

Informe de Clasificación Taxonómica, Variantes, Anotación y Variantes de Interés Grupo 2

Santiago Berty Muñoz
Tomás González Franco
Sofía Cabezas Cruz
Escuela de Ingenieros de Antioquia
Pregrado en Ingeniería Biotecnológica

24 de octubre de 2025

1. Introducción

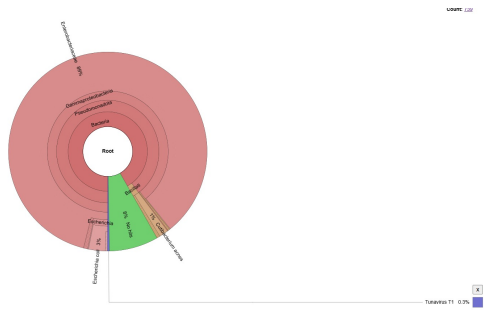
En esta segunda entrega se abordó el análisis de las líneas evolucionadas **EVOL1** y **EVOL2** provenientes de un experimento de evolución experimental con *Escherichia coli*. El objetivo fue identificar los cambios genómicos y taxonómicos surgidos tras varias generaciones bajo condiciones de presión selectiva, utilizando herramientas especializadas para la clasificación y anotación de variantes.

Se realizaron tres etapas principales: 1. Clasificación taxonómica de lecturas no mapeadas mediante **Kraken2** y visualización en **KronaTools**, 2. Llamado y filtrado de variantes con **bcftools** y 3. Anotación funcional de mutaciones mediante **SnpEff**. Los resultados nos permiten evaluar cómo la composición microbiana y las variantes genéticas reflejan los procesos evolutivos ocurridos.

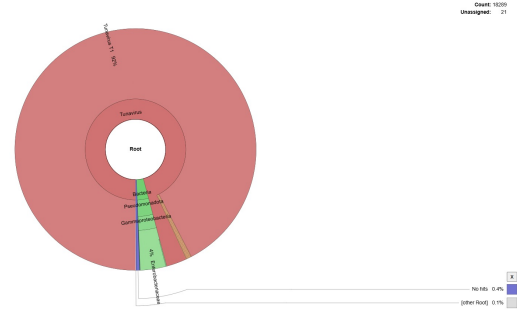
2. Clasificación taxonómica

Comparación global entre EVOL1 y EVOL2

El análisis de las lecturas no mapeadas (unmapped) mostró una diferencia muy marcada en la composición microbiana de ambas líneas. En **EVOL1**, se logró clasificar el 92.02 % de las lecturas, predominando las bacterias (91.75 %) sobre una presencia viral mínima (0.27 %). Por el contrario, en **EVOL2** se observó un cambio bastante drástico: 99.57 % de lecturas clasificadas, donde los virus representaron el 95.83 % y las bacterias tan solo el 3.63 %. Estos resultados evidencian un proceso de amplificación viral masiva (aprox. 8750× respecto a EVOL1) (Wood et al., 2019).



(a) Composición taxonómica de EVOL1.



(b) Composición taxonómica de EVOL2.

Figura 1: Visualización de la clasificación taxonómica mediante **KronaTools** (Ondov et al., 2011).

Componente bacteriano

En **EVOL1**, las lecturas bacterianas estuvieron dominadas por la familia *Enterobacteriaceae* (89.17%), con fuerte presencia de *Escherichia coli*, acompañado con pequeñas proporciones de géneros como *Staphylococcus* y *Cutibacterium acnes*. Este patrón nos dice que hay una población bacteriana estable con contaminación ambiental mínima.

Por otro lado, en **EVOL2**, la proporción bacteriana se redujo drásticamente hasta un 3.63%, aunque *E. coli* continuó siendo la especie predominante. Aparecieron además especies estrechamente relacionadas (*Shigella*, *Citrobacter*, *Salmonella*), reflejando regiones genómicas conservadas y posibles efectos de secuencias profágicas. La notable reducción en contaminantes típicos de la piel (de 2.03% a 0.05%) podría deberse a un manejo del experimento más limpio o una eliminación diferencial durante la evolución.

Componente viral e interpretación biológica

El cambio más significativo ocurrió en el componente viral. En la línea **EVOL1**, apenas se detectaron 2 lecturas virales correspondientes al *Escherichia virus T1*. En cambio, **EVOL2** mostró una presencia viral masiva, con un total de 17,527 lecturas virales, principalmente *Escherichia virus T1* (92.35%), además de otros fagos relacionados como *Escherichia virus ADB2* y *Shigella virus PSf2* (Ackermann, 2009). Todos estos virus pertenecen a la familia *Siphoviridae* (subfamilia *Tunavirinae*), lo que nos indica una expansión clonal de un mismo linaje fágico (De Paepe & Taddei, 2006).

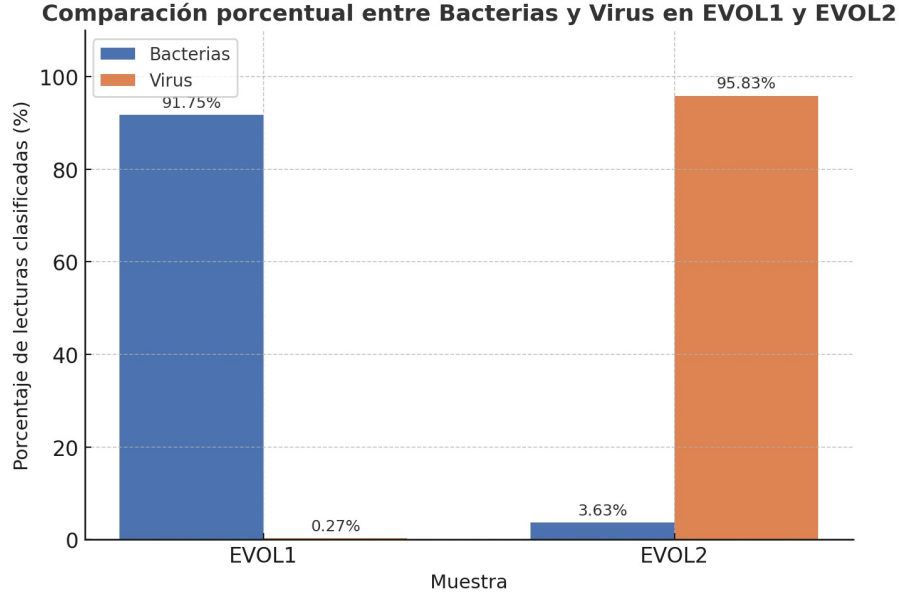


Figura 2: Comparación porcentual entre lecturas bacterianas y virales en EVOL1 y EVOL2.

Este patrón nos sugiere que hubo una infección viral activa e intensa en EVOL2, probablemente de tipo lítico, en donde el fago T1 proliferó de manera significativa, destruyendo la mayor parte de la población bacteriana (Lenski, 1988). Las pocas lecturas restantes de *E. coli* podrían corresponder a subpoblaciones que desarrollaron resistencia o a ADN bacteriano liberado por la lisis celular. Este escenario representa un claro ejemplo de presión selectiva extrema, donde la coevolución entre el hospedador y virus podría haber favorecido la aparición de mutaciones en genes clave para la resistencia bacteriana (Lenski & Levin, 1985).

3. Análisis de variantes y anotación funcional

El proceso de llamado y filtrado de variantes se llevó a cabo sobre ambas líneas mediante la herramienta **bcftools**, aplicando criterios de calidad (**QUAL** >30, **DP** >10, **MQ** >30). Posteriormente, las variantes fueron anotadas con **SnpEff** (Cingolani et al., 2012). Ambas muestras presentaron entre 450 y 470 variantes de alta calidad, de las cuales más del 95 % correspondieron a SNPs puntuales.

Muestra	Total	SNP	INS	DEL	Alto	Moderado	Bajo	Missense
EVOL1	473	458	6	9	6	95	84	95
EVOL2	453	439	4	10	7	96	79	97

Cuadro 1: Resumen de variantes detectadas según **SnpEff**.

La mayoría de las mutaciones detectadas presentaron un impacto moderado, destacándose principalmente variantes del tipo **missense** (o sea, cambios de aminoácidos en las proteínas). También se identificaron algunos casos más disruptivos, como mutaciones **frameshift** o **stop_gained**, que podrían tener consecuencias funcionales importantes dentro de la célula.

4. Variantes seleccionadas y discusión

Variantes de interés

Las variantes de impacto alto o moderado se asociaron principalmente con genes implicados en funciones clave relacionados con estructura de la membrana externa, el transporte de carbohidratos y la respuesta al estrés. Estas funciones están directamente relacionadas con mecanismos que podrían conferir resistencia frente a infecciones por fagos.

Muestra	Gen	Tipo	Efecto	Impacto	Función y relevancia biológica
EVOL1	<i>lamB</i>	SNP	Missense	Moderado	Porina de membrana externa, receptor del fago T1; mutaciones pueden impedir la unión viral (Hantke, 1978).
EVOL1	<i>manY</i>	DEL	Frameshift	Alto	Parte del sistema PTS de manosa, implicado en entrada fágica y transporte de azúcares (Ferenci, 1980).
EVOL2	<i>ompF</i>	SNP	Missense	Moderado	Porina general; su alteración reduce permeabilidad y adsorción viral (Nikaido, 2003).
EVOL2	<i>tolA</i>	INS	Inframe insertion	Alto	Componente del complejo Tol-Pal, esencial para la integridad de la envoltura celular (Lloubès et al., 2001).
EVOL2	<i>rpoB</i>	SNP	Missense	Moderado	Subunidad beta de la ARN polimerasa; regula la respuesta a estrés (Jin & Gross, 1988).

Cuadro 2: Variantes seleccionadas con posible relevancia adaptativa y referencias funcionales.

Discusión biológica

Las diferencias entre ambas líneas reflejan posibles trayectorias adaptativas divergentes frente a la misma presión viral (De Paepe & Taddei, 2006; Lenski, 1988). En el caso de EVOL1, la estructura bacteriana se mantiene relativamente estable, lo que podría indicar una fase previa a la infección o un equilibrio entre el fago y hospedador. EVOL2, en cambio, muestra un escenario de infección lítica más avanzado, evidenciado por mutaciones en genes de membrana externa como (*lamB*, *ompF*) que podrían haber evolucionado para bloquear la entrada del fago T1 (Hantke, 1978; Nikaido, 2003).

Además, la presencia de mutaciones en genes como *tolA* y *rpoB* sugiere que la adaptación no solo ocurrió a nivel de superficie celular, sino afectando también rutas intracelulares clave como la regulación y reparación del material genético (Lloubès et al., 2001; Jin & Gross, 1988). Estos resultados integran de forma coherente los patrones taxonómicos y genómicos, demostrando cómo la presión ejercida por los virus dirige la evolución bacteriana en distintos niveles (Lenski & Levin, 1985).

5. Conclusiones

- La comparación entre las líneas EVOL1 y EVOL2 revela una expansión viral masiva en la EVOL2, compatible con una infección lítica por fago T1.

- Las variantes detectadas indican una adaptación bacteriana frente a la presión ejercida por el fago, especialmente en genes relacionados con la membrana externa y con procesos de regulación transcripcional.
- Los resultados confirman la existencia de una coevolución entre el hospedador y el virus, y destacan la importancia de integrar análisis taxonómicos y genómicos en experimentos evolutivos.

Referencias

- Ondov, B. D., Bergman, N. H., & Phillippy, A. M. (2011). Interactive metagenomic visualization in a web browser. *BMC Bioinformatics*, 12(1), 385.
- Wood, D. E., Lu, J., & Langmead, B. (2019). Improved metagenomic analysis with Kraken 2. *Genome Biology*, 20(1), 257.
- Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., et al. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly*, 6(2), 80–92.
- Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*, 27(21), 2987–2993.
- Hantke, K. (1978). Major outer membrane proteins of *Escherichia coli* K12 serve as receptors for phage lambda and maltose transport. *Molecular and General Genetics*, 164(2), 131–135. <https://doi.org/10.1007/BF00267377>
- Ferenci, T. (1980). The recognition of maltose and maltodextrins by *Escherichia coli*. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 595(2), 255–276. <https://doi.org/10.1111/j.1432-1033.1980.tb04758.x>
- Nikaido, H. (2003). Molecular basis of bacterial outer membrane permeability revisited. *Microbiology and Molecular Biology Reviews*, 67(4), 593–656. <https://doi.org/10.1128/MMBR.67.4.593-656.2003>
- Lloubès, R., Cascales, E., Walburger, A., Bouveret, E., Lazdunski, C., Bernadac, A., & Journet, L. (2001). The Tol–Pal proteins of the *Escherichia coli* cell envelope: an energized system required for outer membrane integrity?. *Research in Microbiology*, 152(6), 523–529. [https://doi.org/10.1016/s0923-2508\(01\)01226-8](https://doi.org/10.1016/s0923-2508(01)01226-8)
- Jin, D. J., & Gross, C. A. (1988). Mapping and sequencing of mutations in the *Escherichia coli* *rpoB* gene that lead to rifampicin resistance. *Journal of Molecular Biology*, 202(1), 45–58.
- Ackermann, H. W. (2009). Phage classification and characterization. In M. H. Abedon (Ed.), *Bacteriophage Ecology* (pp. 1–10). Cambridge University Press.
- De Paepe, M., & Taddei, F. (2006). Viruses’ life history: Towards a mechanistic basis of a trade-off between survival and reproduction among phages. *PLoS Biology*, 4(7), e193. <https://doi.org/10.1371/journal.pbio.0040193>
- Lenski, R. E. (1988). Experimental studies of pleiotropy and epistasis in *Escherichia coli* I: Adaptation to novel environments. *Evolution*, 42(3), 425–432. <https://doi.org/10.1111/j.1558-5646.1988.tb04157.x>

Lenski, R. E., & Levin, B. R. (1985). Constraints on the coevolution of bacteria and virulent phage: A model and experiments with *E. coli* and phage T7. *The American Naturalist*, 125(4), 585–602. <https://doi.org/10.1086/284364>