



UNIVERSITAT DE
BARCELONA



Workflow para encontrar ORFs en el extremo 5' y motivos de deslizamiento de la polimerasa en secuencias genómicas de miembros de la Familia *Potyviridae*

Giannina Stephany Bambarén Capurro

Máster Universitario en Bioinformática y Bioestadística UOC-UB

(Bioinformática y Bioestadística Área 3)

Nombre Consultor/a: Diego Garrido Martín

Nombre Profesor/a responsable de la asignatura: Ferran Prados Carrasco

Tutor CRAG: Juan José López Moya



Workflow para encontrar ORFs en el extremo 5' y motivos de deslizamiento de la polimerasa en secuencias genómicas de miembros de la Familia Potyviridae



- Introducción

- Potyvirus
- Ipomovirus
- Deslizamiento de la polimerasa

- Objetivos

- Métodos y resultados

- Búsqueda de uORFs
- “Motivos” de deslizamiento de la polimerasa
- Algoritmo gkm-SVM
- SPFMV
- Huéspedes de batata
- G. Potyvirus
- F. Potyviridae

- Conclusiones

- Bibliografía

INTRODUCCIÓN

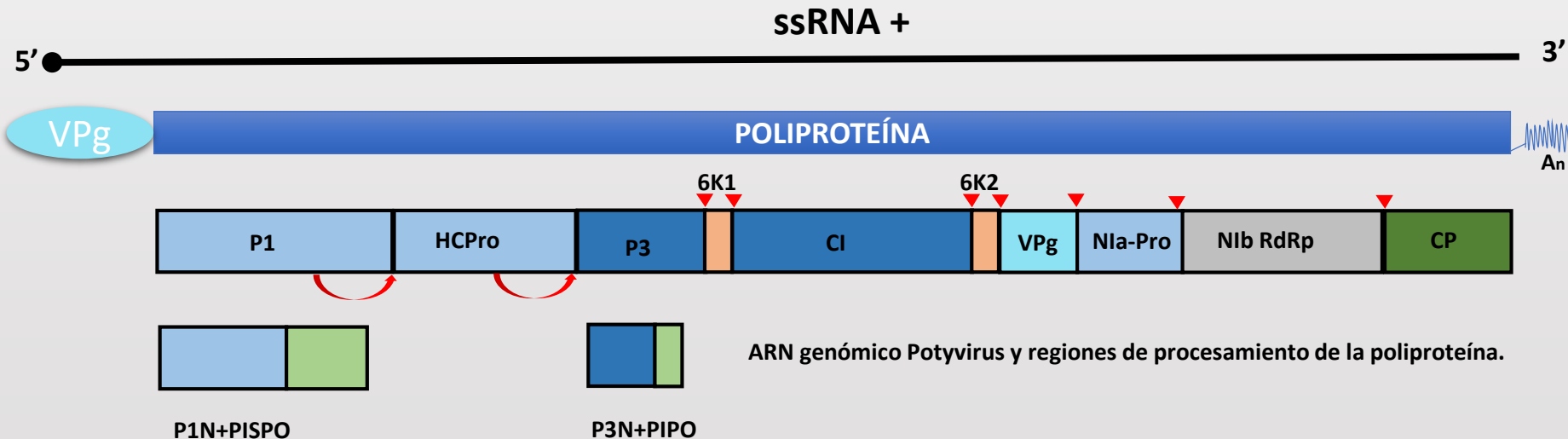
POTYVIRUS



Myzus persicae. Pulgón Verde.
Imagen obtenida: CropScience Bayer



Virión Potyvirus. Imagen obtenida: ViralZone/Potyvirus



Workflow para encontrar ORFs en el extremo 5' y motivos de deslizamiento de la polimerasa en secuencias genómicas de miembros de la Familia Potyviridae



Introducción

- Potyvirus
- **Ipomovirus**
- Deslizamiento de la polimerasa

Objetivos

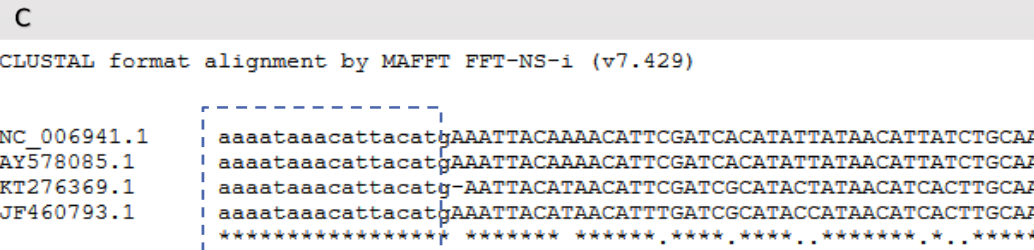
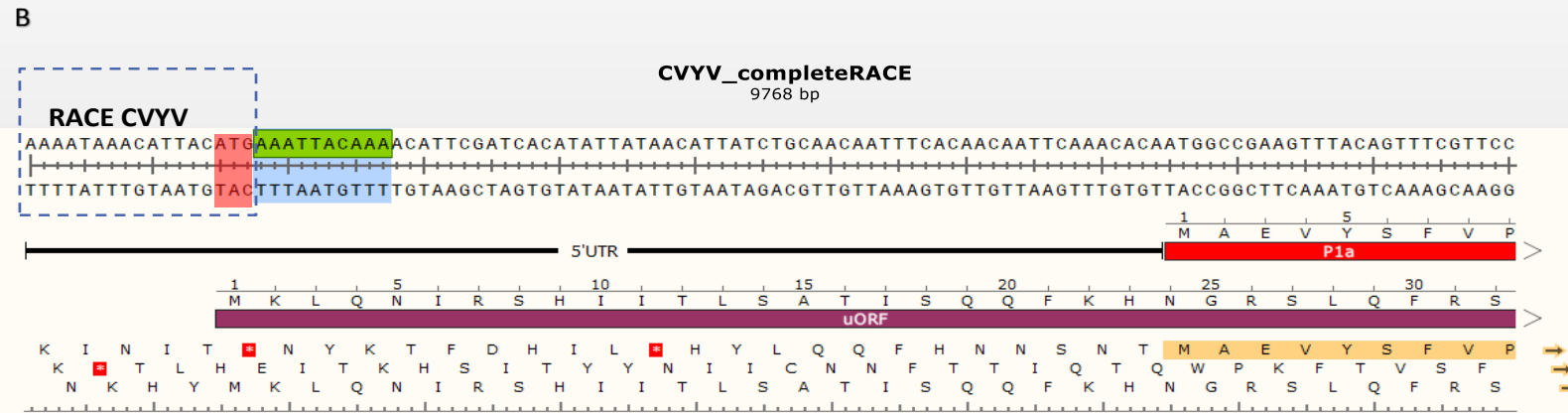
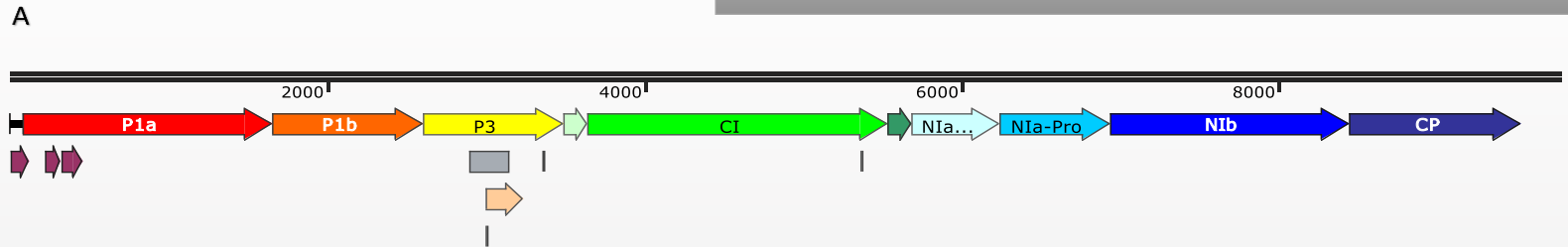
Métodos y resultados

- Búsqueda de uORFs
- "Motivos" de deslizamiento de la polimerasa
- Algoritmo gkm-SVM
- SPFMV
- Huéspedes de batata
- G. Potyvirus
- F. Potyviridae

- Conclusiones
- Bibliografía

INTRODUCCIÓN

IPOMOVIRUS



(A) Visión completa del mapa de CVYV. Imagen obtenida mediante SnapGene. (B) Secuencia de 17 nucleótidos RACE (---) la cual presenta un ATG en 5' UTR (señalado en rojo). Imagen obtenida mediante SnapGene. (C) Salida del alineamiento de todos los aislados de CVYV con RACE. Alineamiento realizado con MAFFT (disponible: <https://www.ebi.ac.uk/Tools/msa/mafft/>) con formato de salida ClustalW.

Workflow para encontrar ORFs en el extremo 5' y motivos de deslizamiento de la polimerasa en secuencias genómicas de miembros de la Familia Potyviridae



Introducción

- Potyvirus
- Ipomovirus
- **Deslizamiento de la polimerasa**

Objetivos

Métodos y resultados

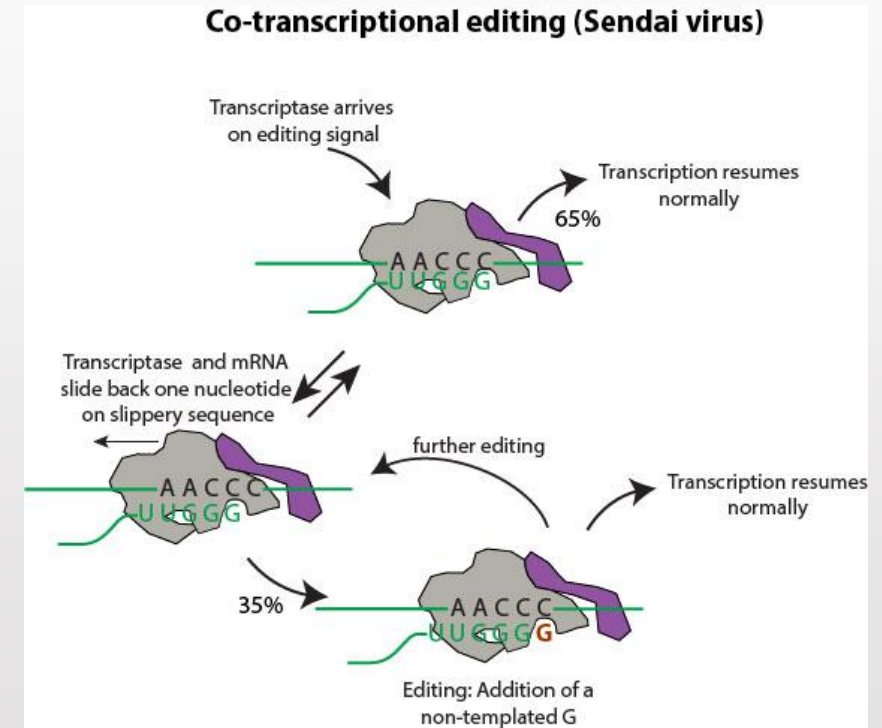
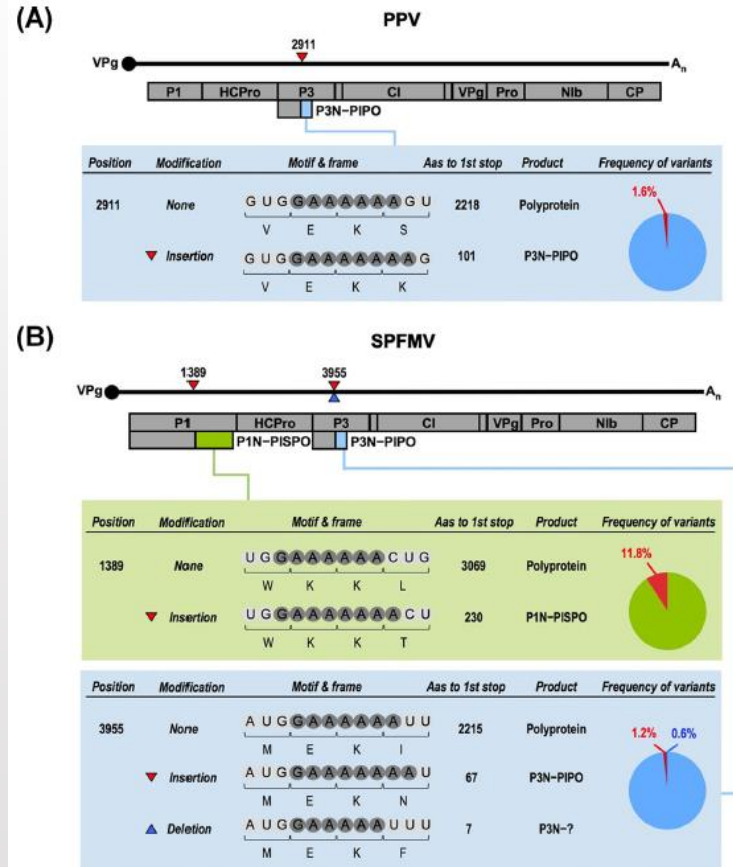
- Búsqueda de uORFs
- "Motivos" de deslizamiento de la polimerasa
- Algoritmo gkm-SVM
- SPFMV
- Huéspedes de batata
- G. Potyvirus
- F. Potyviridae

Conclusiones

Bibliografía

INTRODUCCIÓN

DESLIZAMIENTO DE LA POLIMERASA



Polymerasa slippage. ViralZone

RNA Polymerase Slippage as a Mechanism for the Production of Frameshift Gene Products in Plant Viruses of the *Potyviridae* Family (Rodamilans et al. 2015)

Workflow para encontrar ORFs en el extremo 5' y motivos de deslizamiento de la polimerasa en secuencias genómicas de miembros de la Familia Potyviridae



- **Introducción**

- Potyvirus
- Ipomovirus
- Deslizamiento de la polimerasa

- **Objetivos**

- Métodos y resultados

- Búsqueda de uORFs
- "Motivos" de deslizamiento de la polimerasa
- Algoritmo gkm-SVM
 - SPFMV
 - Huéspedes de batata
 - G. Potyvirus
 - F. Potyviridae

- Conclusiones

- Bibliografía

OBJETIVOS

1. Búsqueda de uORF

- Encontrar ORFs en regiones 5' upstream de miembros del género *Ipomovirus*.

2. Búsqueda de zonas de "slippage"

- Encontrar motivos de deslizamiento de la polimerasa en miembros de la Familia *Potyviridae*

Workflow para encontrar ORFs en el extremo 5' y motivos de deslizamiento de la polimerasa en secuencias genómicas de miembros de la Familia Potyviridae



Introducción

- Potyvirus
- Ipomovirus
- Deslizamiento de la polimerasa

Objetivos

Métodos y resultados

- **Búsqueda de uORFs**
- "Motivos" de deslizamiento de la polimerasa
- Algoritmo gkm-SVM
 - SPFMV
 - Huéspedes de batata
 - G. Potyvirus
 - F. Potyviridae

- Conclusiones

- Bibliografía

MÉTODOS Y RESULTADOS

BÚSQUEDA DE uORFS



TABLA RESUMEN DE LA BÚSQUEDA DE uORFS

Género	Especies	Nº secuencias	Nº ORF	Nº uORF
Ipomovirus	Cassava brown streak virus (CBSV)	12	624	0
Ipomovirus	Cucumber vein yellowing virus (CVYV)	4	274	3
Ipomovirus	Squash vein yellowing virus (SqVYV)	5	269	5
Potyvirus	Plum pox virus (PPV)	423	21227	22
Potyvirus	Tobacco etch virus (TEV)	9	456	0
Ipomovirus	Tomato mild mottle virus (TMMV)	1	60	0

Tabla de uORFs para CVYV.

ID	Name	ORF	start	stop	sequence
KT276369	ORF64_KT276369	ORF64	14	46	MNYTDR
AY578085	ORF56_AY578085	ORF56	14	124	MKNRSHTSATSKHNGRSRNR

Tabla de uORFs para SqVYV.

ID	Name	ORF	start	stop	sequence
KT721735	ORF16_KT721735	ORF16	14	106	MKHNDKSANCYVHTTS
KU953950	ORF54_KU953950	ORF54	14	70	MNHNDSHYTKS
JF897996	ORF15_JF897996	ORF15	14	82	MNHNDKSNHCHSY
EU259611	ORF15_EU259611	ORF15	14	82	MNHNDKSNHCHSY

Caracterización de productos génicos en el tercio N-terminal del genoma del Potyvirus *Sweet Potato Virus 2* de *Ipomea batata*



- Introducción

- Potyvirus
- Ipomovirus
- Deslizamiento de la polimerasa

- Objetivos

- Métodos y resultados

- Búsqueda de uORFs

- "Motivos" de deslizamiento de la polimerasa
- Algoritmo gkm-SVM
 - SPFMV
 - Huéspedes de batata
 - G. Potyvirus
 - F. Potyviridae

- Conclusiones

- Bibliografía

MÉTODOS Y RESULTADOS

RESULTADOS DE LOS uORFS

A.

```
MUSCLE 3.8.31

Call:
  msa(alg_cvyy, "Muscle")

MsaAAMultipleAlignment with 3 rows and 21
  aln          names
[1] MKNRSHTSATSKHNGRSRSNR ORF56_NC_006941
[2] MKNRSHTSATSKHNGRSRSNR ORF56_AY578085
[3] -----MNYTDR ORF64_KT276369
Con MKNRSHTSATSKHNGRSRSNR Consensus
```

B.

```
ClustalOmega 1.2.0

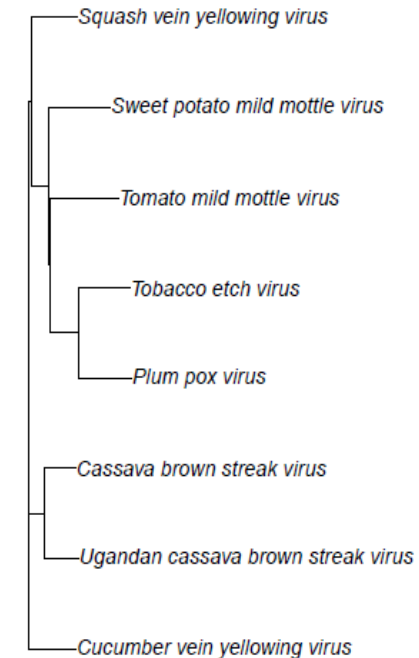
Call:
  msa(alg_svyv, "ClustalOmega")

MsaAAMultipleAlignment with 5 rows and 16
  aln          names
[1] MNHNDSHYTKS----- ORF54_KU953950
[2] MNHNDKSNHCHSY--- ORF15_NC_010521
[3] MNHNDKSNHCHSY--- ORF15_JF897996
[4] MNHNDKSNHCHSY--- ORF15_EU259611
[5] MKHNDKSANCYVHTTS ORF16_KT721735
Con MNHNDKSNHCHSY--- Consensus
```

Salida de R de alineamientos realizados con "msa".

(A) Alineamiento para CVYV **(B)** Alineamiento para SqVYV,

Phylogenetic Tree of Potyviridae



Árbol filogenético de las especies de Ipomovirus y dos de Potyvirus.

Árbol realizado mediante un alineamiento de las secuencias con ClustalOmega (librería Biconductor "msa") y resuelto mediante Neighbor-joining (librería de R "ape"(38)).

Workflow para encontrar ORFs en el extremo 5' y motivos de deslizamiento de la polimerasa en secuencias genómicas de miembros de la Familia Potyviridae



Introducción

- Potyvirus
- Ipomovirus
- Deslizamiento de la polimerasa

Objetivos

Métodos y resultados

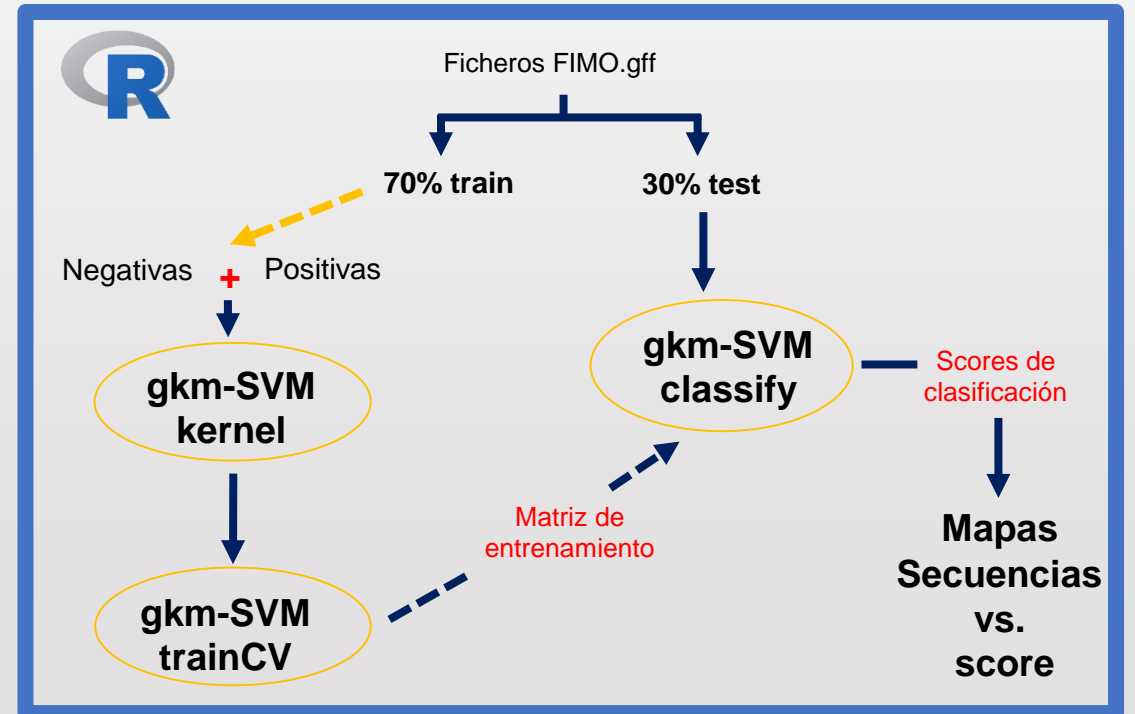
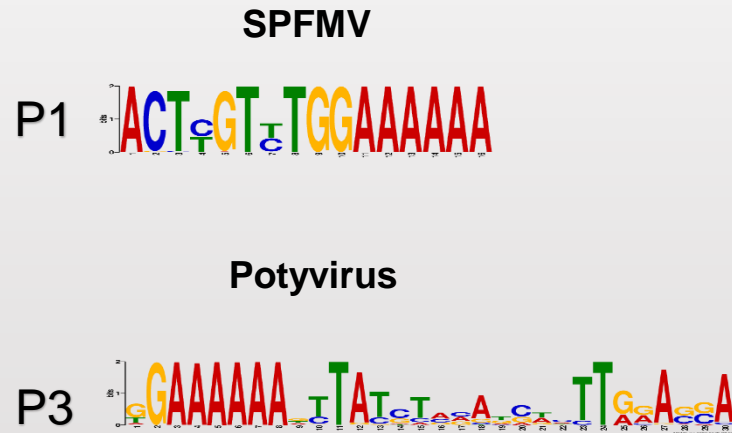
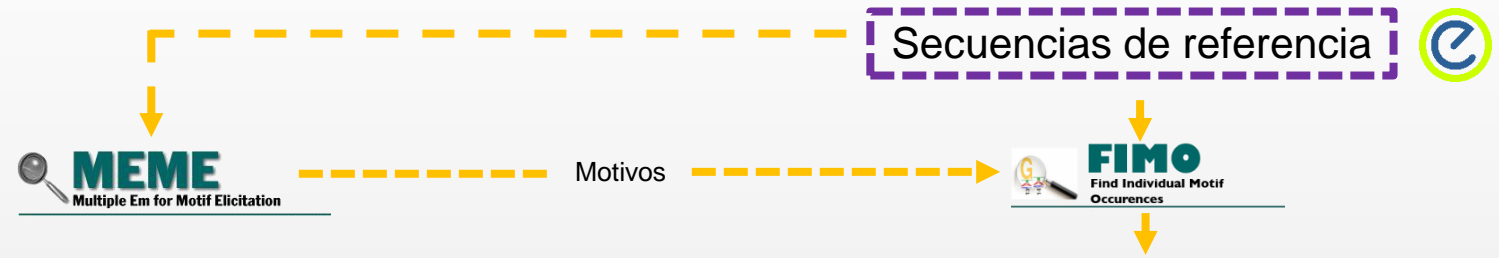
- Búsqueda de uORFs
- “Motivos” de deslizamiento de la polimerasa
- Algoritmo gkm-SVM
 - SPFMV
 - Huéspedes de batata
 - G. Potyvirus
 - F. Potyviridae

Conclusiones

Bibliografía

MÉTODOS Y RESULTADOS

PIPELINE



Workflow para encontrar ORFs en el extremo 5' y motivos de deslizamiento de la polimerasa en secuencias genómicas de miembros de la Familia Potyviridae



- Introducción

- Potyvirus
- Ipomovirus
- Deslizamiento de la polimerasa

- Objetivos

- Métodos y resultados

- Búsqueda de uORFs
- “Motivos” de deslizamiento de la polimerasa
- Algoritmo gkm-SVM
 - SPFMV
 - Huéspedes de batata
 - G. Potyvirus
 - F. Potyviridae

- Conclusiones

- Bibliografía

MÉTODOS Y RESULTADOS

RESULTADOS FIMO

Resumen de FIMO para el motivo en P1.

Motivo	Secuencias utilizadas	Ocurrencias	Patrón real
		FIMO	G2A6 (P1)
P1	Cepas SPFMV (55)	130	55
P1	Todos los Potyvirus huéspedes de batata (112)	285	81
P1	Todas las cepas del g. Potyvirus (2351)	5951	356

Resumen de FIMO para el motivo en P3.

Motivo	Secuencias utilizadas	Ocurrencias	Patrón real
		FIMO	GA6 (P3)
P3	Todos los Potyvirus huéspedes de batata (112)	624	238
P3	Todas las cepas del g. Potyvirus (2351)	11220	2452
P3	Todas las cepas de la fam. Potyviridae (2739)	12504	2634

Workflow para encontrar ORFs en el extremo 5' y motivos de deslizamiento de la polimerasa en secuencias genómicas de miembros de la Familia Potyviridae



Introducción

- Potyvirus
- Ipomovirus
- Deslizamiento de la polimerasa

Objetivos

Métodos y resultados

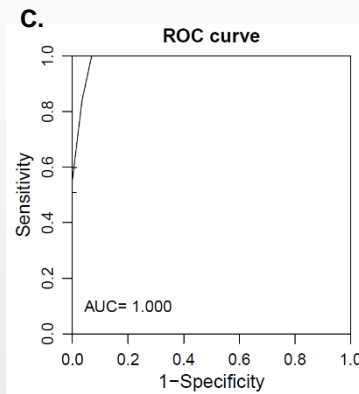
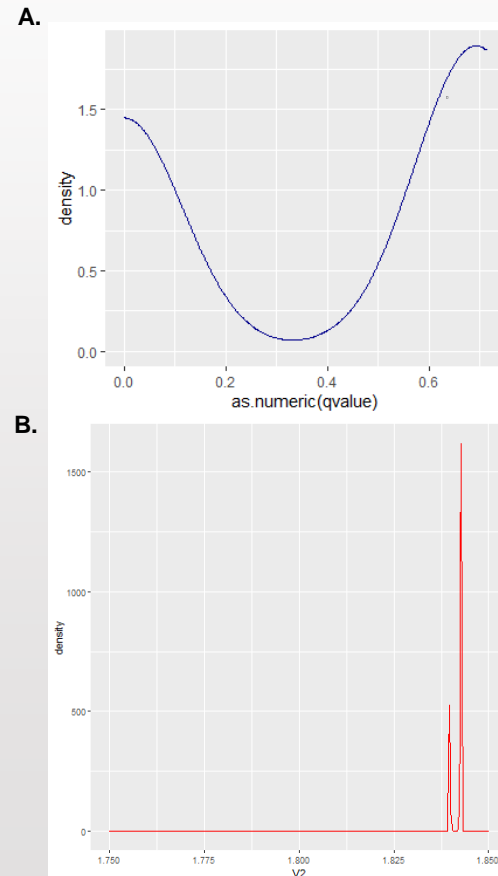
- Búsqueda de uORFs
- "Motivos" de deslizamiento de la polimerasa
- **Algoritmo gkm-SVM**
 - **SPFMV**
 - Huéspedes de batata
 - G. Potyvirus
 - F. Potyviridae

- Conclusiones

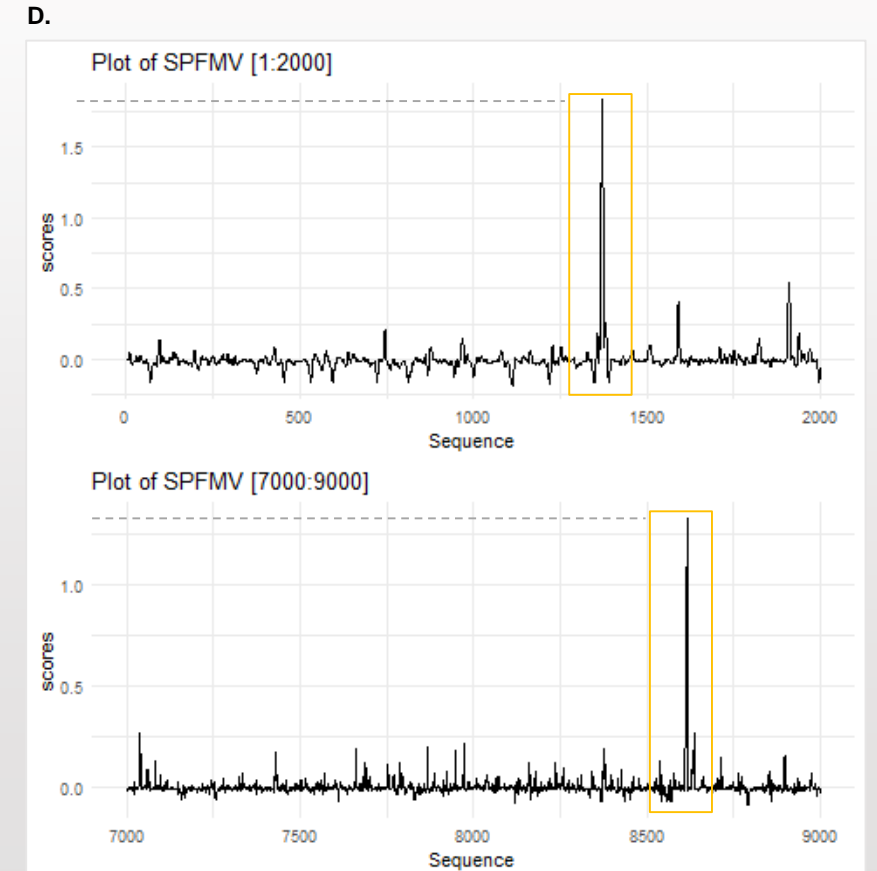
- Bibliografía

MÉTODOS Y RESULTADOS

SECUENCIAS DE CEPAS DE SPFMV



(A) Distribución de los q-values para las ocurrencias encontradas por FIMO para el motivo de P1 en las secuencias de SPFMV, gráfica realizada con "ggplot2". **(B) Gráfica de la distribución de los puntajes** de las secuencias positivas evaluadas mediante el algoritmo de clasificación "gkm-SVM", realizada con "ggplot2". **(C) Curvas ROC.** **(D) Secuencias en mosaico.**



Workflow para encontrar ORFs en el extremo 5' y motivos de deslizamiento de la polimerasa en secuencias genómicas de miembros de la Familia Potyviridae



Introducción

- Potyvirus
- Ipomovirus
- Deslizamiento de la polimerasa

Objetivos

Métodos y resultados

- Búsqueda de uORFs
- "Motivos" de deslizamiento de la polimerasa

Algoritmo gkm-SVM

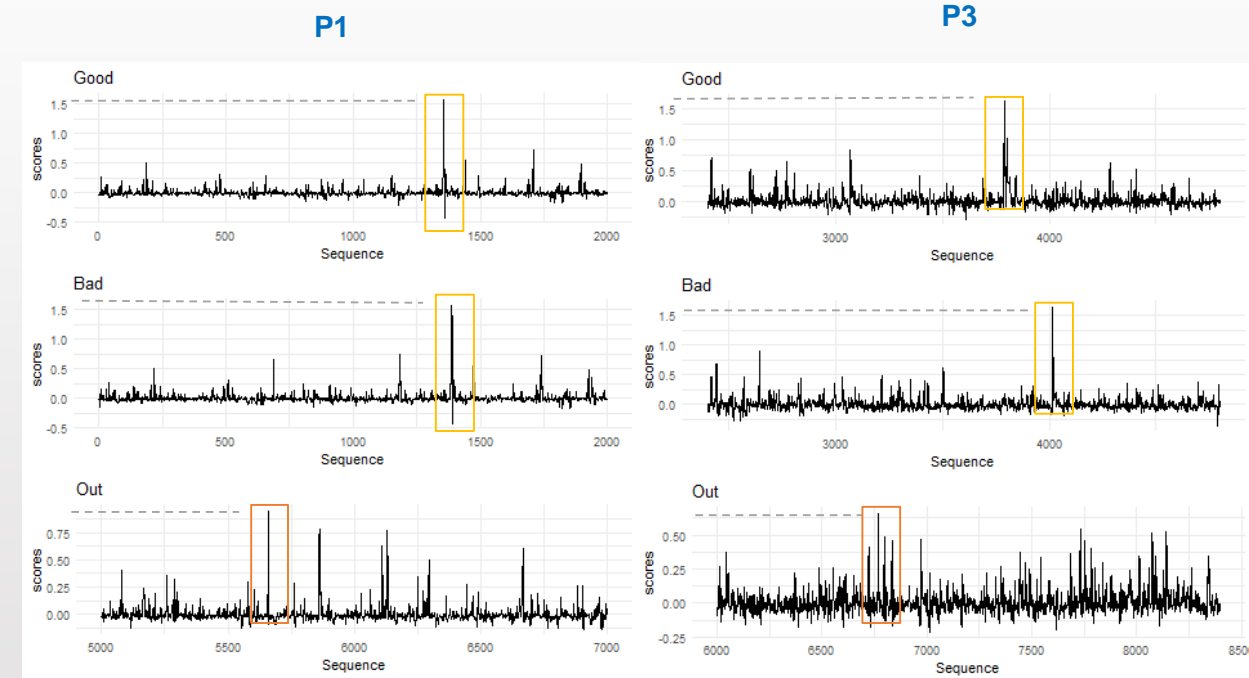
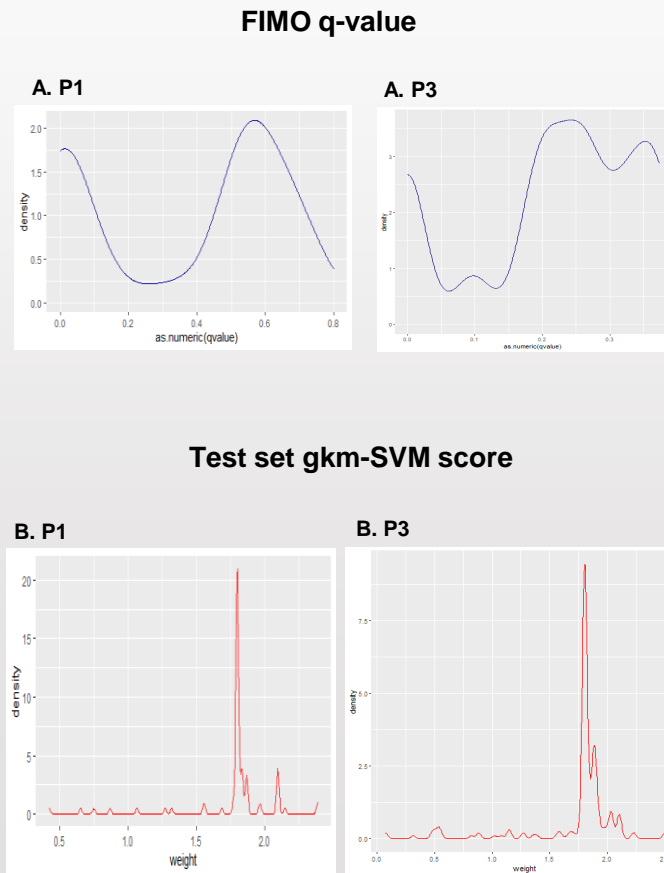
- SPFMV
- **Huéspedes de batata**
- G. Potyvirus
- F. Potyviridae

Conclusiones

Bibliografía

MÉTODOS Y RESULTADOS

SECUENCIAS DE HUÉSPEDES DE BATATA



(A) Diagrama de distribución de q-value de las ocurrencias obtenidas por FIMO de los motivos en secuencia de virus huéspedes de batata. (B) Diagrama de distribución de los puntajes obtenidos del algoritmo de clasificación gkm-SVM para la ocurrencias encontradas por FIMO en secuencias de virus huéspedes de batata. Ambos gráficos se realizaron en "ggplot". (C) Secuencias en mosaico calificadas con gkm-SVM entrenado con secuencias de virus huéspedes de batata. El mapa muestra un cambio de nucleótidos a partir de secuencias de 10-mers, los picos de máximo puntaje corresponden a zonas donde se encuentra el motivo

Workflow para encontrar ORFs en el extremo 5' y motivos de deslizamiento de la polimerasa en secuencias genómicas de miembros de la Familia Potyviridae



Introducción

- Potyvirus
- Ipomovirus
- Deslizamiento de la polimerasa

Objetivos

Métodos y resultados

- Búsqueda de uORFs
- "Motivos" de deslizamiento de la polimerasa

Algoritmo gkm-SVM

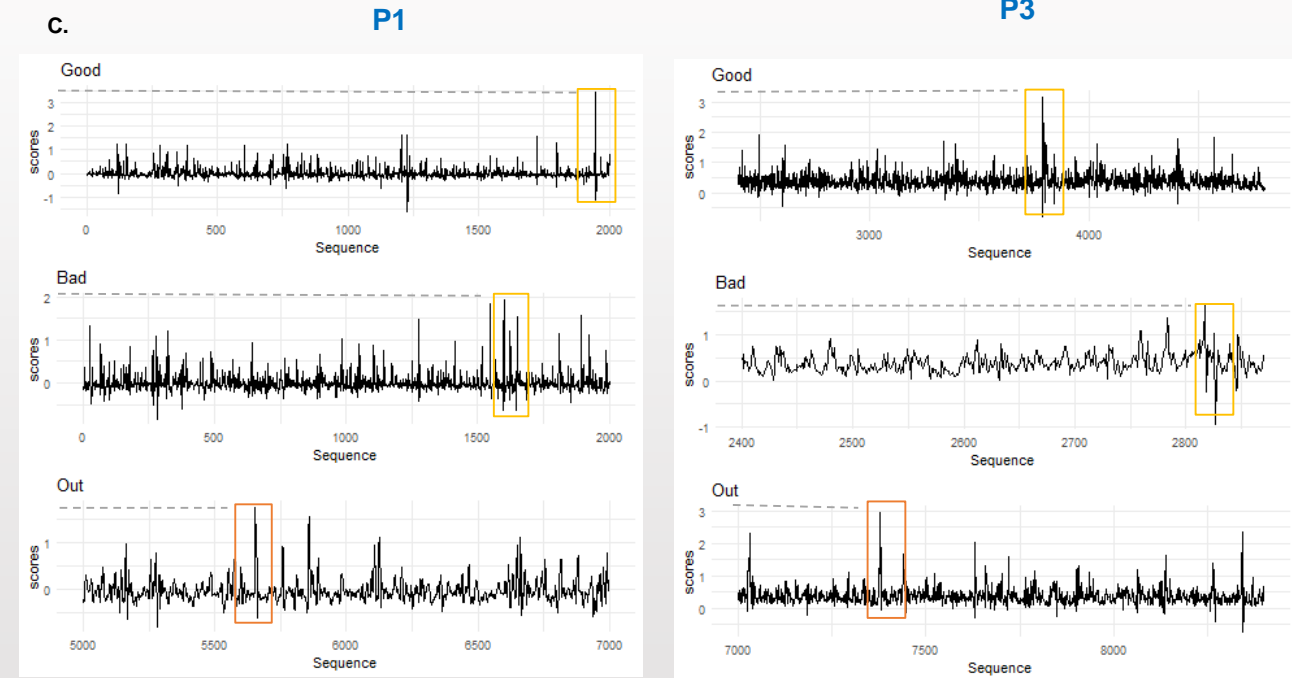
- SPFMV
- Huéspedes de batata
- **G. Potyvirus**
- F. Potyviridae

Conclusiones

Bibliografía

MÉTODOS Y RESULTADOS

SECUENCIAS DEL GENERO POTYVIRUS



(A) Diagrama de distribución de q-value de las ocurrencias obtenidas por FIMO de los motivos en secuencia de virus del género potyvirus. (B) Diagrama de distribución de los puntajes obtenidos del algoritmo de clasificación gkm-SVM para la ocurrencias encontradas por FIMO en secuencias del de potyvirus. (C) Secuencias en mosaico calificadas con gkm-SVM entrenado con secuencias de virus del género potyvirus. El mapa muestra un cambio de nucleótidos a partir de secuencias de 10-mers, los picos de máximo puntaje corresponden a zonas donde se encuentra el motivo.

Workflow para encontrar ORFs en el extremo 5' y motivos de deslizamiento de la polimerasa en secuencias genómicas de miembros de la Familia Potyviridae



- Introducción

- Potyvirus
- Ipomovirus
- Deslizamiento de la polimerasa

- Objetivos

Métodos y resultados

- Búsqueda de uORFs
- "Motivos" de deslizamiento de la polimerasa

- Algoritmo gkm-SVM

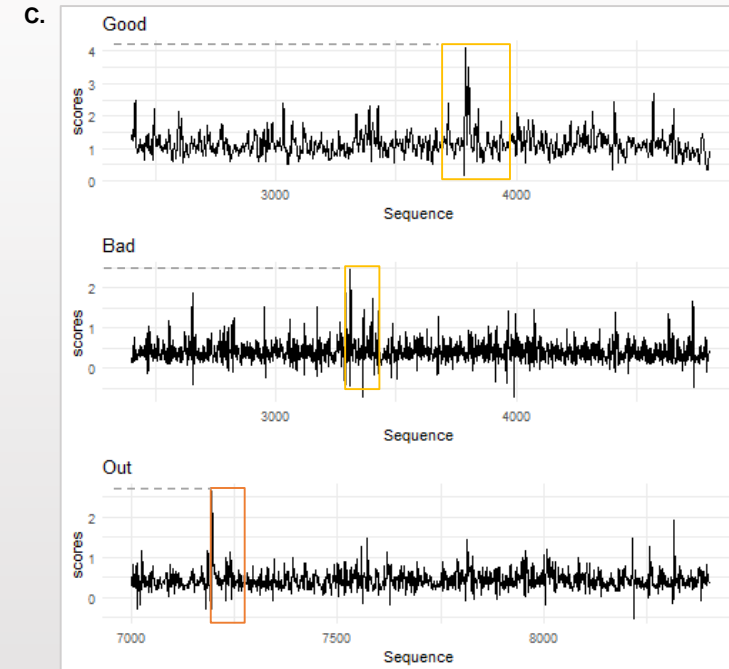
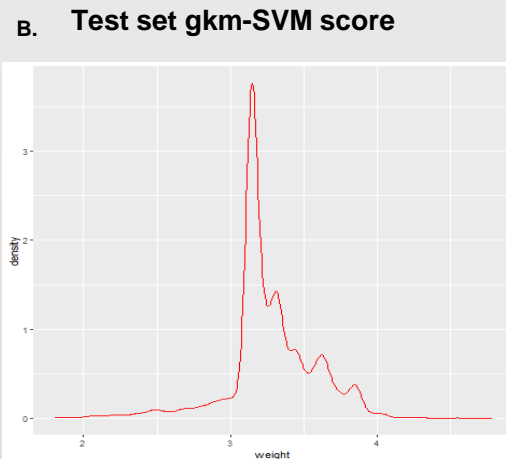
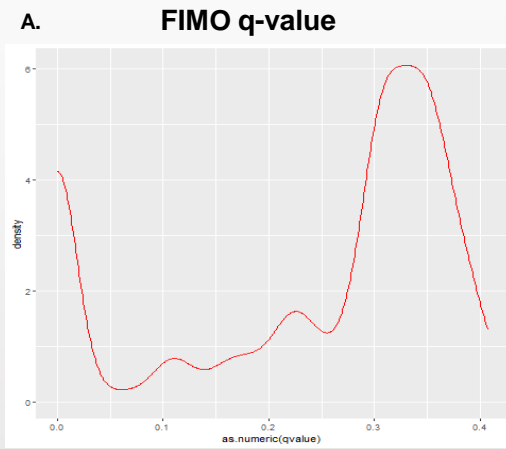
- SPFMV
- Huéspedes de batata
- G. Potyvirus
- **F. Potyviridae**

- Conclusiones

- Bibliografía

MÉTODOS Y RESULTADOS

SECUENCIAS DE LA FAMILIA POTYVIRIDAE



(A) Diagrama de distribución de q-value de las ocurrencias obtenidas por FIMO de los motivos en secuencia de virus de la familia Potyviridae. (B) Diagrama de distribución de los puntajes obtenidos del algoritmo de clasificación gkm-SVM para la ocurrencias encontradas por FIMO en secuencias de virus de la familia Potyviridae. (C) Secuencias en mosaico calificadas con gkm-SVM entrenado con secuencias de virus de la familia Potyviridae. El mapa muestra un cambio de nucleótidos a partir de secuencias de 10-mers, los picos de máximo puntaje corresponden a zonas donde se encuentra el motivo.



- Introducción

- Potyvirus
- Ipomovirus
- Deslizamiento de la polimerasa

- Objetivos

- Métodos y resultados

- Búsqueda de uORFs
- "Motivos" de deslizamiento de la polimerasa
- Algoritmo gkm-SVM
 - SPFMV
 - Huéspedes de batata
 - G. Potyvirus
 - F. Potyviridae

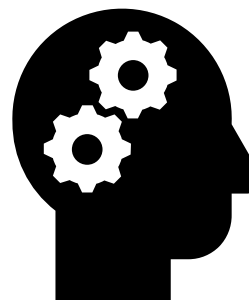
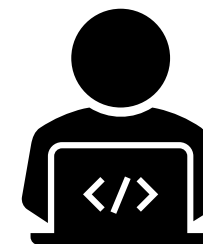
- Conclusiones

- Bibliografía

CONCLUSIONES

- No hemos encontrado indicios claros que nos hagan pensar que CVYV presenta una pauta de lectura nueva en su región 5'.
- **Las herramientas de descubrimiento y búsqueda de motivos son herramienta útiles y gráficas, pero es vital someterlas a un filtrado de sus resultados ya que nos generan muchos falsos positivos.**
- La cantidad de **falsos positivos** es aún mayor cuando el motivo usado presenta muchas posiciones con una baja frecuencia de nucleótidos, como es el caso del motivo de deslizamiento en P3.
- Encontramos que el **algoritmo gkm-SVM** es una potente herramienta de clasificación de motivos de secuencia. Mediante este algoritmo vemos que, a más cantidad de secuencias utilizadas en el entrenamiento, mejor capacidad de discriminación, pero también es más elevado el tiempo de creación del kernel y más pesado el archivo generado.
- El análisis realizado nos ha ayudado a observar que estos virus presentan muchos **zonas potenciales que podrían generar mecanismos de slippage**, ya que como en el caso de P1N-PISPO, estos mecanismos son respuesta al sistema de defensa de las planta.
- Encontramos necesario el desarrollo de herramientas de análisis para virus de plantas.

MUCHAS GRACIAS
POR SU ATENCIÓN



Workflow para encontrar ORFs en el extremo 5' y motivos de deslizamiento de la polimerasa en secuencias genómicas de miembros de la Familia Potyviridae



- **Introducción**

- Potyvirus
- Ipomovirus
- Deslizamiento de la polimerasa

- **Objetivos**

- **Métodos y resultados**

- Búsqueda de uORFs
- "Motivos" de deslizamiento de la polimerasa
- Algoritmo gkm-SVM
 - SPFMV
 - Huéspedes de batata
 - G. Potyvirus
 - F. Potyviridae

- **Conclusiones**

- **Bibliografía**

BIBLIOGRAFÍA

1. Fauquet C, Mayo MA, Maniloff J, Desselberger U, Ball LA. Virus taxonomy - eighth report of the International Committee on the taxonomy of viruses. Vol. 83, The Viruses. 2005. 988–992 p.
2. Stephen J Wylie, Alice Kazuko Inoue-Nagata, Jan Kreuze, Juan José López-Moya, Kristiina Mäkinen KO and AW. Potyviridae [Internet]. Virus Taxonomy: 2019 Release. 2019 [cited 2020 Jun 15]. Available from: https://talk.ictvonline.org/ictv-reports/ictv_online_report/positive-sense-rna-viruses/w/potyviridae
3. Wylie SJ, Adams M, Chalam C, Kreuze J, López-Moya JJ, Ohshima K, et al. ICTV virus taxonomy profile: Potyviridae. J Gen Virol. 2017;98(3):352–4.
4. Adams MJ, Antoniw JF, Fauquet CM. Molecular criteria for genus and species discrimination within the family Potyviridae. Arch Virol [Internet]. 2005;150(3):459–79. Available from: <https://doi.org/10.1007/s00705-004-0440-6>
5. Scholthof KBG, Adkins S, Czosnek H, Palukaitis P, Jacquot E, Hohn T, et al. Top 10 plant viruses in molecular plant pathology. Mol Plant Pathol. 2011;12(9):938–54.
6. Chung BY-W, Miller WA, Atkins JF, Firth AE. An overlapping essential gene in the Potyviridae. Proc Natl Acad Sci. 2008 Apr 15;105(15):5897–902.
7. Mingot A, Valli A, Rodamilans B, San León D, Baulcombe DC, García JA, et al. The P1N-PISPO trans -Frame Gene of Sweet Potato Feathery Mottle Potyvirus Is Produced during Virus Infection and Functions as an RNA Silencing Suppressor . J Virol. 2016;90(7):3543–57.
8. Olsper A, Chung BY, Atkins JF, Carr JP, Firth AE. Transcriptional slippage in the positive-sense RNA virus family Potyviridae . EMBO Rep. 2015;16(8):995–1004.
9. Janssen D, Martín G, Velasco L, Gómez P, Segundo E, Ruiz L, et al. Absence of a coding region for the helper component-proteinase in the genome of cucumber vein yellowing virus, a whitefly-transmitted member of the Potyviridae. Arch Virol. 2005;150(7):1439–47.
10. Valli A, López-Moya JJ, García JA. Recombination and gene duplication in the evolutionary diversification of P1 proteins in the family Potyviridae. J Gen Virol. 2007;88(3):1016–28.
11. Li W, Hilf ME, Webb SE, Baker CA, Adkins S. Presence of P1b and absence of HC-Pro in Squash vein yellowing virus suggests a general feature of the genus Ipomovirus in the family Potyviridae. Virus Res. 2008 Aug;135(2):213–9.
12. Valli A, Martin-Hernandez AM, Lopez-Moya JJ, Garcia JA. RNA Silencing Suppression by a Second Copy of the P1 Serine Protease of Cucumber Vein Yellowing Ipomovirus, a Member of the Family Potyviridae That Lacks the Cysteine Protease HCPro. J Virol. 2006;80(20):10055–63.
13. Frohman MA, Dush MK, Martin GR. Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. Proc Natl Acad Sci U S A. 1988 Dec;85(23):8998–9002.

Workflow para encontrar ORFs en el extremo 5' y motivos de deslizamiento de la polimerasa en secuencias genómicas de miembros de la Familia Potyviridae



- Introducción

- Potyvirus
- Ipomovirus
- Deslizamiento de la polimerasa

- Objetivos

- Métodos y resultados

- Búsqueda de uORFs
- "Motivos" de deslizamiento de la polimerasa
- Algoritmo gkm-SVM
 - SPFMV
 - Huéspedes de batata
 - G. Potyvirus
 - F. Potyviridae

- Conclusiones

- Bibliografía

BIBLIOGRAFÍA

14. Untiveros M, Olsper A, Artola K, Firth AE, Kreuze JF, Valkonen JPT. A novel sweet potato potyvirus open reading frame (ORF) is expressed via polymerase slippage and suppresses RNA silencing. *Mol Plant Pathol*. 2016;17(7):1111–23.
15. Nigam D, LaTourrette K, Souza PFN, Garcia-Ruiz H. Genome-Wide Variation in Potyviruses. *Front Plant Sci*. 2019;10(November):1–28.
16. Bedhomme S, Lafforgue G, Elena SF. Multihost experimental evolution of a plant RNA virus reveals local adaptation and host-specific mutations. *Mol Biol Evol*. 2012;29(5):1481–92.
17. García-Arenal F, Fraile A, Malpica JM. Variability and Genetic Structure of Potato Virus Y Populations. *Annu Rev Phytopathol*. 2001;39(1):157–86.
18. Huang L, Li Z, Wu J, Xu Y, Yang X, Fan L, et al. Analysis of genetic variation and diversity of rice stripe virus populations through high-throughput sequencing. *Front Plant Sci*. 2015;6(MAR):1–7.
19. Athey J, Alexaki A, Osipova E, Rostovtsev A, Santana-Quintero L V., Katneni U, et al. A new and updated resource for codon usage tables. *BMC Bioinformatics*. 2017;18(1):1–10.
20. Bodenhofer U, Bonatesta E, Horejš-Kainrath C, Hochreiter S. msa: an R package for multiple sequence alignment. *Bioinformatics* [Internet]. 2015 Aug 26;31(24):3997–9. Available from: <https://doi.org/10.1093/bioinformatics/btv494>
21. Kans J. Entrez Direct: E-utilities on the UNIX Command Line. In: Entrez Programming Utilities Help. Entrez Program Util Help [Internet]. 2013;(Md):1–86. Available from: <http://www.ncbi.nlm.nih.gov/books/NBK179288/>
22. Zhang H, Wang Y, Lu J. Function and Evolution of Upstream ORFs in Eukaryotes. *Trends Biochem Sci* [Internet]. 2019;44(9):782–94. Available from: <http://www.sciencedirect.com/science/article/pii/S0968000419300556>
23. Carrington JC, Freed DD. Cap-independent enhancement of translation by a plant potyvirus 5' nontranslated region. *J Virol*. 1990 Apr;64(4):1590–7.
24. Roberts R, Zhang J, Mayberry LK, Tatineni S, Browning KS, Rakotondrafara AM. A Unique 5' Translation Element Discovered in Triticum Mosaic Virus. *J Virol*. 2015 Dec;89(24):12427–40.
25. Pagès H, Aboyou P GR. Biostrings: Efficient manipulation of biological strings. R package version 2.56.0. DebRoy S. 2020.
26. D. Charif and J.R. Lobry. Seqin R 1.0-2: a contributed package to the R project for statistical computing devoted to biological sequences retrieval and analysis. Springer Verlag. 2007.

Workflow para encontrar ORFs en el extremo 5' y motivos de deslizamiento de la polimerasa en secuencias genómicas de miembros de la Familia Potyviridae



- Introducción

- Potyvirus
- Ipomovirus
- Deslizamiento de la polimerasa

- Objetivos

- Métodos y resultados

- Búsqueda de uORFs
- "Motivos" de deslizamiento de la polimerasa
- Algoritmo gkm-SVM
 - SPFMV
 - Huéspedes de batata
 - G. Potyvirus
 - F. Potyviridae

- Conclusiones

- Bibliografía

BIBLIOGRAFÍA

27. Bailey TL, Elkan C. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proceedings Int Conf Intell Syst Mol Biol.* 1994;2:28–36.
28. Grant CE, Bailey TL, Noble WS. FIMO: scanning for occurrences of a given motif. *Bioinformatics [Internet].* 2011 Feb 16;27(7):1017–8. Available from: <https://doi.org/10.1093/bioinformatics/btr064>
29. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B.* 1995;57(1):289–300.
30. Yousef M, Khalifa W, Acar İE, Allmer J. MicroRNA categorization using sequence motifs and k-mers. *BMC Bioinformatics [Internet].* 2017;18(1):170. Available from: <https://doi.org/10.1186/s12859-017-1584-1>
31. Ghandi M, Lee D, Mohammad-Noori M, Beer MA. Enhanced Regulatory Sequence Prediction Using Gapped k-mer Features. *PLOS Comput Biol [Internet].* 2014 Jul 17;10(7):e1003711. Available from: <https://doi.org/10.1371/journal.pcbi.1003711>
32. Lantz B. *Machine Learning with R [Internet].* Packt Publishing; 2013. (Community experience distilled). Available from: <https://books.google.es/books?id=ZQu8AQAAQBAJ>
33. Ghandi M, Mohammad-Noori M, Ghareghani N, Lee D, Garraway L, Beer MA. gkmSVM: an R package for gapped-kmer SVM. *Bioinformatics [Internet].* 2016/04/19. 2016 Jul 15;32(14):2205–7. Available from: <https://pubmed.ncbi.nlm.nih.gov/27153639>
34. Lawrence M, Gentleman R, Carey V. rtracklayer: an R package for interfacing with genome browsers. *Bioinformatics [Internet].* 2009 May 25;25(14):1841–2. Available from: <https://doi.org/10.1093/bioinformatics/btp328>
35. Proutski V, Holmes E. SWAN: sliding window analysis of nucleotide sequence variability. *Bioinformatics.* 1998 Jun;14(5):467–8.
36. Janssen D, Martín G, Velasco L, Gómez P, Segundo E, Ruiz L, et al. Absence of a coding region for the helper component-proteinase in the genome of cucumber vein yellowing virus, a whitefly-transmitted member of the Potyviridae. *Arch Virol.* 2005 Jul;150(7):1439–47.
37. Gibbs AJ, Hajizadeh M, Ohshima K, Jones RAC. The Potyviruses: An Evolutionary Synthesis Is Emerging. *Viruses.* 2020 Jan;12(2).
38. Paradis E, Claude J, Strimmer K. APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics [Internet].* 2004 Jan 22;20(2):289–90. Available from: <https://doi.org/10.1093/bioinformatics/btg412>