



**Universidad
Europea MADRID**

Título del Trabajo:

***ANÁLISIS DE DIANAS TERAPÉUTICAS FRENTE A
ANTIMICROBIANOS, EN GENOMAS MÍNIMOS
BACTERIANOS.***

Autor:

Raquel Losada de Paz

Tutor del trabajo:

Dra. María del Rocío González Soltero

Tutor académico:

Dra. María del Rocío González Soltero

***Facultad de Ciencias Biomédicas y de la Salud
Titulación: Máster Universitario en Bioinformática
Curso 2021-2022***

ÍNDICE

| | |
|--|----|
| Glosario ----- | 2 |
| Resumen y palabras clave ----- | 4 |
| <i>Abstract y key words</i> ----- | 4 |
| Introducción ----- | 5 |
| 1. Genoma bacteriano e inestabilidad genómica ----- | 5 |
| 2. Resistencia a antibióticos ----- | 6 |
| 3. Simbiosis y reducción de genomas ----- | 9 |
| 4. Replicación ----- | 10 |
| a. Inicio ----- | 10 |
| b. Elongación (el replisoma) ----- | 11 |
| c. Terminación ----- | 13 |
| Planteamiento de hipótesis y objetivos del trabajo ----- | 14 |
| 1. Justificación ----- | 14 |
| 2. Hipótesis del trabajo ----- | 14 |
| 3. Objetivo general ----- | 15 |
| i. Objetivos específicos ----- | 15 |
| 4. Diseño del estudio ----- | 15 |
| Material y métodos ----- | 16 |
| 1. Obtención de genomas ----- | 16 |
| 2. Obtención de genes ----- | 17 |
| 3. Consulta de bases de datos mediante tBlastn ----- | 18 |
| Resultados ----- | 21 |
| Discusión de resultados ----- | 23 |
| Conclusiones ----- | 27 |
| Fortalezas y limitaciones ----- | 27 |
| Líneas de investigación futuras ----- | 27 |
| Anexos ----- | 28 |
| Bibliografía ----- | 33 |

GLOSARIO

1. **Genoma de referencia:** conjunto de datos para el modelaje y representación de la secuencia de ADN del genoma de una especie.
2. **Genoma mínimo:** tamaño de genoma formado por el mínimo número de genes suficientes para constituir un organismo celular de vida libre y que le permite reproducirse.
3. **Transferencia génica horizontal:** proceso en el que un organismo transfiere material genético a otra célula que no es su descendiente.
4. **Primosoma:** complejo formado por varias proteínas encargado de llevar a cabo los primeros pasos que conducen a la replicación del ADN.
5. **Replisoma:** conjunto de proteínas de una horquilla de replicación que permite la activación de ésta y permite la síntesis de DNA.
6. **Trinquete Müller:** fenómeno descrito en biología evolutiva, el cual ocurre en organismos con reproducción asexual, cuya progenie es idéntica a la inicial. En este proceso, se produce un aumento de la fijación de mutaciones deletéreas debido al reemplazamiento de unos genotipos por otros cada vez más mutados.
7. **Clusters of Orthologous Genes (COG):** Método que establece correspondencia entre grupos de genes entre varios genomas. Los grupos de genes ortólogos se corresponden normalmente con familias génicas que se expandieron antes de la divergencia de las especies comparadas. Este método, sin embargo, es incapaz de distinguir eventos de duplicación recientes de aquellos más antiguos (Tatusov et al. 1997; Tatusov et al. 2001).
8. **Mutación puntual:** mutación que se produce en un genoma cuando se añade, se desplaza o se cambia un par de bases. Muchas mutaciones puntuales son benignas, aunque también pueden tener consecuencias funcionales, incluyendo cambios en la expresión génica o alteraciones en las proteínas para las que codifica. Se producen por un cambio en un solo nucleótido del ADN, de tal manera que un nucleótido puede cambiar a otro nucleótido, se puede eliminar, o se puede insertar en el ADN, haciendo que el ADN sea diferente de la secuencia del gen normal.

9. **FASTA:** programa que compara a partir de secuencias proteicas o de DNA, otras secuencias de proteínas o DNA encontradas en diferentes bases de datos. Esto le permite realizar un alineamiento por discontinuidad. Para realizar la comparación de secuencias, la secuencia proteica o de DNA, tiene que tener un formato determinado, para que el programa pueda reconocerlo.

10. **Blast:** "Basic Local Alignment Sequence Tool". Herramienta que realiza alineamientos locales entre una secuencia problema y todas las secuencias de la base de datos que se están estudiando. Así se obtiene, a partir de la calidad del alineamiento puntuaciones para estos alineamientos. A mejor puntuación del alineamiento mayor similitud habrá entre las dos secuencias. Además, existe un valor E (e-value) el cual muestra si el alineamiento se ha producido o no al azar. Unos valores bajos del valor E indican que es muy probable que no se haya producido al azar. Existen diferentes tipos de algoritmos BLAST, en nuestro caso el utilizado ha sido tBLASTn el cual busca en bases de datos de nucleótidos con queries de nucleótidos.

11. **tBlastn:** algoritmo de BLAST traducido. Utiliza una secuencia de aminoácidos como consulta para compararla con la base de datos de nucleótidos traducida. Esta secuencia de aminoácidos se compara a nivel de proteína con cada secuencia de nucleótidos, y se traduce en seis marcos de lectura.

RESUMEN:

En los últimos años se ha producido un aumento de la resistencia a antibióticos, convirtiéndose en un problema de salud pública. Por ello, actualmente hay muchas investigaciones centradas en el descubrimiento de nuevas dianas terapéuticas que permitan la generación de nuevos antibióticos capaces de luchar contra estas bacterias resistentes. En este trabajo, se espera encontrar un conjunto de genes que conformen el genoma mínimo bacteriano de las bacterias endosimbiontes, y así, utilizarlos como posibles dianas terapéuticas. El estudio se realizó a partir del análisis de los genomas de diferentes familias de bacterias endosimbiontes, para obtener el genoma mínimo bacteriano, y así identificar los genes clave en la replicación del ADN de éstas. Estos genes serían imprescindibles para llevar a cabo el proceso replicativo, y, por tanto, para la supervivencia de estas bacterias endosimbiontes en su hospedador. Los resultados obtenidos muestran 6 genes que conforman el replisoma mínimo, lo que abre la posibilidad de generación de nuevos antibióticos utilizando estos genes como nuevas dianas terapéuticas.

PALABRAS CLAVE:

Bacterias, endosimbiontes, genoma mínimo, genes, replicación, diana terapéutica.

ABSTRACT:

In recent years there has been an increase in antibiotic resistance, which has become a public health problem. For this reason, there is currently a lot of research focused on the discovery of new therapeutic targets that allow the generation of new antibiotics capable of fighting against these resistant bacteria. In this work, we hope to find a set of genes that make up the minimal bacterial genome of endosymbiont bacteria, and therefore, use them as possible therapeutic targets. The study was carried out by analyzing the genomes of different families of endosymbiont bacteria to obtain the minimal bacterial genome and identify the key genes in their DNA replication. These genes would be essential for carrying out the replication process, and therefore, for the survival of these endosymbiont bacteria in their host. The results obtained show 6 genes that make up the minimal replisome, which opens up the possibility of generating new antibiotics using these genes as new therapeutic targets.

KEY WORDS

Bacteria, endosymbionts, minimal genome, genes, replication, therapeutic target.

INTRODUCCIÓN

1. Genoma bacteriano e inestabilidad genómica:

La información genética bacteriana aparece en un único cromosoma que consiste en una molécula de ADN de doble cadena circular. tamaño varía según la especie bacteriana de $0,1 \times 10^9$ a 8×10^9 dalton. Desde el punto de vista genético *Escherichia coli* (*E. coli*) es una de las bacterias más estudiadas. Su ADN tiene 3.200.000 pares de nucleótidos.

El tamaño del genoma está sometido a diferentes procesos de mutación que pueden conducir a cambios en él. Estos cambios pueden suceder a gran escala por duplicación del genoma, mientras que otros ocurren a pequeña escala por pérdida o ganancia de nucleótidos. Otras causas de variaciones en el tamaño del genoma son los elementos genéticos móviles, los cuales son considerados como unos de los causantes de la evolución del tamaño del genoma más importantes a largo plazo. Para que se produzca la fijación de estas mutaciones (*indel*) de manera probable, se necesita que estas afecten a pseudogenes, los cuales son genes no funcionales, o a otras secuencias de DNA que no tienen función. Si esta mutación afecta a genes de copia única, su fijación es poco probable, sobre todo si llevan a cabo funciones esenciales para la supervivencia de las células. Si un gen no esencial desaparece, se suele producir primero por su inactivación mediante una mutación puntual (formación de un pseudogen), seguido de la acumulación de mutaciones en el DNA que forma estos pseudogenes, hasta producirse la pérdida de todo vestigio de su presencia en el genoma (Pérez-Brocal et al., 2006).

Con respecto al aumento del tamaño del genoma de procariotas, uno de los procesos implicados más importante es la transferencia génica horizontal (Ochman et al., 2000). En este proceso, el DNA penetra en la célula de diferentes formas y para poder integrarse en el genoma de tiene que llevarse a cabo la recombinación. Si al introducirse le confieren alguna ventaja, se suelen mantener (mutación positiva). De lo contrario, se pueden producir nuevas mutaciones que llevan a su inactivación. Un ejemplo de la importancia de estos procesos de inestabilidad genómica se observa en *E. coli*. Algunas cepas se han convertido en patogénicas debido a la transferencia de genes de virulencia (González-Soltero, 2010). La atribución de genes les permite colonizar nuevos nichos, así como, la pérdida de genes les beneficia en su adaptación a otros nuevos. Esto se produce en las bacterias que establecen una relación simbiótica con su hospedador (Moran y Wernegreen, 2000).

Por último, los plásmidos son elementos genéticos extracromosómicos presentes en muchos microorganismos tanto de los dominios Bacteria y Archea, como en el Eucariota. Son moléculas

circulares de DNA de doble cadena con replicación autónoma, que aportan funciones a la célula hospedadora. Esto quiere decir que no son portadores de genes esenciales para el crecimiento y desarrollo del hospedador, a no ser que en condiciones ambientales adversas le proporcione ventajas adaptativas aumentando así su supervivencia. Los plásmidos tienen sistemas que les garantizan su autorreplicación. También mecanismos para mantener el número de copias en la célula bacteriana y asegurar así su permanencia en la célula.

2. Resistencia a antibióticos

Los antimicrobianos como son los antibióticos, antivíricos, antifúngicos y antiparasitarios son medicamentos utilizados para el tratamiento y prevención (en el caso de los antiparasitarios) de infecciones tanto en seres humanos como en animales y plantas.

La resistencia a antibióticos se produce por la ausencia de respuesta por parte de los microorganismos patógenos, a los medicamentos por cambios producidos en su genoma durante un periodo de tiempo. Esto dificulta el tratamiento de las infecciones lo cual puede aumentar el riesgo de propagación, así como, incrementar la gravedad de las enfermedades. Debido a esta farmacoresistencia, los antimicrobianos dejan de ser eficaces, incrementando la dificultad en el tratamiento de las enfermedades producidas por estos microorganismos resistentes (Organización Mundial de la Salud, [OMS], 2022).

La resistencia a antimicrobianos, en especial la resistencia a antibióticos, se ha convertido en un problema de salud pública. En la actualidad producen elevados de índices elevados de morbi-mortalidad en los hospitales. Según datos de la Organización Mundial de la Salud (OMS), la tasa de resistencia al ciprofloxacino se encuentra entre el 8,4% al 92,9% para *E. coli* y entre el 4,1% y el 79,4% para *Klebsiella pneumoniae* (en los países que han presentado datos al Sistema Mundial de Vigilancia de la Resistencia a los Antimicrobianos y de su Uso (GLASS)). Otro ejemplo es la resistencia de *E. coli* a las fluoroquinolonas. En la actualidad, en muchos países del mundo este tratamiento es ineficaz en más de la mitad de los pacientes tratados en los hospitales.

Por todo ello, es necesario el desarrollo de nuevos fármacos capaces de frenar el aumento de bacterias resistentes. Tradicionalmente, la identificación de nuevas dianas terapéuticas antimicrobianas se basaba en un proceso lineal, que consistía en nuevas moléculas químicas de productos naturales (Douglas et al., 2004). Actualmente, el desarrollo de nuevas herramientas moleculares en genómica y proteómica, como PCR en tiempo real, pirosecuenciación de ADN, espectrometría de masas, microarrays de ADN y la bioinformática, permiten conocer de forma más estrecha la fisiología y estructura de las bacterias y los mecanismos de resistencia a los

antibióticos. Estos estudios hacen posible identificar nuevas dianas farmacológicas y diseñar antibióticos específicos para suministrar tratamientos más certeros que combatan las infecciones producidas por bacterias. Estas técnicas también permiten la identificación de los genes que confieren la resistencia a los antibióticos y el reconocimiento de las estructuras genéticas como los integrones los cuales permiten que las bacterias se adapten y evolucionen rápidamente a través del almacenamiento y la expresión génica, que intervienen en la diseminación de los genes que producen la multirresistencia (Garza-Ramos et al., 2009).

Las dianas terapéuticas son receptores de membrana, receptores celulares u otras moléculas que están involucradas de manera fundamental en algún punto del desarrollo de la enfermedad. Desde un punto de vista farmacológico, una diana terapéutica es aquella que puede ser inhibida o activada por moléculas que conducen a una modulación de su función, teniendo generalmente otros efectos sobre el metabolismo celular. Así, estas moléculas pueden atacar a las dianas terapéuticas, desencadenando cascadas intracelulares de reacciones bioquímicas seguidas de reacciones celulares. Las potenciales dianas terapéuticas incluyen genes expresados de manera diferencial en diferentes individuos con una enfermedad o condición concreta, genes expresados diferencialmente por individuos sometidos a un tratamiento conocido para atacar los síntomas de interés y genes coexpresados junto con otros genes que se presume que están involucrados en los sistemas y rutas en estudio (Hossain et al., 2018). Por ello, en resumen, se consideran buenas dianas terapéuticas aquellas que tienen fuertes efectos en rutas biológicas específicas y mínimos efectos en otras rutas que no son de interés para las enfermedades en estudio (Jiang et al., 2005).

En la actualidad, el descubrimiento de nuevas dianas terapéuticas precisa de la incorporación de información de diferentes fuentes de datos, entre las que se incluyen datos de expresión génica, datos de secuenciación, polimorfismos, etc. Esta información se proporciona en bases de datos públicas. En este campo, la bioinformática está haciendo importantes contribuciones en la identificación de potenciales dianas terapéuticas. Pero todas ellas necesitan procesos de validación que las vinculen estrechamente con las enfermedades de interés que permitan demostrar que estas dianas terapéuticas tienen un potencial terapéutico relevante. Las bases de datos se han convertido en una herramienta crítica para el descubrimiento e identificación de nuevas dianas, y también como una plataforma muy valiosa para los investigadores y la industria farmacéutica para el desarrollo de nuevos fármacos y vacunas.

Tradicionalmente, la identificación de dianas farmacológicas era un proceso largo, costoso, laborioso y, a menudo, solo se podía identificar unas pocas dianas farmacológicas. Actualmente,

gracias a las técnicas *in silico*, es posible llevar a cabo una gran cantidad de análisis en menos tiempo, aumentando su rentabilidad. A menudo, estas técnicas ofrecen una gran cantidad de objetivos farmacológicos prometedores a partir de una gran cantidad de información. A día de hoy, la utilización de varios enfoques *in silico* para identificar objetivos potenciales de vacunas o fármacos se ha convertido en un requisito previo para el diseño de fármacos y vacunas. Uno de estos enfoques *in silico* es el análisis genómico sustractivo, que intenta descubrir nuevas proteínas u objetivos, que son importantes para la supervivencia del microorganismo patógeno, y no homólogos, en el huésped humano (Hossain et al., 2018). Los estudios basados en los enfoques genómicos sustractivos facilitan la selección, procesamiento y desarrollo de fármacos específicos contra patógenos. El uso de estos análisis sustractivos ayuda a la identificación de genes esenciales de los patógenos, los cuales no tienen homólogos en los genomas de los huéspedes. Estos genes esenciales no homólogos aseguran la supervivencia del patógeno, y por lo tanto, pueden ser el objetivo del desarrollo de fármacos. Se supone que los medicamentos contra tales objetivos potenciales tienen menos efectos secundarios y también se espera que las posibilidades de resistencia a los medicamentos sean bajas en el futuro (Shoukat et al., 2012).

En los últimos años, la bioinformática se ha convertido en una herramienta muy utilizada. El *big data* procedente de los análisis de genomas es utilizado para la búsqueda de nuevas dianas terapéuticas. Las dianas seleccionadas mediante la aproximación genómica a través del *big data*, deben cumplir dos criterios: que al ser eliminadas o inhibidas produzcan la muerte de la bacteria y que cumplan todas el mismo criterio, en nuestro caso, que hayan sufrido reducción del genoma en endosimbiontes.

3. Simbiosis y reducción de genomas

Se llama simbiosis a la relación establecida entre dos organismos que se necesitan mutuamente para sobrevivir. Un gran número de animales y plantas tienen microorganismos simbióticos que contribuyen a su nutrición o defensa contra depredadores y parásitos. Estas relaciones son consideradas como mutualistas, en las que tanto el huésped (animal/planta) como el simbionte microbiano se benefician de la asociación. En estas relaciones, la efectividad del simbionte varía con el genotipo del huésped (Douglas, A., 1998).

Una característica principal de las bacterias endosimbióticas es el tamaño reducido de sus genomas, en comparación con los genomas de bacterias de vida libre con las que están filogenéticamente relacionadas. La reducción de genomas es un proceso evolutivo dominante en endosimbiontes y parásitos de células eucariotas, que habilita a las bacterias a que se adapten al ciclo vital de estas células. Esto es debido a que, para adaptarse a la vida intracelular,

algunos genes considerados como esenciales para la vida libre se convierten en prescindibles, debido a que ya no son necesarios en el nuevo nicho ecológico o porque el producto génico necesario es producido por el hospedador. La idea de genomas mínimos se refiere al grupo de genes más pequeño posible que sería suficiente para mantener una forma de vida celular en funcionamiento en las condiciones más favorables imaginables, es decir, en presencia de un complemento completo de nutrientes esenciales y en ausencia del estrés ambiental (Koonin, 2000). Estas bacterias reducen sus genomas para tener las funciones celulares básicas para sobrevivir en su hospedador. Identificar estos genomas mínimos es una “llave” para identificar estas funciones vitales y poder “atacarlas” a través del desarrollo de nuevos fármacos basados en estos genomas mínimos (dianas terapéuticas). Algunas de las funciones básicas mantenidas en la reducción de genomas pertenecen a las categorías L y D COG, las cuales parecen resistir a los procesos de reducción de genomas. El enfoque COG se basa en la noción de que, cualquier grupo de al menos tres proteínas de genomas distantes, que son más similares entre sí que a cualquier otra proteína de los mismos genomas, probablemente pertenezcan a una familia de ortólogos. Por lo tanto, el enfoque COG se adapta tanto a los genes de evolución lenta como a los de evolución rápida.

La reducción de genomas se produce en una etapa inicial en la que se produce mucha inestabilidad genómica, aquí, se acumulan mutaciones en genes no esenciales, que unido al proceso denominado Trinquete de Müller, aumenta la fijación de mutaciones deletéreas debido al reemplazamiento de unos genotipos por otros cada vez más mutados (Santiago, 2010). Este proceso, explica cómo la acumulación de mutaciones en poblaciones asexuales puede desembocar en la pérdida de viabilidad de la población. Esto es debido a que las mutaciones deletéreas se fijan de forma irreversible (Manrubia, 2009). La probabilidad de que en una población asexual un genoma poco mutado acabe portando al menos una mutación depende en gran parte del tamaño del genoma y de la tasa de mutación. A pesar de ello, las reducciones del genoma, en especial en endosimbiontes, también se pueden producir por la selección producida para eliminar los genes que ya no son funcionales y, por lo tanto, no son necesarios.

A esta fase le sigue la etapa final, en la cual se produce estabilidad genómica por la pérdida de genes necesarios para la recombinación y de las secuencias repetidas durante las primeras etapas. Esta es una de las pocas oportunidades que permiten el intercambio genético debido al secuestro en un ambiente intracelular, lo cual también produce la reducción de genoma.

Estos genomas que han sufrido pérdida de material genético, y por tanto reducción del tamaño genómico, se caracterizan por el sesgo en A+T, la ausencia de sesgo adaptativo en el uso de

codones y la pérdida de elementos repetidos, todo ello junto a una elevada tasa de sustitución (Gonzalez-Soltero, 2009).

4. Replicación

Una de las funciones vitales celulares determinada por los genes esenciales a los que nos hemos referido anteriormente, es la replicación del genoma. Este proceso de replicación se divide en tres etapas: inicio de la replicación, elongación y terminación. Este proceso ha sido muy estudiado por la comunidad científica, siendo en bacterias gram negativas el organismo *E. coli* el utilizado como modelo, mientras que en organismos gram positivos se ha empleado *Bacillus subtilis* (*B. subtilis*).

4.1 Inicio:

En *E. coli*, la replicación se inicia en el origen de replicación *oriC*, que es una secuencia de 245 pares de bases (pb), que contiene una secuencia de 13 pb repetida tres veces en tándem. Además, esta región contiene cuatro zonas de unión a la proteína DnaA, que se encarga de abrir la hélice para comenzar la replicación. La unión de DnaA a sus sitios de fijación, y la fusión de la doble hélice, promueve la asociación secuencial de una serie de proteínas que desenrollan el DNA y forman la horquilla de replicación. Estas proteínas son DnaB y DnaG entre otras. La helicasa DnaB es la encargada de la separación de las hebras, permitiendo que la primasa DnaG genere el cebador necesario para la entrada de la abrazadera de DNA del complejo replicativo β . Una vez que la helicasa DnaB se sitúa en la burbuja de replicación, la primasa DnaG se une al complejo formando lo que se conoce como "primosoma" (González Soltero et al., 2006).

4.2. Elongación (el replisoma)

Se conoce como replisoma o complejo de replicación, al conjunto de proteínas de una horquilla de replicación que permite activarla y posibilitar la síntesis de DNA. Una vez formada la burbuja de replicación, la progresión de la horquilla de replicación en *E. coli* requiere la acción coordinada de cuatro actividades básicas que formarán el replisoma: proteína de unión a cadena sencilla (SSB), DNA Polimerasa, DNA helicasa y DNA primasa.

Para iniciar este proceso de síntesis, la proteína encargada de este proceso dentro del replisoma es la DNA polimerasa Holoenzima Pol III. En *E. coli* es la subunidad α Pol III (polimerasa de ADN de la familia C replicativa). Esta DNA polimerasa III holoenzima, es la principal encargada de mediar el proceso de replicación del ADN cromosómico. Está compuesta por el núcleo enzimático (subunidades α , ϵ y θ), la abrazadera deslizante β y el complejo γ (gen *dnaX* (γ - τ)). La subunidad α , es el centro de actividad de la polimerasa. En organismos del filo Firmicutes,

como *B. subtilis*, la iniciación se produce a través de dos polimerasas de ADN esenciales de la familia C replicativas, *polC* y *dnaE* (también llamada *dnaE3*). *polC*, contiene polimerasa de ADN y actividades correctoras de exonucleasa 3'→5' en una cadena polipeptídica, mientras que *dnaE*, no tiene actividad correctora. *dnaE*, presenta homología con Pol III de *E. coli*, y en menor medida con *polC* (Seco et al., 2017).

En la subunidad Pol III de *E. coli*, la subunidad α participa directamente en la síntesis de DNA y puede darse en dos formas distintas. En organismos como *B. subtilis*, la subunidad *polC* (polimerasa), se encarga de guiar la polimerización de la cadena líder, mientras que la subunidad *dnaE*, polimeriza la cadena rezagada. Los análisis filogenéticos indican que la *polC* es característica del phylum *Firmicutes*. Sin embargo, no todas las bacterias utilizan las dos subunidades en sus mecanismos de replicación. Los análisis bioinformáticos han encontrado que *polC* contiene dominios de exonucleasas correctores, los cuales están presentes en la misma posición de todos los genes de bacterias donde *polC* está presente (Corrales-Santander et al., 2018).

Puesto que, la molécula de DNA es antiparalela, cuando se produce la elongación de las nuevas cadenas se crea asimetría en el origen de replicación, porque la DNA polimerasa tiene actividad en sentido 5'→3'. Por ello, una cadena se sintetiza de manera continua y la otra discontinua en fragmentos de 2Kb, los llamados fragmentos de Okazaki (González-Soltero, 2009). Esto hace que la replicación necesite de dos polimerasas distintas. La DNA polimerasa necesita para iniciar la replicación *de novo*, un cebador sintetizado por la enzima DnaG (primasa). Esta enzima sintetiza pequeñas secuencias de RNA necesarias para iniciar cada fragmento de Okazaki. Ya que cada fragmento de Okazaki se inicia cada pocos segundos, para que la primasa tenga acceso a las cadenas abiertas, se une a la helicasa DnaB (Tougu et al., 1994). La enzima SSB, mantiene la forma de DNA de hebra simple de cada una de las cadenas recién abiertas por la helicasa (Glover y McHenry, 1998). Tras la síntesis de los fragmentos de Okazaki, los cebadores son eliminados por la RNAsa, junto con la DNA polimerasa I y su actividad exonucleasa (Lehman y Uyemura, 1976). Esta polimerasa sintetiza DNA para rellenar el hueco formado. Por último, la DNA ligasa une los fragmentos contiguos.

En *E. coli* se han descrito 4 polimerasas: Pol I (*gen polA*), Pol II (*gen polB*), Pol IV (*gen dinB*) y Pol V (*gen umuDC*).

La DNA polimerasa I en *E. coli* tiene 400 moléculas por célula. La DNA polimerasa II, tiene actividad reparadora de DNA y reinicia la síntesis de DNA en los huecos pequeños. Tiene una

tasa de error baja. La DNA polimerasa IV y la V reparan el DNA y son propensas a error, porque no presentan actividad exonucleasa 3' correctora (González Soltero et al., 2009).

Por último, cabe destacar que el cromosoma de *E. coli* es circular, impidiendo durante la elongación que la molécula de DNA tenga libertad de giro cuando la *DnaB* helicasa abre la doble hélice. Esto hace que el superenrollamiento positivo al que está sometido por la presencia del replisoma, aumente según avance la replicación produciendo un gran impedimento topológico. *gyrB* y *topA*, coordinan sus acciones para que se produzca el avance del replisoma.

La DNA girasa y la topoisomerasa IV, están compuestas por dos subunidades: *gyrA* y *gyrB* para girasa y *parC* y *parE* para la topoisomerasa IV (Corbett et al., 2004). Para cada enzima, estas subunidades se combinan en un complejo heterotetramérico (girasa GyrA2GyrB2 y topo IV ParC2ParE2) para formar las enzimas activas. La subunidad *parE* y la correspondiente subunidad B de girasa (*gyrB*), son las subunidades responsables de la unión e hidrólisis de ATP, por otro lado, *gyrA* y *parC* son responsables de la unión al ADN, reacción de escisión y religación.

gyrB y *ParE* son topoisomerasas de tipo II. Su función es vital para la replicación y reparación del DNA. Están estructuralmente relacionadas, y han sido reconocidas como candidatas principales para el desarrollo de agentes antibacterianos de doble acción con potencial de amplio espectro (Azam et al., 2015). Actúan rompiendo las dos cadenas de DNA y catalizando el paso a otra molécula de DNA a través de la rotura. Todas las topoisomerasas pueden relajar la superhelice, pero solo la girasa (*gyrB*) puede introducir superenrollamientos negativos en presencia de ATP. Por otro lado, la topoisomerasa tipo IV (*parE*) es la única significativamente importante en la decatenación y desnudo de las moléculas de DNA (Barbosa et al., 2012). Son enzimas esenciales y altamente conservadas que regulan los cambios en la topología y la integridad del DNA durante la replicación, la recombinación y la transcripción, en una reacción dependiente de ATP. Ambas enzimas también son esenciales de forma independiente para el crecimiento bacteriano. Aunque existen diferencias estructurales y funcionales específicas entre *gyrB* y *parE*, ambas catalizan las roturas de doble cadena del DNA y la posterior ligadura, y ambas se unen e hidrolizan ATP para proporcionar la energía necesaria para realizar estas tareas (Azam et al., 2015).

4.3. Terminación:

Cuando el complejo de replicación formado en *oriC* llega a la región opuesta, la replicación se detiene. Esto se denomina etapa de terminación. Si un punto de replicación alcanza la zona del término antes que el otro, es frenado por los complejos *Tus-Ter*. Estos complejos se forman por la unión de la proteína Tus (*tus*) a las secuencias terminadoras *Ter*. Estas secuencias se han

localizado tanto en el cromosoma como en algunos plásmidos. Y se ha visto que actúan sobre el replisoma bloqueando su avance de forma polar.

Una vez que la replicación llega a su fin, se produce la segregación de los cromosomas recién replicados. Fenómeno que depende frecuentemente de procesos de recombinación.

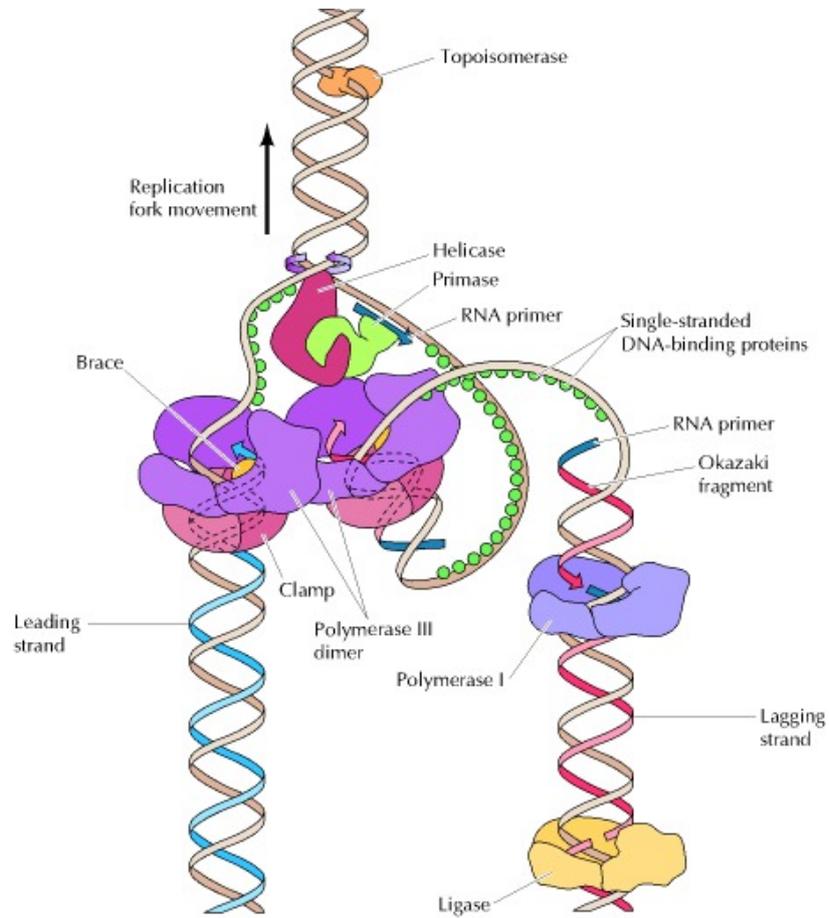


Imagen 1: Modelo de horquilla de replicación de *E. coli*. Cooper, 2000.

PLANTEAMIENTO DE HIPÓTESIS Y OBJETIVOS DEL TRABAJO

1. Justificación

Uno de los principales problemas en salud pública que se está desarrollando en la actualidad, es el aumento de la resistencia a antibióticos. Esto es debido, entre otros factores, al abuso que se ha hecho de estos medicamentos durante los últimos años. Este aumento de la resistencia a antibióticos es causante de elevados de índices de morbi-mortalidad hospitalaria. Por ello, es necesario el desarrollo de nuevos fármacos capaces de frenar el aumento de bacterias resistentes, o que sean capaces de eliminar estas bacterias ya resistentes a los actuales antibióticos presentes en el mercado. Para ello, el estudio de los genomas mínimos de estas bacterias que contienen genes esenciales, y que podría permitir el descubrimiento de nuevas dianas terapéuticas a partir de las cuales se desarrollen dichos fármacos, se ha convertido en una importante fuente conocimiento.

Las bacterias endosimbiontes a lo largo de su evolución, han tendido de manera dominante a reducir sus genomas hasta genomas mínimos, lo cual les ha permitido adaptarse al ciclo vital de sus hospedadores. Esto es debido a que, para adaptarse a la vida intracelular, algunos genes considerados como esenciales para la vida libre se convierten en prescindibles, debido a que ya no son necesarios en el nuevo nicho ecológico o porque el producto génico necesario lo produce el hospedador. Mientras que, para la vida intracelular, algunos genes se convierten en imprescindibles, como se ha dicho anteriormente, y otros genes permanecen como imprescindibles para el ciclo vital de estas bacterias. Entre estos genes esenciales destacan los implicados en la síntesis de las proteínas encargadas de la replicación celular. Dichos genes y proteínas son los que han sido estudiados en este trabajo.

2. Hipótesis de trabajo

La hipótesis principal consiste en que las bacterias endosimbiontes han reducido su genoma al mínimo, manteniendo solo los genes necesarios para actividades mínimas vitales, debido al efecto de los mecanismos mutación/selección natural sufridos durante la evolución de sus genomas y los procesos de adaptación al hospedador. Estas bacterias han conseguido un replisoma mínimo que puede servir como objetivo para el diseño de futuras dianas terapéuticas antimicrobianas.

Esta reducción de genomas se produce en todo tipo de bacterias, incluidas las que actualmente son resistentes a los antibióticos. Debido a esto, si se descubren dianas moleculares para fármacos presentes en todas las bacterias, que atacan a estos genes encargados de las

actividades vitales básicas de las mismas, se estarán desarrollando antibióticos eficaces contra las bacterias resistentes actuales.

Por todo ello, en este trabajo se pretende discernir cuales son las enzimas de replicación más conservadas en las diferentes familias de endosimbiontes estudiadas, por qué se han conservado a lo largo de la evolución, así como saber si existen estudios recientes en los cuales se haya demostrado que estas proteínas conservadas pueden resultar buenas dianas terapéuticas en el desarrollo de nuevos fármacos antimicrobianos.

3. Objetivo general

Identificar genes esenciales relacionados con la replicación del genoma, conservados en genomas mínimos de bacterias endosimbiontes, que puedan ser utilizados como potenciales dianas terapéuticas. Se trata de establecer un set de genes mínimos necesarios para el mantenimiento de la replicación del genoma.

3.1 Objetivos específicos:

3.1.1 Caracterizar el replisoma mínimo, es decir, los elementos funcionales necesarios para la replicación de genomas mínimos.

3.1.2 Teorizar la composición de los componentes del replisoma que puedan servir como dianas terapéuticas.

4. Diseño del estudio

Se trata de un estudio experimental “in silico” de corte descriptivo. Para el mismo, se seleccionaron 41 especies de endosimbiontes a partir de los cuales se obtuvieron 243 genomas conseguidos en la base de datos del NCBI.

Tras esto, se analizaron 26 genes de replicación seleccionados de *E. coli* y *B. subtilis*, para, a partir de ellos, localizar sus homólogos en los 243 genomas de las 41 especies endosimbiontes seleccionadas. Los genes seleccionados para la localización de homólogos, se encuentran representados en la tabla anexo 1. A estas especies de endosimbiontes se les asignó una clase. Esto se llevó a cabo para poder discernir en cuál de ellas estos genes se encuentran más conservados.

A partir de estos resultados, el objetivo es caracterizar el replisoma mínimo de los endosimbiontes estudiados y teorizar si estos componentes pueden actuar como dianas terapéuticas.

MATERIAL Y MÉTODOS:

1. Obtención de genomas

La búsqueda de genomas se hizo a partir de 41 especies bacterianas endosimbiontes, de las cuales se obtuvieron 243 genomas, ya que algunas de estas especies tienen varios genomas depositados en la base de datos. Estos genomas fueron obtenidos directamente del Genbank del NCBI por FTP:

<ftp://ftp.ncbi.nlm.nih.gov/genbank/genomes/Bacteria/>

Las 41 especies bacterianas utilizadas para la búsqueda fueron las siguientes:

| ESPECIES | ESPECIES |
|---|---|
| <i>Blattabacterium</i> | <i>Sodalis_glossinidius_morsitans</i> |
| <i>Blattabacterium</i> | <i>Candidatus_Hamiltonella_defensa</i> |
| <i>Candidatus_Sulcia_muelleri</i> | <i>Candidatus_Moranella_endobia</i> |
| <i>Candidatus_Sulcia_muelleri</i> | <i>Candidatus_Serratia_symbiotica</i> |
| <i>Candidatus_Sulcia_muelleri</i> | <i>Baumannia_cicadellinicola</i> |
| <i>Candidatus_Sulcia_muelleri</i> | <i>Candidatus_Tremblaya_princeps</i> |
| <i>Nitrosococcus_oceani</i> | <i>Candidatus_Tremblaya_princeps</i> |
| <i>Buchnera_aphidicola</i> | <i>Cupriavidus_taiwanensis</i> |
| <i>Buchnera_aphidicola</i> | <i>Nitrosomonas_eutropha</i> |
| <i>Buchnera_aphidicola</i> | <i>Polynucleobacter_necessarius</i> |
| <i>Buchnera_aphidicola</i> | <i>Verminephrobacter_eiseniae</i> |
| <i>Buchnera_aphidicola</i> | <i>Candidatus_Zinderia_insecticola</i> |
| <i>Buchnera_aphidicola</i> | <i>Candidatus_Pelagibacter_ubique</i> |
| <i>Candidatus_Blochmannia_floridanus</i> | <i>Candidatus_Hodgkinia</i> |
| <i>Candidatus_Blochmannia_pennsylvanicus</i> | <i>Wolbachia_endosymbiont_of_Brugia_malayi</i> |
| <i>Candidatus_Carsonella_ruddii</i> | <i>Wolbachia_endosymbiont_of_Culex_quinquefasciatus_Pel</i> |
| <i>Candidatus_Vesicomysocius_okutanii</i> | <i>Wolbachia_endosymbiont_of_Drosophila_melanogaster</i> |
| <i>Candidatus_Ruthia_magnifica</i> | <i>Burkholderia_mallei</i> |
| <i>Wigglesworthia_glossinidiae_endosymbiont_of_Glossina</i> | <i>Pseudomonas_aeruginosa</i> |
| <i>Wigglesworthia_glossinidia</i> | <i>Candidatus_Portiera_aleyrodidarum</i> |

Como punto inicial para el desarrollo del trabajo se comenzó llevando a cabo la selección de genomas, mediante la selección de los genomas de endosimbiontes reportados como genomas mínimos. La búsqueda reportó 243 genomas de endosimbiontes (tabla anexo 2). A continuación, se realizó un listado que incluye por orden: el nombre de la bacteria, el organismo del cual es

endosimbionte, y los links directos de todos los genomas del estudio. Los genomas son de endosimbiontes pertenecientes todos ellos al Phylum Proteobacteria.

Ejemplos e interpretación del listado:

- ***Baumannia_cicadellincola_Hc_Homalodisca_coagulata__uid58111;NC_007984***

Especie: *Baumannia cicadellincola*

Endosimbionte de: *Homalodisca coagulata*

Código de identificación del genoma: uid58111

Código de genoma: NC_007984

Al introducir el código del genoma de la especie en la base de datos del NCBI (BioSample), esta nos devuelve información relevante sobre el genoma, como su tamaño, la definición de la especie o los autores que la describieron, entre otros. En este caso, el genoma de *B. cicadellincola* está formado por 686194 bp, siendo DNA linear. Se encuentra la secuencia completa del genoma. Es una gamma proteobacteria.

- ***Blattabacterium__Blaberus_giganteus__uid165873;NC_017924***

Especie: *Blattabacterium*

Endosimbionte de: *Blaberus giganteus*

Código de identificación del genoma: uid165873

Código de genoma: NC_017924

El genoma de *Blattabacterium blaberus* reportado por BioSample está formado por 629165 bp, siendo DNA circular. Se encuentra la secuencia completa. Es una bacteria del grupo *Bacteroidetes* pertenecientes al orden de las Flavobacterias.

2. Obtención de genes

Se utilizaron como semillas genes componentes de los replisomas de las bacterias *E. coli* (estirpe K-12 MG1655) y *B. subtilis* (168). Para la selección de los grupos de genes implicados en el proceso de replicación y partición de genomas, correspondientes a las categorías COGs L y D (L: replicación, recombinación y reparación; D: división celular, ciclo celular y partición del cromosoma). En el anexo 1 se muestra la tabla en la que aparecen los genes implicados en la replicación, utilizados para realizar el estudio del genoma mínimo bacteriano de endosimbiontes. A continuación, mediante FTP se obtuvieron todos los .fna de estos genes. Las

secuencias FASTA de estos genes se obtuvieron de la base de datos del NCBI en formato de proteína traducida: mediante acceso FTP <ftp://ftp.ncbi.nih.gov/genomes/Bacteria/genes.fna>.

Para ello, se usó un script de descarga automática de ficheros del servidor FTP con dirección <ftp.ncbi.nih.gov>. Para ejecutar este script es necesaria la siguiente línea de comandos:

```
#java -jar "Ruta del script" "Ruta del fichero de nombres a descargar" "Ruta de la carpeta donde se descargan los ficheros" "Formato de los ficheros a descargar".
```

Al ejecutar el script de descarga automática, se obtuvieron las secuencias de proteínas en formato FASTA. Un ejemplo de estas secuencias es el mostrado a continuación:

Ejemplo:

```
>CAB11783.1 DNA gyrase (subunit A) [Bacillus subtilis subsp. subtilis str. 168]
MSEQNTPQVREINISQEMRTSFLDYAMSVIVSRALPDVRDGLKPVHRRILYAMNDLGMTSDKPYKKSARI
VGEVIGKYHPHGDSAVYESMVRMAQDFNYRYMLVDGHGNFGSVDGDSAAAMRYTEARMSKISMEILRDIT
KDTIDYQDNYDGSEREPVMPSRFPNLLVNGAAGIAVGMATNIPPHQLGEIIDGVLAVSENPDITIPELM
EVIPGPDFPTAGQILGRSGIRKAYESGRGSITIRAKAEIEQTSSGKERIIVTELPYQVNAKAKLIEKIADL
VRDKKIEGITDLRDESDRTGMRIVIEIRRDANANVILNPLYKQTALQTSFGINLLALVDGQPKVLTLLKQC
LEHYLDHQKVVIRRRRTAYELRKAEBARAHILEGLRVALDHLDAVISLIRNSQTAEIARTGLIEQFSLTEKQ
AQAILDMRLQRLTGLEREKIEEEYQSLVKLIAELKDILANEYKMLEIIREELTEIKERFNDERRTEIVTS
GLETIEDEDLIERENIVVTLTHNGYVKRLPASTYRSQKRGGKGVQGMGTNEDDFVEHLISTSTHDTILFF
SNKGKVYRAKGYEIPYGRGRTAKGIPIINLLEVEKGEWINAII PVTEFNAELYLFFTTKHGVSKRKTSLSQF
ANIRNNGLIALS LREDEL MGVR L TDG TKQIIIGTKNGLLIRFPETDVREMGR TAAGVKGITLTD DDDVVV
GMEILEEEESHVLIVTEKGYGKRTPAEEYRTQSRGGKGLKTAKITENNGQLVAVKATKGEEDLMIITASGV
LIRMDINDISITGRVQTQVRLIRMAEEEHVATVALVEKNEEDENEEEQEEV
```

Una vez realizada la descarga, se generó una base de datos sobre la cual se realizaron las consultas en Blast. Para ello, la metodología consistió en utilizar una instrucción tipo para generar una base de datos, mientras que en la parte de los ficheros de carga se introdujeron todos los ficheros .fna descargados previamente.

3. Consulta de bases de datos mediante tBLASTn

La consulta de bases de datos se realizó a partir de tBLASTn, que consiste en la búsqueda de secuencias de proteínas frente a secuencias de nucleótidos traducidas.

tBLASTn es un algoritmo BLAST traducido, en el que se utiliza una secuencia de aminoácidos como consulta, para compararla con la base de datos de nucleótidos traducida (secuencia codificante). La secuencia de aminoácidos se compara a nivel de proteína con cada secuencia de nucleótidos del sujeto, traducida en los seis marcos de lectura. Por ello, tBLASTn es muy útil para la búsqueda de homólogos de proteínas en datos de nucleótidos no anotados, como etiquetas de secuencias expresadas (mantenidas en la base de datos BLAST "est") y borradores de registros del genoma (ubicados en la base de datos BLAST "htgs"), que permanecen sin anotaciones en las respectivas bases de datos. El objetivo de este algoritmo es el de buscar las secuencias de proteínas homólogas de un par de secuencias de proteínas dadas (NP_001028007, NP_001028008) (Mukhopadhyay et al., 2017).

Esta consulta se realizó a partir de un script que se ejecuta a partir de la siguiente línea de comandos:

```
tblastn -db endosimbiontes -query /Users/antibioticos/Desktop/Script\ NCBI/
CAB11783.1.fna -outfmt 7 -out /Users/antibioticos/Desktop/resultados/
CAB11783.1.txt -evalue 0.000001
```

El e-value seleccionado es de 0.000001 para tBLASTn. Este umbral fue determinado por la tutora de este trabajo, María del Rocío González Soltero, durante su estancia en la Universidad de Roskilde, Dinamarca.

A continuación, se llevó a cabo el desarrollo de un nuevo script, a partir de un directorio donde se almacenaron los resultados de Blast. Se preparó el formato para que pudieran ser insertados en un libro Excel en el que se identificaron los nombres de cada bacteria.

```
/Users/antibioticos/NetBeansProjects/preparaResultados/dist/preparaResulta
dos.jar
"/Users/antibioticos/Documents/ResultadosBlast/ResultadosEndosimbiontes"
```

Tras la creación del libro de Excel, se generaron dos tablas: una tabla de equivalencia entre el código de los genomas (p.ej. NC_017381) y el código legible, y otra entre el código de los genes (p.ej. NP_415692) y el nombre legible. Estas tablas de equivalencia fueron insertadas en el siguiente libro de Excel: ResultadosEndosimbiontes.xlsx.

A partir de aquí, se crearon campos de búsqueda mediante funciones Excel para incluir en los resultados los nombres legibles, y a continuación se extrajo un fichero CSV de tres columnas que correspondía a conjuntos de gen, genomas y familias filogenéticas.

Una vez creados los campos de búsqueda y extraído el fichero CSV, se creó otro script para el análisis de los resultados que generó 2 ficheros de resultados:

1. EndosimbiontesTotal.csv: indica la probabilidad de que aparezca un gen concreto del conjunto de los genomas endosimbiontes analizados.
2. endosimbiontesFamilia.csv: Indica la probabilidad de que aparezca un gen concreto en un filo concreto dentro del conjunto de genomas endosimbiontes analizados.

A partir de los ficheros de resultado .csv, se creó un libro de excel al que se denominó: relacion_GEN-FAMILIA_Replicacion.xlsx, en el cual se recogen en las distintas columnas las proteínas, su código, el gen, y las familias en las que se encuentran presentes, así como un *count* de estas (número de veces que aparece ese gen en esa familia).

RESULTADOS

Tras llevar a cabo el análisis y la búsqueda de genes que conformarían el genoma mínimo bacteriano de endosimbiontes implicados en la replicación en diferentes familias, se obtuvo como resultado, que los genes más conservados en las bacterias, es decir, aquellos que se encuentran presentes en mayor número de familias de endosimbiontes, y que por tanto, podrían formar parte del replisoma mínimo necesario para la replicación de los genomas son:

- ***parC*** para la topoisomerasa IV subunidad α
- ***PolC*** de la polimerasa III
- ***dnaE*** de la polimerasa III subunidad α
- ***gyrA*** topoisomerasa II subunidad A
- ***parE*** topoisomerasa IV subunidad β
- ***ligA*** *ligasa*

Se seleccionaron estos genes como los componentes esenciales del posible replisoma mínimo, debido a que de todos los genes estudiados, estos son los que se encontraron en el mayor número de familias de entre todas las que fueron estudiadas, por lo que serán los genes mejor conservados evolutivamente.

Como se puede observar en la imagen 2, tanto *parC* como *polC* y *gyrA* se encontraron presentes en las 12 familias estudiadas. *dnaE* no forma parte del genoma mínimo de *actinobacterias* ni *tenericutes*, y *parE* no se encuentra en clamidias. Por último, *ligA* se encuentra presente en todas las familias excepto en *Spirochaetes*.

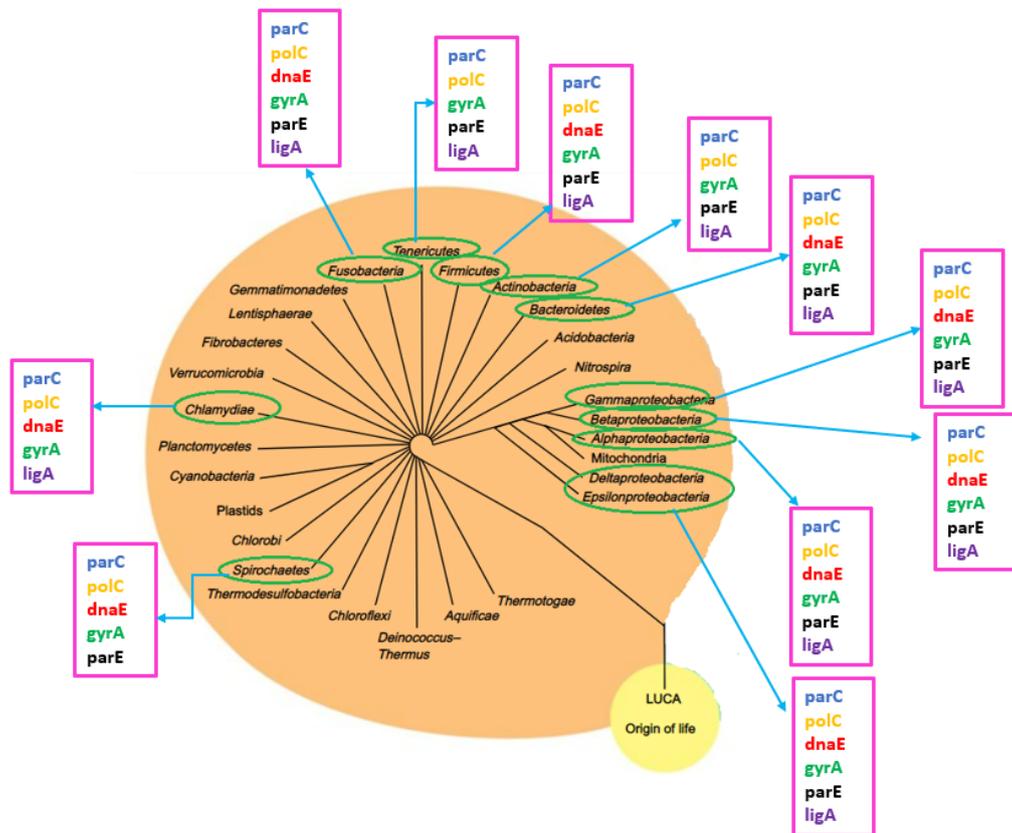


Imagen 2: representación de las familias de bacterias endosimbiotes estudiadas, junto con los genes que conforman su genoma mínimo. Se puede observar, como en todas estas familias, tanto parC, como polC y gyrA, forman parte de todos los genomas de las familias.

DISCUSIÓN DE RESULTADOS

Como se ha descrito anteriormente en los resultados, el replisoma mínimo caracterizado estaría conformado por un set de 6 genes. Estos genes están implicados de forma relevante e imprescindible en la maquinaria de replicación de los endosimbiontes estudiados. Sin estos, no podría llevarse a cabo.

En las bacterias, la familia C de ADN polimerasas se presenta en dos formas principales, *polC* y *dnaE*. *polC* representa un grupo compacto evolutivo, pero *dnaE* se puede subdividir en al menos tres grupos (*dnaE1-3*) (Zhao et al., 2006). *polC* contiene polimerasa de ADN y actividades de exonucleasa 3' → 5' correctoras en una cadena polipeptídica de *B. subtilis*. A diferencia de *polC*, *dnaE* no tiene actividad correctora y presenta homología con la subunidad Pol III de *E. coli*, y en menor medida con *polC* (Seco et al., 2017). Las polimerasas *dnaE3* combinadas con *polC* se encuentran en la clase *Bacilli*, mientras que en *Firmicutes*, *polC* se acompaña de *dnaE1* o *dnaE3* (Zhao et al., 2006).

Se han desarrollado estudios recientes que demuestran que *polC* es la encargada de la síntesis de la hebra principal durante la replicación de manera independiente a *dnaE*. A pesar de esto, *dnaE* extiende el cebador de RNA sintetizado por *dnaG* antes de que *polC* lo desplace, ya que esta, sintetiza la mayor parte de DNA en ambas hebras. Con esto se sugiere que *dnaE*, es el homólogo de la polimerasa en eucariotas en la síntesis de cadena retrasada, ya que, además, interactúa directamente con las helicasas replicativas *dnaC* y *dnaG* formando el primosoma (Seco et al., 2017). Estas dos enzimas (*polC* y *dnaE*), son imprescindibles para iniciar la replicación del DNA, ya que sin ellas no se puede llevar a cabo la síntesis de las nuevas hebras. Por este motivo, tiene sentido que no desaparezcan a lo largo de la evolución, sino al contrario, que se mantengan como parte esencial del genoma mínimo de los endosimbiontes.

Otros estudios anteriores sugieren que la polimerasa Pol III de *E. coli*, evolucionó para extender los cebadores de DNA y RNA (McHenry, 2011), mientras que en *B. subtilis*, *dnaE3* evolucionó siendo propenso a errores y a extender los cebadores de RNA. *polC* evolucionó para tener fidelidad, para extender solo los cebadores de DNA y para ser estimulada por *dnaG* en la cadena principal (Seco et al., 2017).

Por último, cabe destacar que *dnaE*, coevoluciona con el ribosoma, la maquinaria de transcripción, el núcleo de las enzimas del metabolismo intermediario y también está conectado a la enzima ahorradora de energía necesaria para la degradación del RNA, la polinucleótido fosforilasa. La mayoría de las proteínas de este conjunto coevolutivo pertenecen al conjunto

persistente en los proteomas bacterianos. Por el contrario, *polC* coevoluciona con enzimas de degradación de RNA que están presentes solo en el clado *Firmicutes*, rico en A+T (Engelen et al., 2012).

Todos estos resultados y conocimientos sobre ambas enzimas permiten entender por qué están tan conservadas en las familias estudiadas, ya que forman parte de la maquinaria básica replicativa para estos endosimbiontes, en los cuales, si alguna de ellas es inhibida no se puede iniciar la replicación. Para llevar a cabo el inicio de la replicación, tanto en *E. coli* como en *B. subtilis*, es necesario la presencia de *dnaG*, que es una primasa. En nuestro estudio bioinformático, esta enzima se encuentra presente en 10 de las familias estudiadas. Por lo tanto, se puede afirmar que tanto como *dnaE*, *polC*, y *dnaG*, son enzimas imprescindibles para la maquinaria de replicación de *B. bacillus*, así como la subunidad α Pol III de *E. coli*. Por ello, a pesar de que estas bacterias reduzcan sus genomas para la vida simbiote, esta reducción no afecta a los genes que codifican para estas enzimas, ya que, sin ellas, su maquinaria de replicación no funcionaría.

Recientemente, se ha descubierto un inhibidor de la proteína tirosin-fosfatasa CpsB, FQE, la cual inhibe el crecimiento de las bacterias gram positivas. CpsB, es una proteína miembro del dominio de la familia polimerasa e histidinol fosfatasa (PHP), que al igual que *polC*, está presente en una variedad de 22 patógenos bacterianos gram positivos. En este estudio de Standish et al., de la Universidad de Adelaide, se purificó el dominio PHP de *polC* (PolC PHP) demostrando que elimina la inhibición que hace FQE de la actividad de la fosfatasa CpsB. Además, se demostró que este dominio hidroliza el éster-nitrofenílico 5'-p25 de timidina-5'-monofosfato (pNP-TMP), utilizado como medida de la actividad exonucleasa. Por último, en este estudio se demostró que FQE no solo inhibe la actividad fosfatasa de CpsB, sino también la capacidad de PolC PHP para catalizar la hidrólisis de pNP-TMP (p-nitrofenil timidina 5' monofosfato). Todos estos resultados, sugirieron que *polC* puede ser el objetivo esencial de FQE, y que el dominio PHP puede representar un objetivo aún sin explotar para el desarrollo de nuevos antibióticos, debido a que al inhibirlo se suprimirían funciones críticas del patógeno bacteriano, al igual que procesos celulares esenciales tales como la replicación.

Otros genes implicados de manera imprescindible en la replicación celular de los obtenidos en nuestro estudio son *parE* y *gyrA*. *parE*, codifica para la topoisomerasa IV (enzima con cuatro subunidades) al igual que *parC*, mientras que *gyrA*, codifica para la DNA girasa compuesta por cuatro subunidades (dos tipos A y dos tipos B). *parC* y *parE*, son homólogas a *gyrA* y *gyrB* respectivamente (Mosquito et al., 2011). Se sabe que estas son inhibidas de manera satisfactoria

mediante antibióticos basados en quinolonas. Los antimicrobianos de quinolona se usan ampliamente en medicina clínica y son la única clase actual de agentes que inhiben directamente la síntesis de DNA bacteriano. Las quinolonas se dirigen, de manera dual, a la unión de la DNA girasa y la topoisomerasa IV con dominios y conformaciones específicos para bloquear la catálisis del paso de la hebra de DNA y estabilizar los complejos de DNA-enzima que bloquean el aparato de replicación del DNA, generando así, roturas dobles en el DNA que subyacen a su actividad bactericida. La resistencia a antimicrobianos ha surgido con el uso clínico de estos agentes. Los mecanismos de resistencia, incluyen alteraciones mutacionales en la afinidad por el objetivo del fármaco, en la expresión de la bomba de expulsión y en la adquisición de genes que confieren resistencia. Las mutaciones de resistencia en una o ambas de las dos enzimas diana del fármaco, se encuentran comúnmente en un dominio localizado de las subunidades *gyrA* y *parC* de la girasa y la topoisomerasa IV, respectivamente, y reducen la unión del fármaco al complejo enzima-DNA (Hooper et al., 2016).

La DNA girasa (*gyrB*) y las topoisomerasas IV (*parE*) han sido reconocidas durante mucho tiempo como los principales candidatos para el desarrollo de agentes antibacterianos de amplio espectro (Azam et al., 2015). Entre estos antibacterianos, destacan la delafloxacina y JNJ-Q2 (antibacteriano de fluoroquinolona de amplio espectro, que ha sido desarrollado para el tratamiento de infecciones bacterianas agudas de la piel, infecciones bacterianas agudas de la estructura de la piel, y neumonía adquirida en la comunidad (Jones et al., 2016)). Específicamente, JNJ-Q2 se está estudiando activamente para el tratamiento de infecciones por *Staphylococcus aureus* resistente a la metilicina. En los últimos años, la disponibilidad de la estructura cristalina de alta resolución de *gyrB* y *parE*, y la explotación de enfoques computacionales, han impulsado nuevos esfuerzos de investigación para desarrollar agentes antibacterianos robustos. (Azam et al., 2015).

Entre los antimicrobianos comercializados basados en la inhibición de girasas y topoisomerasas, está la novobiocina. En los estudios de Azam et al., se reveló que la resistencia a la novobiocina en *Staphylococcus aureus* se debe, predominantemente, a las mutaciones puntuales sucesivas que ocurrieron primero de manera específica en el gen *gyrB*, seguidas de una mutación puntual en el gen *parE* y luego una mutación puntual adicional en el gen *gyrB* (Nakamura et al., 2005). Estos hallazgos, revelaron que la DNA girasa es el objetivo principal y que la topoisomerasa IV es el objetivo secundario de la novobiocina. Los resultados de este estudio también sugirieron que la acumulación de mutaciones puntuales en los genes *gyrB* y *parE*, está asociada con un alto nivel de resistencia a la novobiocina en *S. aureus* (Azam et al., 2015).

Además de los anteriormente mencionados, otro gen implicado de manera imprescindible en la replicación bacteriana es *ligA*, el cual codifica para la ligasa de DNA. Las ligasas de DNA catalizan la formación de un enlace fosfodiéster entre un hidroxilo en 3' y un fosfato en 5'. Se trata de una reacción que es esencial para mantener la integridad del genoma durante la replicación y la reparación. Las ligasas de DNA de las eubacterias usan NAD⁺ como cofactor y poseen una homología estructural y de secuencia baja, en relación con las ligasas de DNA eucarióticas que usan ATP como cofactor. Esta diferencia es la que permite emplear las DNA ligasas bacterianas como estrategia antibacteriana (Alomari et al., 2021). La dihidrobenzoxazepina tricíclica y los derivados de indol tetracíclicos, se caracterizan por su capacidad para distinguir entre ligasas de DNA bacterianas y humanas, y por ello, están siendo estudiados como antimicrobianos cuya diana son las ligasas. A partir de diferentes estudios basados en la inhibición in vivo que emplearon *E. coli* GR501 y *S. typhimurium* LT2, rescatadas por DNA ligasa T4 dependiente de ATP o DNA ligasa dependiente de NAD (+) de *Mycobacterium tuberculosis* (Mtb LigA) respectivamente, se ha demostrado que estos dos compuestos pueden inhibir específicamente la *ligA* bacteriana (Yadav et al., 2015).

CONCLUSIONES:

Este estudio experimental *in silico* de corte descriptivo muestra la existencia de un conjunto de genes imprescindibles para la replicación de bacterias endosimbiontes.

Estos genes estudiados permiten el establecimiento de nuevas líneas de investigación para la generación de nuevos antimicrobianos utilizándolos como dianas terapéuticas.

FORTALEZAS Y LIMITACIONES

Cabe destacar que se trata de un estudio grande, que ha empleado un gran tamaño muestral de genomas (243), dando como resultado un conjunto de genes que pueden ser utilizados junto con estudios posteriores para el desarrollo de nuevos fármacos y líneas de investigación, a pesar de que se encuentra limitado con respecto a la extracción de datos. Esto ocurre ya que se ha llevado a cabo de manera manual, lo cual no permite la actualización de estos de manera constante, y por ello, el sistema no se retroalimenta y los datos son del momento puntual en el que esta extracción se llevó a cabo.

LÍNEAS DE INVESTIGACIÓN FUTURAS:

La existencia de un conjunto de genes que conforman el replisoma mínimo de bacterias endosimbiontes, permite la apertura de nuevas líneas de investigación que impulsen el desarrollo de nuevos fármacos antimicrobianos, capaces de combatir a las bacterias resistentes, y así, poder frenar el avance de esta resistencia en el ámbito de la salud pública.

ANEXOS

ANEXO 1: Tabla de los 26 genes implicados en la replicación seleccionados de *E. coli* y *B. subtilis*, utilizados para realizar la localización de homólogos en los 243 genomas de las 21 especies endosimbiontes seleccionadas. Los valores *not found* que aparecen para algunos genes, se interpretan como la ausencia de dicho gen en esa especie.

| Genes | Número de acceso <i>E.coli</i> | Número de acceso <i>B.</i> <i>subtilis</i> |
|------------------|---|---|
| <i>dnaA</i> | NP_417618 | CAB11777.1 |
| <i>hda</i> | YP_787988 | Not found |
| <i>dnaB</i> | NP_418476 | AAC00358 |
| <i>dnaC/dnaI</i> | NP_418781 | CAB14858 |
| <i>dnaG</i> | NP_417538 | NP_390400 |
| <i>gyrA</i> | NP_416734 | CAB11783.1 |
| <i>gyrB</i> | YP_026241 | CAB11782.1 |
| <i>parE</i> | NP_417502 | CAB13692 |
| <i>parC</i> | NP_417491 | CAB13693 |
| <i>topA</i> | NP_415790 | CAB13485 |
| <i>ligA</i> | NP_416906 | CAB12482 |
| <i>polC</i> | Not found | NP_389540 |
| <i>dnaE</i> | NP_414726 | CAB13531 |
| <i>dnaN</i> | NP_418156 | NP_387883 |
| <i>dnaQ</i> | NP_414751 | WP_004398027 |
| <i>dnaX</i> | NP_415003 | CAB11795 |
| <i>holA</i> | NP_415173 | CAB14498.1 |
| <i>holB</i> | NP_415617 | CAB11807 |
| <i>holC</i> | NP_418680 | Not found |
| <i>holE</i> | NP_416356 | Not found |
| <i>polA</i> | NP_418300 | Not found |
| <i>polB</i> | NP_414602 | Not found |
| <i>umuC</i> | NP_415702 | Not found |
| <i>umuD</i> | NP_415701 | Not found |
| <i>dinB</i> | NP_414766 | Not found |

ANEXO 2: Tabla de los genomas de las 21 especies de endosimbiontes utilizadas. Se reportan más genomas que especies, debido a que en las bases de datos aparecen depositados varios genomas para las mismas especies, como se puede observar en la tabla.

| Espece | Genoma | Grupo |
|--|-----------|---|
| <i>Candidatus_Pelagibacter</i> | uid58401 | <i>a-proteobacteria</i> |
| <i>Wolbachia_endosymbiont_of_Culex_quinquefasciatus</i> | uid61645 | <i>a-proteobacteria</i> |
| <i>Wolbachia_endosymbiont_of_Drosophila_melanogaster_</i> | uid57851 | <i>a-proteobacteria</i> |
| <i>Candidatus_Hodgkinia_cicadicola</i> | uid59311 | <i>a-proteobacteria</i> |
| <i>Blattabacterium__Blaberus_giganteus</i> | uid165873 | <i>a-proteobacteria</i> |
| <i>Blattabacterium__Blatta_orientalis</i> | uid188115 | <u><i>Bacteroidetes/Chlorobi group;</i></u> |
| <i>Blattabacterium__Blattella_germanica</i> | uid41533 | <u><i>Bacteroidetes/Chlorobi group;</i></u> |
| <i>Blattabacterium__Cryptocercus_punctulatus</i> | uid81083 | <u><i>Bacteroidetes/Chlorobi group;</i></u> |
| <i>Blattabacterium__Mastotermes_darwiniensis</i> | uid77127 | <u><i>Bacteroidetes/Chlorobi group;</i></u> |
| <i>Blattabacterium__Panesthia_angustipennis_spadica</i> | uid193717 | <u><i>Bacteroidetes/Chlorobi group;</i></u> |
| <i>Blattabacterium__Periplaneta_americana</i> | uid41287 | <u><i>Bacteroidetes/Chlorobi group;</i></u> |
| <i>Candidatus_Sulcia_muelleri</i> | uid52535 | <u><i>Bacteroidetes/Chlorobi group;</i></u> |
| <i>Candidatus_Sulcia_muelleri</i> | uid47075 | <u><i>Bacteroidetes/Chlorobi group;</i></u> |
| <i>Candidatus_Sulcia_muelleri</i> | uid58943 | <u><i>Bacteroidetes/Chlorobi group;</i></u> |
| <i>Candidatus_Sulcia_muelleri</i> | uid59393 | <u><i>Bacteroidetes/Chlorobi group;</i></u> |
| <i>Rhodothermus_marinus</i> | uid41729 | <u><i>Bacteroidetes/Chlorobi group;</i></u> |
| <i>Rhodothermus_marinus</i> | uid72767 | <u><i>Bacteroidetes/Chlorobi group;</i></u> |
| <i>Burkholderia_mallei_ATCC_23344</i> | uid57725 | <u><i>Bacteroidetes/Chlorobi group;</i></u> |
| <i>Burkholderia_mallei_NCTC_10229</i> | uid58383 | <i>b-proteobacteria</i> |
| <i>Burkholderia_mallei_NCTC_10247</i> | uid58385 | <i>b-proteobacteria</i> |
| <i>Burkholderia_mallei_SAVP1_</i> | uid58387 | <i>b-proteobacteria</i> |
| <i>Candidatus_Tremblaya_princeps</i> | uid68741 | <i>b-proteobacteria</i> |
| <i>Candidatus_Tremblaya_princeps</i> | uid159519 | <i>b-proteobacteria</i> |
| <i>Candidatus_Zinderia_insecticola</i> | uid52459 | <i>b-proteobacteria</i> |
| <i>Cupriavidus_taiwanensis</i> | uid61615 | <i>b-proteobacteria</i> |
| <i>Nitrosomonas_eutropha</i> | uid58363 | <i>b-proteobacteria</i> |
| <i>Polynucleobacter_necessarius_asympioticus</i> | uid58611 | <i>b-proteobacteria</i> |
| <i>Polynucleobacter_necessarius</i> | uid58967 | <i>b-proteobacteria</i> |
| <i>Verminephrobacter_eiseniae</i> | uid58675 | <i>b-proteobacteria</i> |
| <i>Candidatus_Vesicomysocius_okutanii</i> | uid59427 | <i>b-proteobacteria</i> |
| <i>Candidatus_Ruthia_magnifica_Cm__Calyptogena_magnifica</i> | uid58645 | <i>Gamma</i> proteobacteria; |
| <i>Candidatus_Portiera_aleyrodidarum</i> | uid173859 | <i>Gamma</i> proteobacteria; |
| <i>Candidatus_Portiera_aleyrodidarum</i> | uid176373 | <i>Gamma</i> proteobacteria; |

| | | |
|--|-----------|-----------------------------|
| <i>Candidatus_Portiera_aleyrodidarum</i> | uid175570 | <i>Gammaproteobacteria;</i> |
| <i>Candidatus_Portiera_aleyrodidarum</i> | uid176374 | <i>Gammaproteobacteria;</i> |
| <i>Candidatus_Portiera_aleyrodidarum</i> | uid195460 | <i>Gammaproteobacteria;</i> |
| <i>Nitrosococcus_oceani</i> | uid58403 | <i>Gammaproteobacteria;</i> |
| <i>Baumannia_cicadellinicola__Homalodisca_coagulata</i> | uid58111 | <i>Gammaproteobacteria;</i> |
| <i>Buchnera_aphidicola__Cinara_tujafilina</i> | uid68101 | <i>Gammaproteobacteria;</i> |
| <i>Buchnera_aphidicola__Acyrtosiphon_pisum__</i> | uid59285 | <i>Gammaproteobacteria;</i> |
| <i>Buchnera_aphidicola_Ak__Acyrtosiphon_kondoi__</i> | uid158533 | <i>Gammaproteobacteria;</i> |
| <i>Buchnera_aphidicola_APS__Acyrtosiphon_pisum__</i> | uid57805 | <i>Gammaproteobacteria;</i> |
| <i>Buchnera_aphidicola_Bp__Baizongia_pistaciae__</i> | uid57827 | <i>Gammaproteobacteria;</i> |
| <i>Buchnera_aphidicola_Cc__Cinara_cedri__</i> | uid58579 | <i>Gammaproteobacteria;</i> |
| <i>Buchnera_aphidicola_JF98__Acyrtosiphon_pisum__</i> | uid158845 | <i>Gammaproteobacteria;</i> |
| <i>Buchnera_aphidicola_JF99__Acyrtosiphon_pisum__</i> | uid158847 | <i>Gammaproteobacteria;</i> |
| <i>Buchnera_aphidicola_LL01__Acyrtosiphon_pisum__</i> | uid158843 | <i>Gammaproteobacteria;</i> |
| <i>Buchnera_aphidicola_Sg__Schizaphis_graminum__</i> | uid57913 | <i>Gammaproteobacteria;</i> |
| <i>Buchnera_aphidicola_TLW03__Acyrtosiphon_pisum__</i> | uid158849 | <i>Gammaproteobacteria;</i> |
| <i>Buchnera_aphidicola_Tuc7__Acyrtosiphon_pisum__</i> | uid59283 | <i>Gammaproteobacteria;</i> |
| <i>Buchnera_aphidicola_Ua__Uroleucon_ambrosiae__</i> | uid158535 | <i>Gammaproteobacteria;</i> |
| <i>Candidatus_Blochmannia_floridanus__</i> | uid57999 | <i>Gammaproteobacteria;</i> |
| <i>Candidatus_Blochmannia_pennsylvanicus</i> | uid58329 | <i>Gammaproteobacteria;</i> |
| <i>Candidatus_Carsonella_ruddii_CE_isolate</i> | uid172732 | <i>Gammaproteobacteria;</i> |
| <i>Candidatus_Carsonella_ruddii_CS_isolate</i> | uid172733 | <i>Gammaproteobacteria;</i> |
| <i>Candidatus_Carsonella_ruddii_HC_isolate</i> | uid172734 | <i>Gammaproteobacteria;</i> |
| <i>Candidatus_Carsonella_ruddii_HT_isolate</i> | uid172735 | <i>Gammaproteobacteria;</i> |
| <i>Candidatus_Carsonella_ruddii_PC_isolate</i> | uid172736 | <i>Gammaproteobacteria;</i> |
| <i>Candidatus_Carsonella_ruddii</i> | uid58773 | <i>Gammaproteobacteria;</i> |
| <i>Candidatus_Hamiltonella_defensa_5AT__Acyrtosiphon_pisum</i> | uid59289 | <i>Gammaproteobacteria;</i> |
| <i>Candidatus_Moranella_endobia_PCIT</i> | uid68739 | <i>Gammaproteobacteria;</i> |
| <i>Candidatus_Moranella_endobia_PCVAL</i> | uid197215 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli__BL21_Gold_DE3_pLysS_AG</i> | uid59245 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli__clone_D_i14</i> | uid162049 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli__clone_D_i2</i> | uid162047 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_042</i> | uid161985 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_536</i> | uid58531 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_55989</i> | uid59383 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_ABU_83972</i> | uid161975 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_APEC_O1</i> | uid58623 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_APEC_O78</i> | uid187277 | <i>Gammaproteobacteria;</i> |

| | | |
|---|-----------|-----------------------------|
| <i>Escherichia_coli_ATCC_8739</i> | uid58783 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_B_REL606</i> | uid58803 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_BL21_DE3</i> | uid161947 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_BL21_DE3</i> | uid161949 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_BW2952</i> | uid59391 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_CFT073</i> | uid57915 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_DH1</i> | uid161951 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_DH1</i> | uid162051 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_E24377A</i> | uid58395 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_ED1a</i> | uid59379 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_ETEC_H10407</i> | uid161993 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_HS</i> | uid58393 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_IAI1</i> | uid59377 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_IAI39</i> | uid59381 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_IHE3034</i> | uid162007 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_K_12_substr__DH10B</i> | uid58979 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_K_12_substr__MDS42</i> | uid193705 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_K_12_substr__MG1655</i> | uid57779 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_K_12_substr__W3110</i> | uid161931 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_KO11FL</i> | uid162099 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_KO11FL</i> | uid52593 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_LF82</i> | uid161965 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_NA114</i> | uid162139 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_O103_H2_12009</i> | uid41013 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_O104_H4_2009EL_2050</i> | uid175905 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_O104_H4_2009EL_2071</i> | uid176128 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_O104_H4_2011C_3493</i> | uid176127 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_O111_H__11128</i> | uid41023 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_O127_H6_E2348_69</i> | uid59343 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_O157_H7_EC4115</i> | uid59091 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_O157_H7_EDL933</i> | uid57831 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_O157_H7</i> | uid57781 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_O157_H7_TW14359</i> | uid59235 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_O26_H11_11368</i> | uid41021 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_O55_H7_CB9615</i> | uid46655 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_O55_H7_RM12579_</i> | uid162153 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_O7_K1_CE10</i> | uid162115 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_O83_H1_NRG_857C</i> | uid161987 | <i>Gammaproteobacteria;</i> |

| | | |
|--|-----------|-----------------------------|
| <i>Escherichia_coli_P12b</i> | uid162061 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_S88</i> | uid62979 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_SE11</i> | uid59425 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_SE15</i> | uid161939 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_SMS_3_5</i> | uid58919 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_UM146</i> | uid162043 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_UMN026</i> | uid62981 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_UMNK88</i> | uid161991 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_UTI89</i> | uid58541 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_W</i> | uid162011 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli</i> | uid162101 | <i>Gammaproteobacteria;</i> |
| <i>Escherichia_coli_Xuzhou21</i> | uid163995 | <i>Gammaproteobacteria;</i> |
| <i>Wigglesworthia_glossinidia_endosymbiont_of_Glossina_brevipalpis</i> | uid57853 | <i>Gammaproteobacteria;</i> |
| <i>Wigglesworthia_glossinidia_endosymbiont_of_Glossina_morsitans_Yale_colony</i> | uid88075 | <i>Gammaproteobacteria;</i> |
| <i>Pseudomonas_aeruginosa_B136_33</i> | uid196598 | <i>Gammaproteobacteria;</i> |
| <i>Pseudomonas_aeruginosa_DK2</i> | uid168996 | <i>Gammaproteobacteria;</i> |
| <i>Pseudomonas_aeruginosa_LESB58</i> | uid59275 | <i>Gammaproteobacteria;</i> |
| <i>Pseudomonas_aeruginosa_M18</i> | uid162089 | <i>Gammaproteobacteria;</i> |
| <i>Pseudomonas_aeruginosa_NCGM2_S1</i> | uid162173 | <i>Gammaproteobacteria;</i> |
| <i>Pseudomonas_aeruginosa_PA7</i> | uid58627 | <i>Gammaproteobacteria;</i> |
| <i>Pseudomonas_aeruginosa_PAO1</i> | uid57945 | <i>Gammaproteobacteria;</i> |
| <i>Pseudomonas_aeruginosa_RP73</i> | uid209328 | <i>Gammaproteobacteria;</i> |
| <i>Pseudomonas_aeruginosa_UCBPP_PA14</i> | uid57977 | <i>Gammaproteobacteria;</i> |
| <i>Candidatus_Pelagibacter_ubique_HTCC1062</i> | uid58401 | <i>Gammaproteobacteria;</i> |

BIBLIOGRAFÍA

Alomari A, Gowland R, Southwood C, Barrow J, Bentley Z, Calvin-Nelson J, Kaminski A, LeFevre M, Callaghan AJ, Vincent HA, Gowers DM. *Identification of Novel Inhibitors of Escherichia coli DNA Ligase (LigA)*. *Molecules*. 2021 Apr 25;26(9):2508. doi: 10.3390/molecules26092508. PMID: 33923034; PMCID: PMC8123306.

Barbosa LC, Garrido SS, Garcia A, Delfino DB, Santos Ldo N, Marchetto R. *Design and synthesis of peptides from bacterial ParE toxin as inhibitors of topoisomerases*. *Eur J Med Chem*. 2012 Aug;54:591-6. doi: 10.1016/j.ejmech.2012.06.008. Epub 2012 Jun 15. PMID: 22749642.

Brilli, M., Fondi, M., Fani, R., Mengoni, A., Ferri, L., Bazzicalupo, M., & Biondi, E. G. (2010). *The diversity and evolution of cell cycle regulation in alpha-proteobacteria: a comparative genomic analysis*. *BMC Systems Biology*, 4(1), 52.

Cooper GM. *The Cell: A Molecular Approach*. 2nd edition. Sunderland (MA): Sinauer Associates; 2000. DNA Replication. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK9940/> Azam, Mohammed Afzal; Thathan, Janarthanan; Jubie, Selvaraj (2015). *Dual targeting DNA gyrase B (GyrB) and topoisomerase IV (ParE) inhibitors: A review*. *Bioorganic Chemistry*, 62(), 41–63. doi:10.1016/j.bioorg.2015.07.004

Douglas, A. *Host benefit and the evolution of specialization in symbiosis*. *Heredity* 81, 599–603 (1998).

Elena, S. F. (2010). *Una aproximación experimental a la evolución viral: desentrañando los papeles de la mutación, la selección y el azar*.

Engelen, S., Vallenet, D., Médigue, C., & Danchin, A. (2012). *Distinct co-evolution patterns of genes associated to DNA polymerase III DnaE and PolC*. *BMC Genomics*, 13(1), 69. doi:10.1186/1471-2164-13-69

Garza-Ramos, Ulises, Jesús Silva-Sánchez, y Esperanza Martínez-Romero. «Genética y genómica enfocadas en el estudio de la resistencia bacteriana». *Salud Pública de México* 51 (2009): s439-46.

González-Soltero, R. (2009). *Análisis bioinformático de genes aplicados en Replicación de genomas de bacterias Endosimbiontes (trabajo de fin de master) Universidad Internacional de Andalucía*.

González-Soltero, R., Botello, E., & Jiménez-Sánchez, A. (2006). *Initiation of heat-induced replication requires DnaA and the L-13-mer of oriC*. *Journal of bacteriology*, 188(23), 8294-8298.

Hooper DC, Jacoby GA. *Topoisomerase Inhibitors: Fluoroquinolone Mechanisms of Action and Resistance*. Cold Spring Harb Perspect Med. 2016 Sep 1;6(9):a025320. doi: 10.1101/cshperspect.a025320. PMID: 27449972; PMCID: PMC5008060.

Hossain, M. U., Omar, T. M., Alam, I., Das, K. C., Mohiuddin, A. K. M., Keya, C. A., & Salimullah, M. (2018). *Pathway based therapeutic targets identification and development of an interactive database CampyNIBase of Campylobacter jejuni RM1221 through non-redundant protein dataset*. *PLOS ONE*, 13(6), e0198170.

Jarvis, T. C., Beaudry, A. A., Bullard, J. M., Janjic, N., & McHenry, C. S. (2004). *Reconstitution of a Minimal DNA Replicase from Pseudomonas aeruginosa and Stimulation by Non-cognate Auxiliary Factors*. *Journal of Biological Chemistry*, 280(9), 7890–7900.

Jiang, Z., & Zhou, Y. (2005). *Using gene networks to drug target identification*. *Journal of Integrative Bioinformatics*, 2(1).

Jones TM, Johnson SW, DiMondi VP, Wilson DT. *Focus on JNJ-Q2, a novel fluoroquinolone, for the management of community-acquired bacterial pneumonia and acute bacterial skin and skin structure infections*. *Infect Drug Resist*. 2016 Jun 7;9:119-28. doi: 10.2147/IDR.S105620. PMID: 27354817; PMCID: PMC4908938.

K.D. Corbett, J.M. Berger, *Annu. Rev. Biophys. Biomol. Struct.* 33 (2004) 95e118.

Koonin, E. V. (2000). *HOW MANY GENES CAN MAKE A CELL: The Minimal-Gene-Set Concept*. *Annual Review of Genomics and Human Genetics*, 1(1), 99–116.

M.F. Nakamura, H. Ito, Y. Oyamada, T. Nishino, J. Yamagishi, *Antimicrob. Agents Chemother.* 49 (2005) 3810-3815.

Manrubia, S. C. (2009). *Biodiversidad, asexualidad y virus: tres esbozos sobre la complejidad de la evolución*. *Apuntes de Ciencia y Tecnología*, 32, 29-36.

McHenry, C.S. (2011) *Breaking the rules: bacteria that use several DNA polymerase IIIs*. *EMBO Rep.*, 12, 408–414.

Moran, Wernegreen. *Lifestyle evolution in symbiotic bacteria: insights from genomics*. *Trends Ecol. Evol. (Amst.)*. 2000;15(8):321-326.

Mosquito, S., Ruiz, J., Bauer, J. L., & Ochoa, T. J. (2011). *Mecanismos moleculares de resistencia antibiótica en Escherichia coli asociadas a diarrea*. *Revista peruana de medicina experimental y salud pública*, 28, 648-656.

Mukhopadhyay, C. S., Choudhary, R. K., & Iquebal, M. A. (2017). *Basic Applied Bioinformatics*. John Wiley & Sons.

Ochman, Howard, B.Jones, Isaac (2000). *Dinámica evolutiva del contenido completo del genoma en Escherichia coli*. *El Diario EMBO vol. 19 No. 24 págs. 6637±6643, 20*

Organización Mundial de la Salud (13 de octubre 2020). Resistencia a los antimicrobianos. <https://www.who.int/es/news-room/fact-sheets/detail/antimicrobial-resistance>

Perez-Brocal, V., Gil, R., Ramos, S., Lamelas, A., Postigo, M., Michelena, J. M., ... Latorre, A. (2006). *A Small Microbial Genome: The End of a Long Symbiotic Relationship?* *Science*, 314(5797), 312–313. doi:10.1126/science.1130441

Rubio A., Jimenez J., Pérez-Pulido A. J., Assessment of selection pressure exerted on genes from complete pangenomes helps to improve the accuracy in the prediction of new genes, *Briefings in Bioinformatics*, Volume 23, Issue 2, March 2022, bbac010.

Sabater-Muñoz, B., & Toft, C. (2020). *Evolution from Free-Living Bacteria to Endosymbionts of Insects: Genomic Changes and the Importance of the Chaperonin GroEL*. *Symbiosis: Cellular, Molecular, Medical and Evolutionary Aspects*, 77–103. doi:10.1007/978-3-030-51849-3_3

Seco, E. M., & Ayora, S. (2017). *Bacillus subtilis DNA polymerases, PolC and DnaE, are required for both leading and lagging strand synthesis in SPP1 origin-dependent DNA replication*. *Nucleic Acids Research*, 45(14), 8302–8313.

Sharples, G. J. (2009). *For absent friends: life without recombination in mutualistic γ -proteobacteria*. *Trends in Microbiology*, 17(6), 233–242

Shoukat, K., Rasheed, N., Sajid, M. *Subtractive genome analysis for in silico identification and characterization of novel drug targets in c. trachomatis strain d/uw-3/cx*. *International Journal of Current Research* vol.4, Issue, 07, pp.017-021, July, 2012.

Van Ham, R. C. H. J., Kamerbeek, J., Palacios, C., Rausell, C., Abascal, F., Bastolla, U., ... Moya, A. (2003). *Reductive genome evolution in Buchnera aphidicola*. *Proceedings of the National Academy of Sciences*, 100(2), 581–586.

Yadav N, Khanam T, Shukla A, Rai N, Hajela K, Ramachandran R. *Tricyclic dihydrobenzoxazepine and tetracyclic indole derivatives can specifically target bacterial DNA ligases and can distinguish them from human DNA ligase I*. *Org Biomol Chem*. 2015 May 21;13(19):5475-87. doi: 10.1039/c5ob00439j. PMID: 25875403.

Zhao, X.-Q., Hu, J.-F., & Yu, J. (2006). *Comparative Analysis of Eubacterial DNA Polymerase III Alpha Subunits*. *Genomics, Proteomics & Bioinformatics*, 4(4), 203–211. doi:10.1016/s1672-0229(07)60001-1

Zhao,X.Q., Hu,J.F. and Yu,J. (2006) *Comparative analysis of eubacterial DNA polymerase III alpha subunits*. *Genom. Proteom. Bioinf.*, 4, 203–211.