGTGGACAT
 TCCCACCTGGGGAACATAACATCCTTTGGGGATGCTTTGGCACCATAAGGTGTCAACCCTGATTGATGAAATGGTGTGCG
 GTCGAATGTACCCATCATGGAAAAAGCAGGGCAGGCTGCCTGGAAACAGGTGGTGAGCGAGGCTACGCTGTCTCGCATT
 AGTAGTTTGGATGTGGTGGCTCATTTTTCAGCATCTAGCCGCCATTGAAGCCGAGACCTGTAATATTTGGCCTCCCGGCT
 GCCCATGTACACAACCTGCGCATGACAGGGTCAAATGTAACCATAGTGTATAATAGCACTTTGAATCAGGTGTTTGCTA
 TTTTCCAACCCCTGGTTCCCGGCCAAAGCTTCATGATTTTTCAGCAATGGTTAATAGCTGTACATTCTCCATATTTTCC
 TCTGTTGACGCTTCTGTACTCTTTTGTGTGCTGTGGTTGCGGGTCCAATACTACGTAAGTGTGTTTGGTTTCCGCTG
 GTTAGGGCAATTTTCTTTCGAACTCACAGTGAATTACACGGTGTGTCCACCTTGCTCACCAGCAAGCAGCCACAGA

Fig. 1M-4

GATCTACGAACCCGGTAGGTCTCTTTGGTGCAGGATAGGGTATGACCGATGTGGGGAGGACGATCATGACGAGCTAGGGT
 TTATGATACCGCTGGCCTCTCCAGCGAAGGCCACTTGACTGGTGTTCACGECTGGTTGGCGTTCTTGTCCTTCAGCTAC
 ACGGCCCAGTTCCATCCCGAGATATTCGGGATAGGGAATGTGAGTCGAGTTTATGTTGACATCAAACATCAACTCATCTG
 CGCCGAACATGACGGGCAGAACACCACCTTGCCCTCGTCATGACAACATTCAGCCGTGTTTCAGACCTATTACCAACATC
 AAGTCGACGGCGGCAATTGGTTTCACCTAGAATGGCTTCGTCCCTTCTTTCTCGTGGTTGGTTTTAAATGTCTCTTGG
 TTTCTCAGGCGTTTCGCCTGCAACCATGTTTCAGTTCGAGTCTTGACAGATTAAGACCAACACCACCGCAGCGGCAAGC
 TTTGCTGTCTCCAAGACATCAGTTGCCTTAGGCATCGCGACTCGGCCTCTGAGGCGATTGCGAAATCCCTCAGTGCCG
 TACGGCGATAGGGACACCCGTGTATGTTACCATCACAGCCAATGTGACAGATGAGAATTATTTACATTCTTGATCTCC
 TCATGCTTTCTTCTGCTTTCTATGCTTCTGAGATGAGTGAAAAGGGATTAAAGGTGGTATTTGGCAATGTGTGAGGC
 ATCGTGGCTGTGTGTCAATTTTACCAGCTACGTCACATGTCAAGGAGTTTACCCAACGCTCCCTGGTGGTCGACCA
 TGTGCGGTTGCTCCATTTATGACACCTGAGACCATGAGGTGGGCAACTGTTTTCAGCTGTCTTTTGGCATTCTGTTGG
 CAATTTGAATGTTTAAATGATGTTGGAGAAATGCTTGACCGCGGGCTGTTGCTCGCGATTGCTTTCTTTGGTGTATCGT
 GCCGTTCTGTTTGTCTGCTCGCAACGCCAGCAACGACAGCAGCTCCCCTACAGCTGATTACAACCTTGACGCTAT
 GTGAGCTGAATGGCACAGATTGGCTAGCTAACAAATTTGATTGGGCAGTGGAGAGTTTGTGTCATCTTTCCCGTTTGA
 CACATTGTCTCCTATGGTGCCCTCACTACCAGCCATTTCTTGACACAGTCGCTTTAGTCACTGTGTCTACCGCGGGT
 TGTTACGGGCGGTATGTCCTAAGTAGCATCTACGCGGTCTGTGCCCTGGCTGCGTTGACTTGCTTCGTATTAGGTTTG
 CAAAGAATTGCATGTCCTGGCGCTACGCGGTGACAGATATACCAACTTTCTTCTGGACACTAAGGGCAGACTCTATCGT
 TGGCGGTGCGCTGTATCATAGAGAAAAGGGCAAGTTGAGGTGGAAGGTGATCTGATCGACCTCAAAGAGTTGTGCT
 TGATGGCTCCGTGGCAACCCCTATAACCAGAGTTTCAGCGGAACAATGGGGTCGTCCTTAGATGACTTCTGTACGATAG
 CACGGCTCCACAAAAGGTGCTTTTGGCGTTTCTATTACCTACACGCCAGTGATGATATATGCCCTAAAGGTGAGTCGCG
 GCCGACTGTAGGGCTTCTGCACCTTTTGTCTTCTGAATTTGCTTTTACCTTCGGGTACATGACTTTTCGCGCACTTT
 CAGAGTACAAATAAGGTGCGGCTCACTATGGGAGCAGTAGTTGCACTCCTTTGGGGGTGTACTCAGCCATAGAAACCTG
 GAAATTCATCACCTCCAGATGCCGTTTGTGCTTGTAGGCCGCAAGTACATTCTGGCCCTGCCACCACTTGAAAGTG
 CCGCAGGCTTTATCCGATTGCGGCAATGATAACACGCAATTTGTCTCGTCCGGCTCCCGGCTCCACTACGGTCAACGGC
 ACATTGGTGCCCGGGTAAAAAGCCTCGTGTGGGTGGCAGAAAGCTGTTAAACAGGGAGTGGTAAACCTTGTCAAATA
 TGCCAAATAACAACGGCAAGCAGCAGAGAGAAAGAGGGGGATGGCCAGCCAGTCAATCAGCTGTGCCAGATGCTGGGT
 AAGATCATCGCTCAGCAAAACAGTCCAGAGGCAAGGGACCGGAAAGAAAAATAAGAAGAAAAACCCGGAGAGCCCA
 TTTTCTCTAGCGACTGAAGATGATGTGACACATCACTTTACCCCTAGTGAGCGGCAATTGTGTCTGTCTGTCATCCAGA
 CCGCTTTAATCAAGGCGCTGGGACTTGACCCCTGTGAGATTGAGGGAGGATAAGTTACACTGTGGAGTTAGTTTGCCT
 ACGCATCATACTGTGCGCTGATCCGCTCACAGCATCACCTCAGCATGATGGGCTGGCATTCTTGAGGCATCTCAGTG
 TTTGAATTGGAAGAATGTGTGGTGAATGGCACTGATTGACATTGTGCTCTAAGTCACCTATTCAATTAGGGCGACCGTG
 TGGGGTGAGATTTAATTGGCGAGAACCATGCGGCCGAAATTAATAAAAAA

Fig. 2

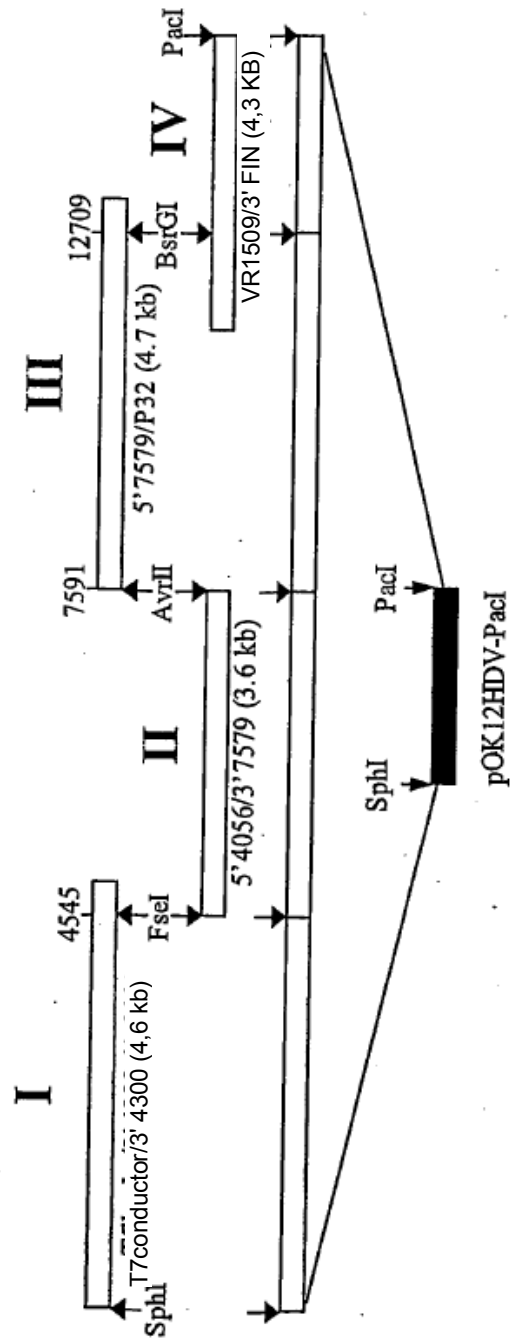


Fig. 3

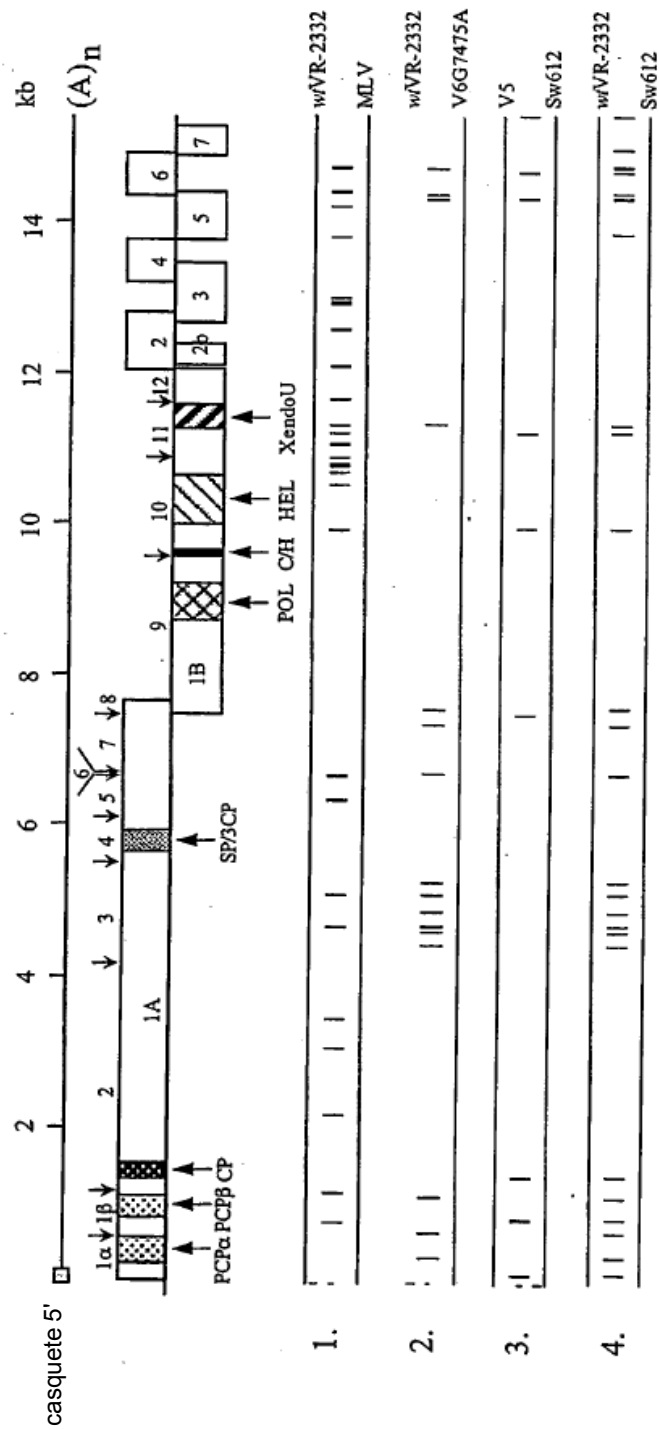


Fig. 4

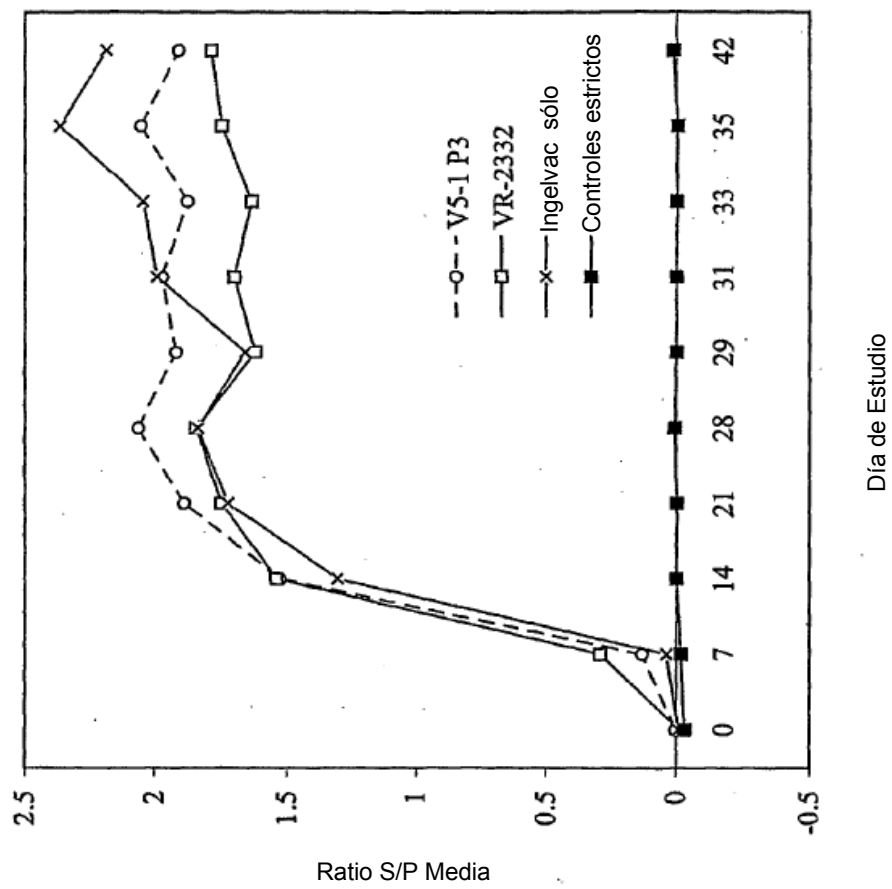


Fig. 5A

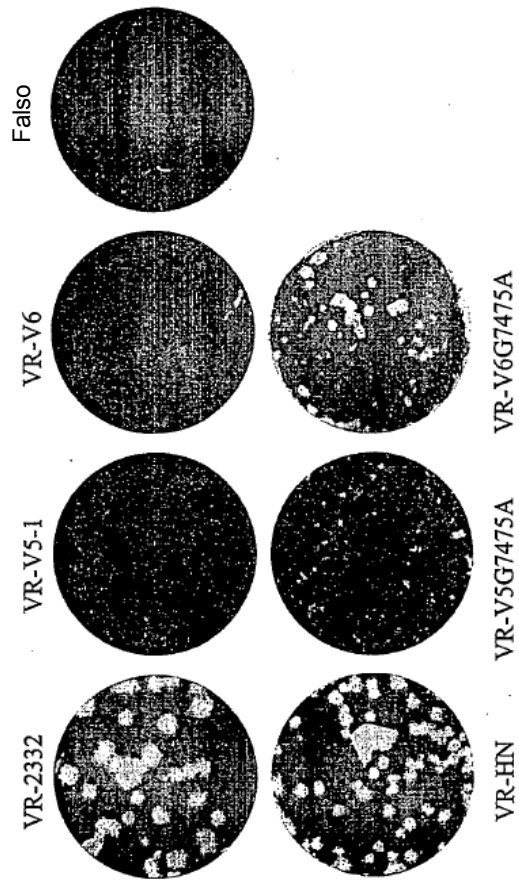


Fig. 5B

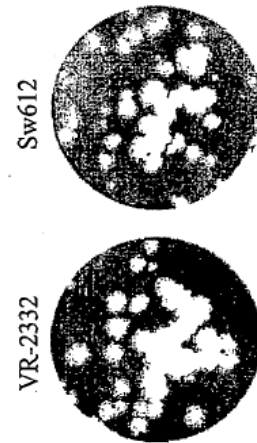


Fig. 6A

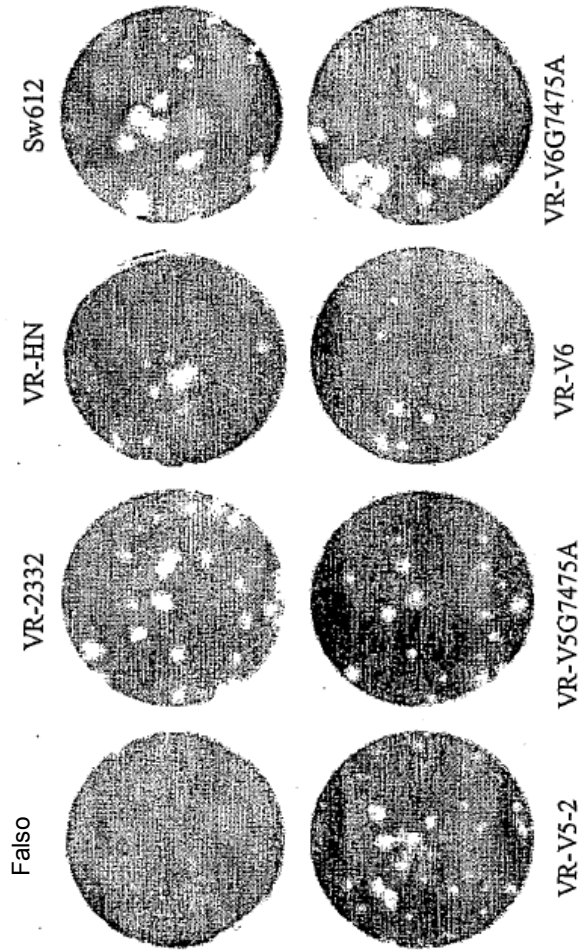


Fig. 6B

	VR-2332	Sw612	VR-HN	VR-V5-2	VR-V5G7475A	VR-V6	VR-V6G7475A	Falso
Título de Virus	2.5×10^5	9.8×10^4	1.5×10^4	9.8×10^3	2×10^3	1.5×10^3	3.25×10^3	-

Fig. 7A

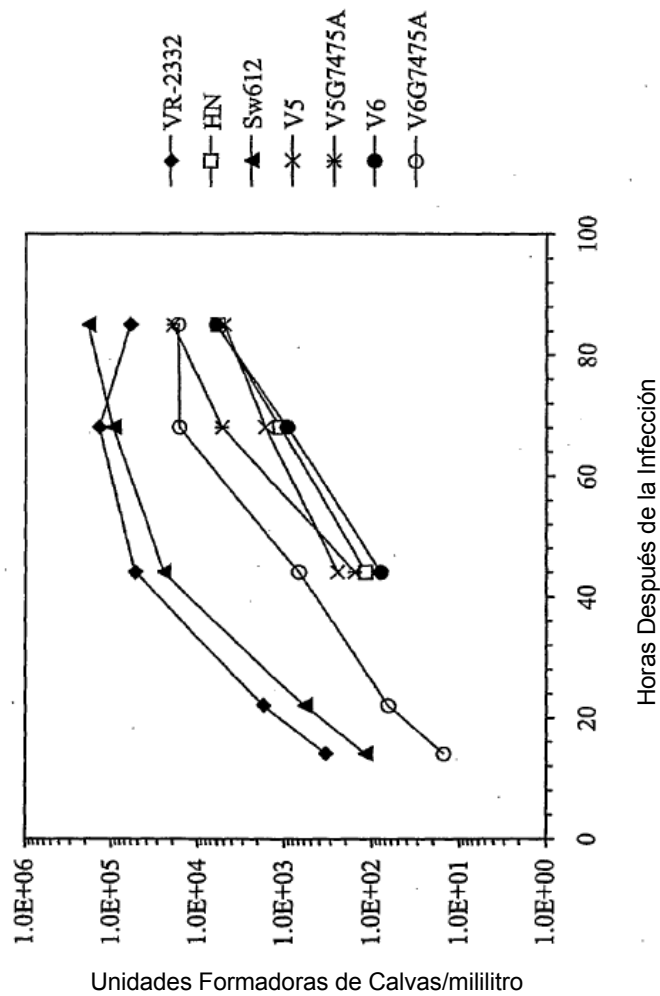
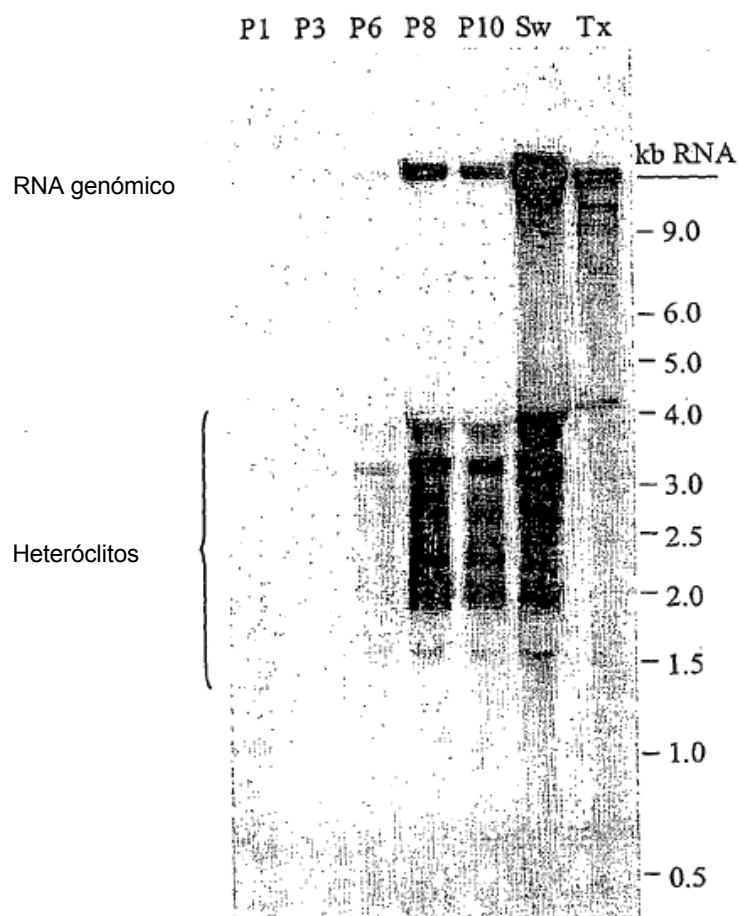


Fig. 7B

	VR-2332	Sw612	VR-HN	VR-V5	VR-V5G7475A	VR-V6	VR-V6G7475A	Falso
Título de Virus	8.12×10^4	2.0×10^5	5.5×10^3	5.0×10^3	2.5×10^4	4.4×10^3	1.9×10^4	

Fig. 8



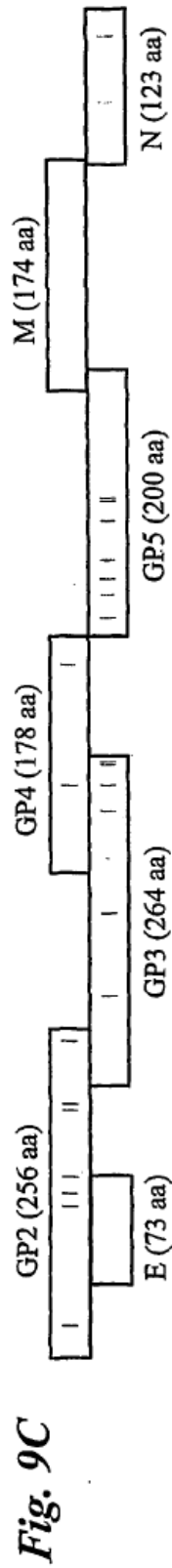
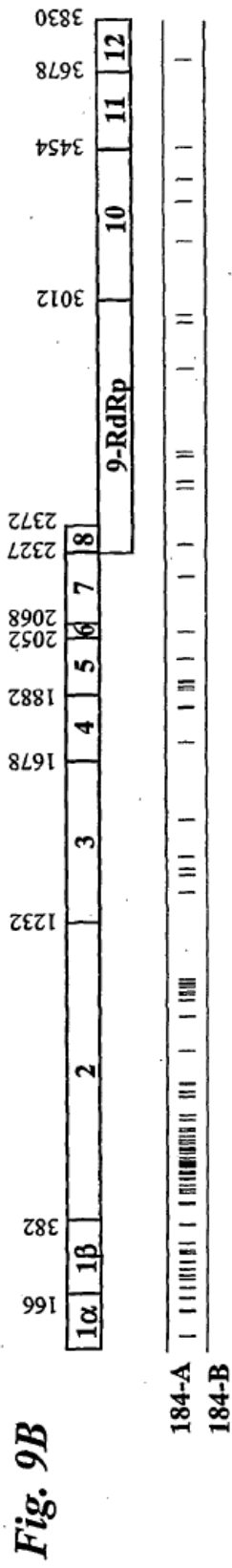
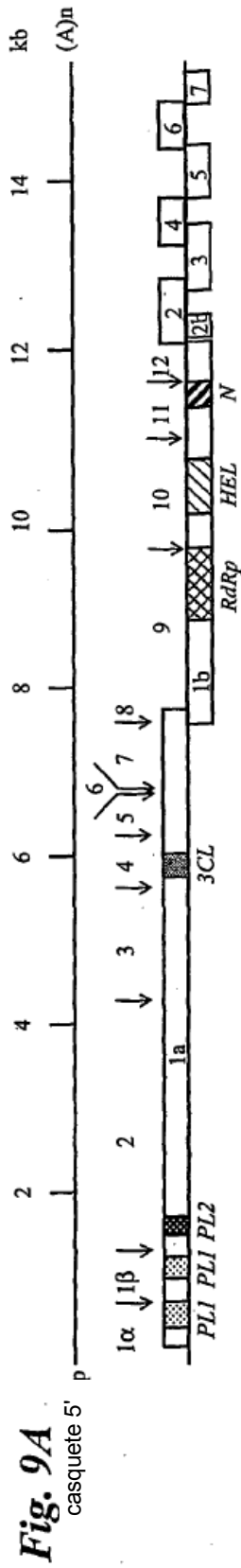


Figura 11

EuroPRRSV	QACROPFCPPEE	NSG	VRWKKFVIFSDSP	NGOSRI	WTKSDSAA	IELEP	ETEROVEI	IRSPRAH	PINLADEL	USHEN	EFNNTSHSC	HIRVNSM	POKXAL	112													
Lelystad	QACROPFCPPEE	NSG	VRWKKFVIFSDSP	NGOSRI	WTKSDSAA	IELEP	ETEROVEI	IRSPRAH	PINLADEL	USHEN	EFNNTSHSC	HIRVNSM	POKXAL	112													
MN184A	RPKPEDFCPPEE	NSG	VRWKKFVIFSDSP	NGOSRI	WTKSDSAA	IELEP	ETEROVEI	IRSPRAH	PINLADEL	USHEN	EFNNTSHSC	HIRVNSM	POKXAL	106													
MN184B	RPKPEDFCPPEE	NSG	VRWKKFVIFSDSP	NGOSRI	WTKSDSAA	IELEP	ETEROVEI	IRSPRAH	PINLADEL	USHEN	EFNNTSHSC	HIRVNSM	POKXAL	106													
Ingelvac	RPKPEDFCPPEE	NSG	VRWKKFVIFSDSP	NGOSRI	WTKSDSAA	IELEP	ETEROVEI	IRSPRAH	PINLADEL	USHEN	EFNNTSHSC	HIRVNSM	POKXAL	106													
VR-2332	RPKPEDFCPPEE	NSG	VRWKKFVIFSDSP	NGOSRI	WTKSDSAA	IELEP	ETEROVEI	IRSPRAH	PINLADEL	USHEN	EFNNTSHSC	HIRVNSM	POKXAL	106													
PL97-1	RPKPEDFCPPEE	NSG	VRWKKFVIFSDSP	NGOSRI	WTKSDSAA	IELEP	ETEROVEI	IRSPRAH	PINLADEL	USHEN	EFNNTSHSC	HIRVNSM	POKXAL	106													
HN1	RPKPEDFCPPEE	NSG	VRWKKFVIFSDSP	NGOSRI	WTKSDSAA	IELEP	ETEROVEI	IRSPRAH	PINLADEL	USHEN	EFNNTSHSC	HIRVNSM	POKXAL	106													
16244B	RPKPEDFCPPEE	NSG	VRWKKFVIFSDSP	NGOSRI	WTKSDSAA	IELEP	ETEROVEI	IRSPRAH	PINLADEL	USHEN	EFNNTSHSC	HIRVNSM	POKXAL	106													
PA-8	RPKPEDFCPPEE	NSG	VRWKKFVIFSDSP	NGOSRI	WTKSDSAA	IELEP	ETEROVEI	IRSPRAH	PINLADEL	USHEN	EFNNTSHSC	HIRVNSM	POKXAL	106													
SP	RPKPEDFCPPEE	NSG	VRWKKFVIFSDSP	NGOSRI	WTKSDSAA	IELEP	ETEROVEI	IRSPRAH	PINLADEL	USHEN	EFNNTSHSC	HIRVNSM	POKXAL	106													
JA142	RPKPEDFCPPEE	NSG	VRWKKFVIFSDSP	NGOSRI	WTKSDSAA	IELEP	ETEROVEI	IRSPRAH	PINLADEL	USHEN	EFNNTSHSC	HIRVNSM	POKXAL	106													
CH-1	RPKPEDFCPPEE	NSG	VRWKKFVIFSDSP	NGOSRI	WTKSDSAA	IELEP	ETEROVEI	IRSPRAH	PINLADEL	USHEN	EFNNTSHSC	HIRVNSM	POKXAL	106													
P129	RPKPEDFCPPEE	NSG	VRWKKFVIFSDSP	NGOSRI	WTKSDSAA	IELEP	ETEROVEI	IRSPRAH	PINLADEL	USHEN	EFNNTSHSC	HIRVNSM	POKXAL	106													
HB-2	RPKPEDFCPPEE	NSG	VRWKKFVIFSDSP	NGOSRI	WTKSDSAA	IELEP	ETEROVEI	IRSPRAH	PINLADEL	USHEN	EFNNTSHSC	HIRVNSM	POKXAL	106													
HB-1	RPKPEDFCPPEE	NSG	VRWKKFVIFSDSP	NGOSRI	WTKSDSAA	IELEP	ETEROVEI	IRSPRAH	PINLADEL	USHEN	EFNNTSHSC	HIRVNSM	POKXAL	106													
EuroPRRSV	TCFLGOSV	VRCH	EHCHL	NA	FCYQUR	WVSH	SKYLORE	QVRE	RAV	VPD	SPIT	VEAL	CSQWIR	EL	NDVOT	REFV	TSR	IV	PNTE	TTTS	---	QI	PRFCA	HWWE	219		
Lelystad	TCFLGOSV	VRCH	EHCHL	NA	FCYQUR	WVSH	SKYLORE	QVRE	RAV	VPD	SPIT	VEAL	CSQWIR	EL	NDVOT	REFV	TSR	IV	PNTE	TTTS	---	QI	PRFCA	HWWE	219		
MN184A	SLFSSIP	LE	VEY	KE	IR	HA	IRFQYQUR	WVSH	SKYLORE	QVRE	RAV	VPD	SPIT	VEAL	CSQWIR	EL	NDVOT	REFV	TSR	IV	PNTE	TTTS	---	QI	PRFCA	HWWE	217
MN184B	SLFSSIP	LE	VEY	KE	IR	HA	IRFQYQUR	WVSH	SKYLORE	QVRE	RAV	VPD	SPIT	VEAL	CSQWIR	EL	NDVOT	REFV	TSR	IV	PNTE	TTTS	---	QI	PRFCA	HWWE	217
Ingelvac	SLFSSIP	LE	VEY	KE	IR	HA	IRFQYQUR	WVSH	SKYLORE	QVRE	RAV	VPD	SPIT	VEAL	CSQWIR	EL	NDVOT	REFV	TSR	IV	PNTE	TTTS	---	QI	PRFCA	HWWE	217
VR-2332	SLFSSIP	LE	VEY	KE	IR	HA	IRFQYQUR	WVSH	SKYLORE	QVRE	RAV	VPD	SPIT	VEAL	CSQWIR	EL	NDVOT	REFV	TSR	IV	PNTE	TTTS	---	QI	PRFCA	HWWE	217
PL97-1	SLFSSIP	LE	VEY	KE	IR	HA	IRFQYQUR	WVSH	SKYLORE	QVRE	RAV	VPD	SPIT	VEAL	CSQWIR	EL	NDVOT	REFV	TSR	IV	PNTE	TTTS	---	QI	PRFCA	HWWE	217
HN1	SLFSSIP	LE	VEY	KE	IR	HA	IRFQYQUR	WVSH	SKYLORE	QVRE	RAV	VPD	SPIT	VEAL	CSQWIR	EL	NDVOT	REFV	TSR	IV	PNTE	TTTS	---	QI	PRFCA	HWWE	217
16244B	SLFSSIP	LE	VEY	KE	IR	HA	IRFQYQUR	WVSH	SKYLORE	QVRE	RAV	VPD	SPIT	VEAL	CSQWIR	EL	NDVOT	REFV	TSR	IV	PNTE	TTTS	---	QI	PRFCA	HWWE	217
PA-8	SLFSSIP	LE	VEY	KE	IR	HA	IRFQYQUR	WVSH	SKYLORE	QVRE	RAV	VPD	SPIT	VEAL	CSQWIR	EL	NDVOT	REFV	TSR	IV	PNTE	TTTS	---	QI	PRFCA	HWWE	217
SP	SLFSSIP	LE	VEY	KE	IR	HA	IRFQYQUR	WVSH	SKYLORE	QVRE	RAV	VPD	SPIT	VEAL	CSQWIR	EL	NDVOT	REFV	TSR	IV	PNTE	TTTS	---	QI	PRFCA	HWWE	217
JA142	SLFSSIP	LE	VEY	KE	IR	HA	IRFQYQUR	WVSH	SKYLORE	QVRE	RAV	VPD	SPIT	VEAL	CSQWIR	EL	NDVOT	REFV	TSR	IV	PNTE	TTTS	---	QI	PRFCA	HWWE	217
CH-1	SLFSSIP	LE	VEY	KE	IR	HA	IRFQYQUR	WVSH	SKYLORE	QVRE	RAV	VPD	SPIT	VEAL	CSQWIR	EL	NDVOT	REFV	TSR	IV	PNTE	TTTS	---	QI	PRFCA	HWWE	217
P129	SLFSSIP	LE	VEY	KE	IR	HA	IRFQYQUR	WVSH	SKYLORE	QVRE	RAV	VPD	SPIT	VEAL	CSQWIR	EL	NDVOT	REFV	TSR	IV	PNTE	TTTS	---	QI	PRFCA	HWWE	217
HB-2	SLFSSIP	LE	VEY	KE	IR	HA	IRFQYQUR	WVSH	SKYLORE	QVRE	RAV	VPD	SPIT	VEAL	CSQWIR	EL	NDVOT	REFV	TSR	IV	PNTE	TTTS	---	QI	PRFCA	HWWE	217
HB-1	SLFSSIP	LE	VEY	KE	IR	HA	IRFQYQUR	WVSH	SKYLORE	QVRE	RAV	VPD	SPIT	VEAL	CSQWIR	EL	NDVOT	REFV	TSR	IV	PNTE	TTTS	---	QI	PRFCA	HWWE	217

