



“REVISIÓN BIBLIOGRÁFICA DEL FENÓMENO DE DUPLICACIÓN GENÉTICA RECIENTE EN ENTEROBACTERIAS”

*Dpto. de Desarrollo Biotecnológico, Instituto de Higiene,
Facultad de Medicina, UdelaR.*

Autores: Valentina Ximénez

Tania Rodríguez

Milagros Vital

Rodrigo Tais

Yuliana Sanguinetti

Tutor: Andrés Iriarte

Metodología Científica II, año 2015

ÍNDICE

Resumen.....	Página 3
Introducción.....	Página 3
Justificación.....	Página 6
Profundización Teórica.....	Página 6
Bloque Experimental.....	Página 10
Duplicación y Resistencia a Antibióticos.....	Página 16
Transferencia Horizontal de Genes y Duplicación Génica....	Página 18
Análisis de Inparálogos en Genomas de Salmonella.....	Página 21
Conclusiones.....	Página 25
Referencias.....	Página 26

RESUMEN

La duplicación genética en bacterias no es el proceso más frecuente para la adquisición de nuevo material genético, pero es más relevante de lo que una vez se pensó, y contribuye positivamente a la adaptación del microorganismo a los cambios ambientales

Un gen Parálogo es aquel que se duplica a partir de otro gen preexistente; y cuando ocurre dentro del mismo genoma o especie, el producto de la duplicación recibe el nombre de Inparálogo. Miles de estos genes fueron identificados en un importante número de procariotas gracias a la secuenciación de genomas completos. Estos fueron asociados con la producción de energía, la motilidad celular, el transporte de iones y mecanismos de defensa, entre otras funciones importantes para la supervivencia y la proliferación de organismos específicos.

A pesar de que el genoma en procariotas está constantemente cambiando, el tamaño relativo del mismo, y por ende el número de genes, es constante. Esto se debe a que la proporción de genes por cantidad de ADN es altamente conservada.

Las familias más grandes contienen genes de transporte y genes de regulación, mientras que las familias más pequeñas están involucradas en el metabolismo y la producción de energía. Los Parálogos se correlacionan bien con el tamaño del genoma, cuanto más grande sea este, más Parálogos tendrá.

Se revisaron trabajos experimentales para obtener información que sustente las múltiples teorías que se postulan en torno a la función de la duplicación en los organismos procariotas. La duplicación génica, al generar amplificación, confiere resistencia antibiótica por provocar la sobreproducción de enzimas, la modificación de moléculas diana y bombas de eflujo.

La transferencia horizontal, el mecanismo más utilizado por las bacterias para incrementar su material genético, es un proceso de gran importancia evolutiva, en el cual bacterias intercambian material genético y sufren recombinaciones posteriores.

INTRODUCCIÓN

Los organismos procariotas pueden sobrevivir en los más diversos escenarios ecológicos del planeta. Su rápida adaptación es, en parte, el resultado de la capacidad de adquirir nueva información genética; es decir nuevas funciones (1). Estos microorganismos tienen cantidades de genes altamente variables, desde poco más de 100 a 10.000 genes. Tales variaciones en el tamaño del genoma se asocian típicamente con expansiones y contracciones de las familias de

genes. Estas variaciones estarían asociadas con la adquisición de nuevas funciones, nuevas estructuras y sistemas de regulación (2).

La composición del genoma es moldeada durante la evolución mediante la aparición de nuevos genes por duplicación, la transferencia horizontal y la pérdida de los mismos (2). Esto significa que los procariotas utilizan dos principales caminos para ampliar su repertorio de familias de proteínas: adquirir genes por transferencia horizontal en cuyo caso se llaman Xenólogos; o se pueden duplicar a partir de un gen pre-existente, en el cual los genes homólogos se llaman Parálogos, específicamente Imparálogos. Entendiendo a la duplicación génica como un fenómeno de mutación, se genera la copia de un gen que luego tendrá la posibilidad de diferenciarse en cuanto a la función original y que se hereda en forma vertical a las células hijas; si no existe selección directa, las copias adicionales redundantes se pierden rápidamente. La ganancia y la pérdida de genes en procariotas son procesos equilibrados; así, en promedio, el tamaño del genoma se mantiene constante (1). Si bien en eucariotas la duplicación es el mecanismo más importante de adquisición de genes, en procariotas se cree que el principal responsable es la transferencia horizontal; aunque la duplicación es un proceso esencial para la conservación y evolución de genoma y también juega un rol en la determinación de las características fenotípicas de los microorganismos (3).

Los genes homólogos son clasificados según su origen: los Homólogos-Ortólogos u Ortólogos están relacionados a través de la descendencia vertical a partir de un ancestro común; los Homólogos-Parálogos o Parálogos simplemente, se relacionan a través de acontecimientos de duplicación posteriores a la especiación. Los Parálogos se definen a menudo como secuencias de codificación de proteínas que tienen al menos 30% de identidad de secuencia en más de 60% de su longitud (cobertura) (4, 5), aunque estos criterios varían entre los distintos trabajos. A su vez los Parálogos son clasificados en Inparálogos (intra parálogo) los cuales corresponden a genes duplicados dentro de la misma especie o genoma; y en Outparálogos (extra parálogos): genes homólogos - parálogos que en su origen surgieron como producto de una duplicación en un linaje ancestral y se encuentran en genomas distintos; es decir que ocurrió antes del evento de separación de linajes actuales. Otras subdivisiones dentro de Parálogos se han generado, como por ejemplo, Coortólogos y Ecoparálogos (6).

Miles de genes Parálogos fueron identificados en un importante número de procariotas gracias a la secuenciación de genomas completos. Estos fueron asociados con la producción de energía, la motilidad celular, el transporte de iones y mecanismos de defensa, entre otras funciones. Los Parálogos a menudo se asocian con propiedades que son importantes para la supervivencia y la

proliferación de los organismos específicos. Si bien estas definiciones actualmente son claras, antes no estaban tan uniformemente usadas por los investigadores (7).

Las familias de genes surgen durante la historia evolutiva a través de un proceso de duplicación de un gen ancestral seguido de divergencia aleatoria o de una especialización funcional y estructural (divergencia dirigida por selección) en una o ambas copias.

Las duplicaciones genéticas en tándem en bacterias y bacteriófagos ocurren espontáneamente a una frecuencia de 10^{-3} a 10^{-5} por locus por generación, y puede ser de tamaño ilimitado (8). Estas duplicaciones espontáneas también pueden ocurrir por otros mecanismos, no en tándem, como por ejemplo: duplicación cromosómica, disyunción, translocación y retrotransposición, entre otros. Claro que la frecuencia de cada uno de ellos en procariotas es significativamente distinta.

En general las duplicaciones espontáneas en genes se producen con relativa frecuencia; por lo tanto en el proceso evolutivo de la duplicación de genes y la divergencia, el paso limitante de la velocidad se produce probablemente en el nivel de la selección natural después de la duplicación. Por esto, un gen duplicado es probable que se pierda a menos que adquiera una función nueva y útil. La divergencia es el resultado de las diferencias genéticas producidas por distintos eventos en la evolución, es decir la de diferencias en la secuencias producto de mutaciones independientes acumuladas con paso del tiempo..

El fenotipo resultante de una mutación por duplicación puede ser debido al aumento de la dosis génica, que por lo general causa una ganancia de función o de las consecuencias funcionales de un nuevo elemento.

La amplificación es usualmente un proceso de dos pasos en el cual la duplicación genética (paso uno) promueve un aumento adicional en el número de copias (paso dos). Ambos pasos tienen importancia evolutiva para la aparición de funciones genéticas innovadoras.

Las propiedades genéticas básicas de duplicaciones en bacterias se exponen por primera vez por Campbell (1963- 1965), utilizando la duplicación de la región Gal. Estos y otros trabajos sobre los principios de la genética de duplicaciones han sido revisados por Anderson y Roth (1977) y más recientemente por Petes y Hill (1988) y por Roth et. al (9).

El objetivo del presente trabajo es revisar el impacto del fenómeno de duplicación en las características fenotípicas de los microorganismos, con especial interés en las ventajas biológicas que esto podría determinar y las consecuentes repercusiones clínicas y

epidemiológicas. La lectura se centrará en la duplicación de genes en bacterias, más precisamente en Enterobacterias, dada su importancia Médica.

La revisión se dividirá en las siguientes secciones: i) profundización teórica; ii) documentación de trabajos experimentales, que evidencian el proceso en estudio y sus propiedades; iii) duplicación y la resistencia a antibióticos; iv) transferencia horizontal y duplicación, contraposición de mecanismos generadores de nuevos genes en los genomas procariotas.

En paralelo se incluirá una sección que presenta el análisis de la duplicación de genes recientes o “Inparálogos” en el género *Salmonella*, identificando qué genes son los que más frecuentemente se duplican; para evaluar de esta forma si existe un mecanismo de respuesta al entorno o si los genes se duplican de forma aleatoria de acuerdo a su función. Al mismo tiempo se analizará si la frecuencia de duplicación está determinada por la importancia funcional de dichos genes, sin tener en cuenta el costo energético para la célula; o si únicamente se duplican genes con secuencias cortas, lo cual implica que el gasto energético sea menor. Dicho análisis se hará sobre una base de datos presente en el Instituto de Higiene, Montevideo.

JUSTIFICACION

Debido a que la duplicación génica en procariotas no es considerado un proceso tan relevante como la transferencia horizontal de genes, se profundizará sobre este mecanismo genético a través de una revisión bibliográfica con el fin de determinar la relevancia que tiene este proceso en fenómenos como la resistencia a antibióticos, la supervivencia de las bacterias y la adaptación a entornos, entre otros.

Esto permitirá estimar la importancia de dicho proceso, lo que podría servir como inicio para nuevas investigaciones en la esfera Bio-Médica que permitan conocer en profundidad el fenómeno de duplicación y con la hipótesis de desarrollar nuevas estrategias respecto al control de la patogenicidad de estos microorganismos.

PROFUNDIZACIÓN TEÓRICA

Como ya hemos visto, la duplicación de genes es un proceso evolutivo normal, que junto con la pérdida de genes y la transferencia horizontal se consideran importantes procesos que dan forma a los genomas de procariotas (3). A pesar de estos procesos, que implican cambios constantemente a nivel del genoma en procariotas, el tamaño relativo del mismo, y por ende el número de genes, se mantiene constante. Esto se debe a la proporción de genes por cantidad de ADN, la cual es altamente conservada y en promedio 1kb por gen (2).

Amplificación génica

En términos generales la amplificación genética, tiene importancia en cuanto a la perspectiva evolutiva, (10) (11) y contribuye en genética médica a muchas enfermedades humanas, así como a la variabilidad fenotípica entre los individuos y la susceptibilidad humana a las enfermedades infecciosas. (12) (13) Además, la amplificación de genes juega una parte importante en la generación de la variabilidad genómica, que está disponible para la adaptación genética frente a condiciones de crecimiento alterados. Por lo tanto la amplificación, luego de la duplicación de genes, permite a las células adaptarse a un nuevo entorno. (14)

Se han realizado varios estudios acerca de las duplicaciones genéticas, especialmente en *Salmonella* Entérica serotipo Typhimurium LT2 y en *E. Coli*. El método que se utiliza actualmente es la PCR en tiempo real, ya que es la manera de medir el número de copias de una región genómica específica de interés (15) o directamente la secuenciación del genoma (14).

De acuerdo a estos estudios, duplicaciones genéticas en tándem se han encontrado en todas las regiones del cromosoma (16), en tamaños que van desde unas pocas kilobases a varias megabases (17-19). Se estima que aproximadamente un 10% de las células bacterianas en un cultivo en crecimiento contiene una duplicación de genes en algún lugar de su genoma (Ver “Bloque Experimental”) (20).

En cuanto al mecanismo de duplicación, existen al menos dos:

Mecanismo RecA- dependiente, en el cual duplicaciones surgen por recombinación homóloga no iguales, entre largas secuencias de repetición directamente orientadas, tales como operones ribosómicos ARN, (16, 21) secuencias de inserción o elementos transponible (22, 23) y secuencias repetitivas tales como secuencias palindrómicas extragénicas (24,25) que están presentes en las cromátidas hermanas.

Mecanismo RecA-independiente; en este caso las duplicaciones carecen de secuencias repetidas en el punto de unión a la duplicación, (14) aunque a través de esta forma independiente, se han implicado secuencias repetitivas cortas (<20pb), principalmente asociadas a la formación de duplicaciones genéticas, eliminaciones y recombinación entre homologías de secuencia corta (26, 27). Otro mecanismo, es la replicación de círculo rodante (RCR), en el cual la amplificación puede producirse sin ningún intermedio de duplicación. (20, 28) Este mecanismo, es común en plásmidos y bacteriófagos, y permite la síntesis rápida de múltiples copias de un

genoma o región específica, pudiendo generar una gran disposición en tándem en una sola generación (11, 18, 29).

En cuanto al tamaño de las unidades amplificadas, el mismo varía desde, unas pocas kilobases hasta varias megabases, como se muestra, a través de estudios en *S. Typhimurium* y *E. Coli*, en los que aumentó la dosis de genes, a través de DGA (duplicación de genes y amplificación) cuando fueron seleccionados. (18, 30-32)

Las duplicaciones en tándem y amplificaciones son intrínsecamente inestables, ya que tienen tasas de pérdidas muy altas por cada generación de células, esto se debe a la recombinación homóloga, que ocurre entre las unidades amplificadas (20, 33). Estas duplicaciones y amplificaciones pueden estabilizarse por efecto de la selección, la cual va a enriquecer continuamente la población de células que mantienen la duplicación promotora del crecimiento o de amplificación (18, 34-39); obteniendo como resultado el aumento de la actividad del gen (40). Por lo antedicho, podemos asumir que las duplicaciones de genes y la amplificación, podrían servir como un intermedio y facilitador en el desarrollo de una adaptación genética estable, (14) siempre que exista dicha selección.

Parálogos y su divergencia

En todos los genomas, la presencia de familias y superfamilias de proteínas sugiere una historia de duplicación de genes y divergencia en la evolución (7). Dentro de estas familias, las más grandes contienen genes de transporte y de regulación, mientras que las más pequeñas están involucradas en el metabolismo y la producción de energía (3). En lo que respecta a vías metabólicas, las mismas pueden ser utilizadas por muchos organismos, por ejemplo, para la síntesis de ácidos grasos; pero hay vías que están restringidas solo a organismos entéricos, tales como la emulsificación de ácidos biliares, biosíntesis de ácido colánico, lípido A, enterobactina y antígeno común enterobacteriano (41).

A través de diversos estudios, se ha analizado la existencia de familias parálogas, es decir familias de secuencias relacionadas dentro de su genoma. Dentro de las enterobacterias, en *E.Coli* más de la mitad de sus proteínas, caen en el grupo de secuencias relacionadas (41-45). Como ejemplo, en *E.Coli K-12*, el 30% de sus proteínas tienen al menos una secuencia paróloga en su genoma. (5) La existencia de estas familias de proteínas, de secuencia similar, sugiere que la duplicación y posterior divergencia, son mecanismos importantes de la evolución molecular, presentes también en procariotas. (41)

Identificar las características que están sobrerrepresentadas en parálogos, permite evaluar como la composición del genoma de los procariotas responde a las presiones del medio ambiente, por ejemplo, adquiriendo, por transferencia horizontal o duplicación, genes vinculados a la locomoción, la quimiotaxis, el transporte de iones o la fotosíntesis, permitiendo así la adaptación a un nuevo entorno. (7) (Ver “Bloque experimental)

Particularmente en la especie *E. Coli*, los genes sobrerrepresentados fueron los implicados en el transporte de metabolitos, seguidos de los de transcripción y replicación/repación. Estas características, pueden contribuir como una ventaja competitiva para el organismo, en el caso de transporte de iones y el metabolismo, al tener más copias de genes en esta categoría, por ejemplo pueden aumentar más la absorción de metales (cobre y hierro), los cuales son importantes para la supervivencia (7).

Un nuevo concepto: ecoparálogos

La duplicación de genes y posterior divergencia son un importante contribuyente a la génesis de genes (73), sin embargo existe un proceso evolutivo con poca divergencia que conduce a nuevos genes con funciones redundantes, a ecoparálogos. Los ecoparálogos son genes relacionados que realizan una misma función celular, bajo diferentes condiciones ecológicas, esto permite que las bacterias puedan adaptarse a condiciones ambientales cambiantes (por ejemplo, salinidad o temperatura). *Salinibacter Ruber* es un ejemplo de organismo con ecoparálogo, siendo la primera bacteria halófila secuenciada (74). Las proteínas típicas halófilas acumulan residuos ácidos en su superficie y muestran su óptima actividad y estabilidad en alta salinidad; en consecuencia no es sorprendente que *S. Ruber* presente el proteoma más ácido (punto isoelectrico -pI- bajo) entre las bacterias (74).

Los ecoparálogos, difieren poco en secuencia, ya que la función de sus proteínas codificadas es idéntica, sin embargo frente a cualquier variación, deben conferir un carácter ecológico diferente (75). La redundancia genética se relaciona a la robustez de los organismos, ya que mantienen un fenotipo estable frente a cambios ambientales, o bajo la forma hereditaria, por ejemplo, frente a mutaciones genéticas (76). Aunque no está claro cuál de estos factores generan el efecto amortiguador de los genes redundantes (77, 78), el alto número de secuencias parálogas en bacterias que habitan en varios nichos (3) y la presencia de ecoparálogos en respuesta a los cambios externos, apuntan hacia el medio ambiente como un factor de vital importancia en mantener la redundancia en los genomas procariotas.

Mediante tres procesos selectivos se lograría dicha redundancia (75):

- 1- Dosis altas de proteínas, dando lugar a genes duplicados idénticos.
- 2- Diversificación de proteínas, dando lugar a parálogos divergentes.
- 3- Adaptación a variaciones ambientales, dando lugar a ecoparálogos de divergencia intermedia.

Todos los procesos mencionados anteriormente, contribuyen a la adaptación en bacterias (79). A pesar de que muchos genes duplicados experimentan aceleración en la tasa de la evolución, la contribución relativa de la selección positiva y la deriva al azar a la retención y posterior evolución de genes duplicados, aún no está claro. (80)

BLOQUE EXPERIMENTAL

El estudio de la duplicación en procariotas ha sido un tema de interés que motivó varios trabajos experimentales en el área con el objetivo de comprender el origen y mecanismo de este proceso. La motivación de varios de los trabajos que recorreremos en este bloque es la búsqueda de evidencia empírica que sustente las múltiples teorías que se postulan en torno a la función de la duplicación en los organismos procariotas.

La primera idea elaborada en torno a este tema provino del genetista Sumusu Ohno cuando planteó en 1970 que “el surgimiento de nuevos genes ocurre durante la división celular produciendo una copia adicional de un gen existente, y que esta copia extra es libre de mutar y adquirir nuevas funciones”. Esta simple pero elegante idea abrió el camino en el estudio de este fenómeno donde cada experimento realizado formula nuevas interrogantes; ¿Qué ocurre con un gen una vez duplicado?, ¿Existe un criterio para la duplicación genética?, ¿Qué mecanismos están implicados en la selección de la duplicación?, ¿Puede un organismo duplicar su material genético con una tasa constante?...

Uno de los agregados más importante al modelo propuesto por Ohno proviene del trabajo experimental realizado por Näsval y Anderson publicado en el año 2012 en la revista Science (81). Dichos investigadores propusieron un modelo llamado modelo de innovación-amplificación-divergencia. Mediante este modelo los investigadores plantean que “el mejor candidato para la duplicación de genes es un gen cuyo único producto de proteína no sólo lleva a cabo su trabajo, sino que también podría llevar a cabo una función secundaria no esencial” (81). A partir de que esta segunda función cobra relevancia para el organismo, el gen duplicado se mantendría por un proceso de selección positiva resultando finalmente en la divergencia de ambas copias génicas hacia las dos funciones propuestas.

Para demostrar este modelo Näsval y Anderson realizaron un estudio en tiempo real en una población de *Salmonella*. Los investigadores centraron su estudio en el comportamiento de la población de *Salmonella* en relación a un medio carente de los aminoácidos histidina y triptófano. Para probar su modelo seleccionaron una *Salmonella* que poseía una mutación espontánea del gen que codifica la enzima biosintética de histidina (HisA) que al mismo tiempo había adquirido un bajo nivel de actividad TrpF suficiente para apoyar el crecimiento lento en un medio que carece de histidina y triptófano. Dicho gen bifuncional fue colocado bajo el control de un promotor constitutivo que co-transcribe fluoruro amarillo como marcador de la duplicación y se localizó cerca del operón lac del plásmido F`128 para asegurar su duplicación a un bajo costo energético.

Para probar el modelo se introdujo el plásmido en una cepa de *Salmonella Entérica* (S. Entérica) carente de los genes codificantes de las enzimas encargadas de la síntesis de HisA y TrpF en un medio carente de dichos aminoácidos y se siguió el comportamiento de las sucesivas generaciones así como el tiempo de replicación de cada una.

Los resultados del estudio arrojaron datos muy interesantes que apoyan el modelo propuesto por los investigadores. En primer lugar al estudiar el tiempo necesario para el crecimiento de las poblaciones encontraron que, dentro de unos pocos cientos de generaciones, la tasa de crecimiento se incrementó de 5 horas por división a 1.9 a 2.5 horas, dependiendo del linaje (81). Sumado a este fenómeno se observó el aumento gradual de la expresión del gen bifuncional debido a la amplificación en la mayoría de los cultivos (81). Esto se explica, acorde a la interpretación de los autores, por la necesidad de varios linajes independientes de esta cepa de evolucionar bajo la selección continua para mejorar el crecimiento y el aumento de las actividades de HisA y TrpF.

Otro de los hallazgos importantes para el modelo está relacionado con el hecho de que después de la evolución en más de 3.000 generaciones, todos los linajes adquirieron mutaciones resultantes en un crecimiento más rápido (81). Al analizar los diferentes linajes fueron puestas en evidencia 22 variantes del gen mutado bifuncional original; estas variantes pueden clasificarse en tres grandes grupos: los genes especializados con fuerte actividad mejorada HisA y la pérdida de actividad TrpF, los genes especializados con fuerte mejora de la actividad TrpF y pérdida de actividad HisA, y genes no-especializados cuya enzima codificada mostró un aumento moderado en ambas actividades. De esta manera cada una de las fases del modelo de innovación, amplificación y divergencia (IAD) fueron respaldadas por la evidencia obtenida en el estudio permitiendo que el modelo IAD sea considerado como uno de los aportes más

importantes a la teoría inicial de Ohno y permita continuar ahondando en los mecanismos que guían la duplicación de genes en procariotas.

En el año 2010 Marit S Bratlie y Jostein Johansen realizaron un estudio para comprobar de manera experimental la importancia de los ecoparálogos en la adaptación al medio de los organismos procariotas (7). Para ello los investigadores utilizaron la secuencia de genomas procariotas de NCBI (National Center for Biotechnology Information), mediante análisis de secuencia, utilizando todos los genomas procariotas completamente secuenciados (arqueas y bacterias) disponibles en el Gen Bank (Banco de Genes), centrándose en parálogos recientemente surgidos ó bien conservados (7), en la cual encontraron parálogos en 897 procariotas, y seleccionaron los 200 organismos que presentaban una mayor fracción de parálogos. En el análisis, se vio que el organismo con fracción paráloga más alta fue *Aster Yellow Fitoplasmas* (cepa AYWB), con un 12,12%, mientras que el organismo con la fracción paráloga más baja fue *Streptococcus Pyogenes* (cepa MGAS8232) en un 2,28%.

Para asegurar que los parálogos a estudiar no correspondían a genes con neo-funcionalización, sub-funcionalización o genes obtenidos por transferencia horizontal, los investigadores tomaron criterios estrictos en la homología de la secuencia (> 75 % de homología y 65% de cobertura) (7). Una vez seleccionados los genes duplicados estos fueron analizados mediante el trazado del número de genes con parálogos, versus la longitud del genoma, y su distribución dentro de las categorías COG (agrupaciones de grupos ortólogos) (7). Luego mediante el análisis de enriquecimiento diferencial con el programa web DAVID se identificaron los términos funcionales sobrerrepresentados, que posteriormente se estudiaron a través de términos de ontología de genes (GO). El análisis GO obtiene una información más unificada, estableciendo una categoría para los genes en base a sus funciones en eucariotas y procariotas. De estos resultados se extrajeron la clasificación de dichos genes en clusters. Por último se comparó la similitud obtenida por términos GO versus la distancia genómica para terminar de concluir si los genes estudiados fueron en efecto duplicados en una escala de la evolución previa o si responden a una duplicación inducida por las condiciones ambientales.

En primer lugar se encontró que el número de parálogos está relacionado lineal y positivamente con el tamaño del genoma, $r=0.8$ (7). Por otro lado al analizar las categorías COG a las que pertenecen los parálogos, versus aquellos genes que no se encuentran duplicados se encontraron las siguientes diferencias: los parálogos fueron preferentemente asociados con la conversión y producción de energía (COG categoría C), la motilidad celular (N), el metabolismo y transporte de iones inorgánicos (P), la traducción de señales (T), y los mecanismos de defensa (V). Los

genes no duplicados, por otra parte, fueron relacionados con el control del ciclo celular y la mitosis (D), el metabolismo y transporte de nucleótidos (F), modificaciones post-traduccionales (O). Para ambos grupos la replicación y reparación (L) es importante del mismo modo que la transcripción (K) (7). De esta manera los investigadores concluyen que la duplicación es un proceso que está bajo estricto control, preservando funciones fundamentales para el organismo. Al mismo tiempo, las categorías en las que fueron asociados los parálogos, permiten asumir que estos están relacionados con mecanismos de adaptación.

Del mismo modo mediante el “biclustering” (agrupación de conjuntos de datos en dos dimensiones simultáneamente) de los genes sobrerrepresentados por los términos GO se identificaron 104 clusters, muchos correspondientes a categorías generales, aunque se diferencian 6 clusters que pueden relacionarse de mejor manera con la adaptación ambiental. Dichos clusters fueron asociados con: motilidad celular, transporte de iones, dominios de GTPasas, glicósido hidrolasas, ligantes de iones metálicos y metiltransferasas.

Cluster 1 - Unión de iones metálicos: este grupo incluye principalmente términos GO implicados en la unión de cobre. Los organismos dependientes de cobre, tiene un desafío en mantener niveles adecuados de cobre, ya que niveles altos a nivel intracelular, por ejemplo, podrían ser tóxicos para el organismo (51).

* *Dinoroseobacter shibae* (cepa *DFL 12*), con sus proteínas multicobre oxidasa 3 y un regulador transcripcional de la familia Merr

* *Mycobacterium gilvum* con las proteínas citocromo c oxidasa y proteína de resistencia de cobre COPC.

* *N. hamburgensis*, con proteínas citocromo c oxidasa.

Las proteínas de la familia Merr detectan reguladores que se unen a metales y activan la transcripción de las proteínas implicadas en la desintoxicación de iones de metal (52), lo cual hace pensar de que la bacteria (*D. shibae*- cepa *DFL 12*) conserva múltiples parálogos implicados en la unión para la captación y regulación más eficiente de cobre. La proteína COPC, se encarga de la resistencia al cobre, por lo tanto puede mejorar la supervivencia de la bacteria bajo estrés extremo de cobre, como es el caso de la proteína COPC en *E. coli* (53).

Cluster 2- motilidad celular: todos los organismos en este grupo tienen flagelos, que además de la motilidad, juega un papel clave en la expresión génica de muchas bacterias, además podría

participar en la adherencia a las células huésped, invasión huésped celular, formación de biopelículas y la secreción de proteínas (54).

* *Agrobacterium Tumefaciens* (cepa mutante flagelo-menos) el mismo muestra una virulencia atenuada.

* *Yersinia*, los flagelos en esta bacteria actuarían como un sistema de secreción de un factor de virulencia. (55)

* *Xanthomonas Oryzae*, causa enfermedad vascular, enfermedad de la plaga bacteriana del arroz. Estudios han demostrado que esta bacteria se mueve hacia exudados de plantas susceptibles de arroz, mientras que no se produce la quimiotaxis hacia exudados en las plantas de arroz resistentes. Esto implica que la quimiotaxis juega un rol importante en la patogenicidad de las bacterias antes de la penetración en la hoja de arroz. También la mutación en uno de los ORFs de la región operón flagelar en *X. Oryzae* conduce a la quimiotaxis débil. (55, 56)

Cluster 3- glucósido hidrolasa: este grupo incluye lisozimas, hialuronidasas, chinitases, esterases y xilosidasas (57). La lisozima está involucrada en defensa antibacteriana, y la neuraminidasa en mecanismos patogénicos, ambas están implicadas en la degradación de celulosa y la hemicelulosa.

* *Thermophilum Anaerocellum* (cepa DSM 6725), la cual utiliza una variedad de polisacáridos como celulosa cristalina y la hemicelulosa (58).

* *Streptomicetos*, es una de las bacterias que viven en el suelo, y degradan restos insolubles de otros organismos de otros organismos como la quitina y lignocelulosa (59).

* Enterobacterias: *Sodalis Glossinidus* (cepa *Morsitans*) y las dos cepas de *E. coli* O157: H7. Ambas representadas con las proteínas implicadas en la degradación de la pared celular. *E. coli* O157: H7 es una enterobacteria hemorrágica que causa colitis hemorrágica (60) mientras que *S. Glossinidus*, un endosimbionte transmitido por vía materna, se encuentran para infectar a una amplia gama de tejidos en las moscas tsé-tsé, el vector que causa la enfermedad del sueño en los seres humanos (61).

* *Streptococcus Pyogenes* pertenecen al grupo A estreptococo (SGA) causando una variedad de enfermedades en los seres humanos. Las bacterias se cubren con una cápsula externa hecha de ácido hialurónico para evitar la fagocitosis y para facilitar la adherencia a las células epiteliales

(62). La enzima hialuronidasa reduce la viscosidad de ácido hialurónico, aumentando así la permeabilidad del tejido. Según Starr y Engleberg (62), la hialuronidasa puede jugar un papel nutricional para el organismo bajo las condiciones de escasez de nutrientes, por lo cual es posible utilizar dicha enzima o su propia cápsula como una fuente de energía de carbono.

Cluster 4- Metil transferasas:

* *Orientia Tsutsugamushi* (cepa *Ikeda*), con proteínas ADN-metiltransferasa-adenina específica N6.

* *Orientia Tsutsugamushi* (cepa *Boryong e Ikeda*), con proteínas adenina metilasa de ADN de sitio específica.

* *B. Tribocorum*, sus genes se clasifican como helicasa /metiltransferasa, fago relacionado modificación metilasa y como tipo III de sistema de restricción metilasa.

* *Cyanothece* (cepa PCC 7424), se encuentran en el, varias familias de transposases.

* *Frankia sp.* (cepa *EAN1 pec*), con ADN (citocina-5) metiltransferasa y transposases.

La metilación de ADN en bacterias, mediante metiltransferasas hace que éste se pueda separar a partir de ADN extraño, dando a estas bacterias una protección frente a invasiones, como los bacteriófagos. La metilación de adenina o citosina es una parte del sistema de modificación de restricción en bacterias. La metilación de adenina es uno de los mejores mecanismos epigenéticos, que consta de regulación del ciclo celular (63). Sin embargo, hay metiltransferasas que no son parte de este sistema de modificación de restricción, como Presa (ADN metilasa adenina) en gammaproteobacteria y CCRM (ciclo celular regulado metiltransferasa) en alphaproteobacteria (7). Mutaciones en CCRM es letal en *A. tumefaciens*, *Brucella Abortus*, *Caulobacter Crescentus*, y *Rhizobium Meliloti*. Estas metiltransferasas son parte del sistema de regulación en una célula, incluyendo el control de la virulencia bacteriana. Mutantes dam (enzima ADN adenina metiltransferasa) de *Salmonella entérica*, *Caenorhabditis Elegans*, *Haemophilus Influenzae*, *Yersinia Pseudotuberculosis* e *Y. Pestis* tienen virulencia atenuada (64, 65).

Cluster 5- transporte de iones: * *E.Coli K-12*, incluye enterobactina, transportadores ABC de nitrato/nitrito, esta cepa se encuentra en el intestino inferior de animales y se caracteriza por sobrevivir en el medio ambiente natural (5).

* *Ralstonia Metallidurans* (cepa CH34), presenta proteínas que actúan como bombas de flujo de salida de metales pesados, proteínas de resistencia a metales pesados, proteínas de transporte de mercurio, Hg (II) proteína de resistencia MERP y bacterioferritina. La resistencia a metales pesados caracteriza a esta bacteria (66).

* *Methanosarcina Barkeri* (cepa Fusaro), es un archaea que fija nitrógeno y ha sido aislada de muestras de barro y rumen de ganado (67, 68). En esta cepa se encuentran proteínas transportadoras de molibdeno, transportadores ABC de hierro (III), y las proteínas de unión de iones de mercurio.

Tanto el molibdeno como el hierro son esenciales para la mayoría de los organismos vivos. La disponibilidad del hierro esta limitada, debido a su baja solubilidad, por lo que los transportadores de hierro (III) ABC hacen posible vivir donde la disponibilidad de hierro es baja. Un estudio realizado por Flo et al. (69) ha demostrado que el hierro es secuestrado por lipocalina 2 como una respuesta inmune durante la infección, que posteriormente limita el crecimiento bacteriano. Para adquirir hierro, los microorganismos liberan sideróforos. Enterobactina es uno de los sideróforos de unión a Fe^{3+} más fuertes que se conocen hoy en día (70). La duplicación de genes en enterobactina hace posible que *E. coli* K-12 colonice sitios pobres en hierro como lo es el intestino inferior. Por otra parte el molibdeno es el metal de transición mas abundante en agua de mar (71).

Cluster 6- dominio GTPasa: La mayor parte de las proteínas de este grupo contienen un dominio GTPasa como-Ras, y por lo tanto son anotados por procedimientos automáticos como funcionalmente relacionados. Sin embargo, el dominio GTPasa Ras se conserva en muchas proteínas diferentes (72), y las mismas no se encuentran necesariamente en las mismas vías.

Gracias a la evidencia arrojada por el estudio los autores concluyen que los parálogos están relacionados con estilos de vida individuales de procariontes y son propensos a dar una ventaja competitiva para el organismo. Parálogos y genes no duplicados dominan diferentes categorías de clasificación funcional, donde los parálogos en particular, parecen asociarse con los procesos que implican la interacción con el entorno (7).

DUPLICACIÓN Y LA RESISTENCIA A ANTIBIÓTICOS

Es ampliamente conocido el rol de las mutaciones o de la transferencia horizontal en la adaptación de las bacterias a un medio hostil, tal como lo es la presencia de niveles tóxicos de antibióticos (82, 83) estos procesos no son los únicos que contribuyen a esta causa, la

duplicación de genes viene siendo considerada hace mucho tiempo una importante fuente de potencial genético adaptativo (Ohno en 1970, Zhang en 2003, Dittmar and Liberles en 2010). Se ha demostrado que la misma, al generar amplificación genética, confiere resistencia al provocar la sobreproducción de enzimas, la modificación de moléculas diana y bombas de eflujo. Los primeros ejemplos de aumento de los niveles de resistencia a los antibióticos, resultante del aumento del número de copias de genes, se encontraron en la década de 1970. En ese año se evidencio que diferentes plásmidos R (plásmidos que llevan los genes que confieren resistencia a los antibióticos) aumentan en tamaño cuando se cultivaron en presencia de antibióticos. En la mayoría de los casos en los que la resistencia a antibióticos se confiere por la amplificación del gen, el aumento de la resistencia está ligado directamente a un aumento de los niveles del producto génico del gen amplificado.

Sin embargo, debido a la alta frecuencia con la que las amplificaciones se producen en una población, estas duplicaciones en tándem pueden ser el primer paso hacia el aumento de la resistencia, que se sigue posteriormente por mutaciones puntuales menos frecuentes en el gen amplificado. Es decir la probabilidad de que esto ocurra se incrementa por el aumento del número de copias del gen diana (14).

La duplicación y amplificación genética son considerados mecanismos adaptativos en bacterias (84) porque permiten al organismo habituarse al cambio ambiental modificando la resistencia antibiótica, la virulencia de los patógenos, la eficacia de las vacunas y la biosíntesis de los antibióticos.

Recientemente los análisis de genomas secuenciados muestran como la plasticidad del mismo puede afectar la evolución y las propiedades inmunogenéticas de las bacterias patógenas. Por ejemplo, al amplificarse los genes *acrAB* estos pueden aumentar la resistencia a fármacos en *E. Coli* (85).

La duplicación en *Streptococo Pneumoniae* afecta la producción de la cápsula y la resistencia a macrólidos. La duplicación espontánea controla la variación de fase capsular durante las infecciones neumocócicas. Esta cápsula además de conferir resistencia al complemento mediante opsonización y fagocitosis, inhibe la adhesión y la invasión de algunas células eucariotas. La capacidad de cambiar de una forma capsular a una forma sin cápsula puede facilitar la transición bacteriana hacia la enfermedad invasiva. En un aislado clínico serotipo 3 de *S. Pneumoniae*, la duplicación provoca resistencia a los macrólidos. Los macrólidos inhiben la traducción bacteriana mediante la unión a la subunidad ribosomal 23S. La duplicación inserta seis aminoácidos adyacentes al sitio de unión del macrólido y confiere resistencia a una variedad de diferentes antibióticos de este grupo (85).

La caracterización genómica de *Borrelia Burgdorferi* (la bacteria causante de la enfermedad de Lyme), reveló que un gen (cspA) involucrado en la evasión de las defensas antibacterianas del huésped está comprendido dentro de una disposición en tándem de parálogos pertenecientes a la familia PFam54 (86).

Un estudio realizado con el patógeno oportunista *Streptococo Agalactiae* identificó dos cadenas con amplificación genética, una con cuatro veces más amplificaciones en tándem en una región de 13500 y otra de 92000 pares de bases. La primera incluía 5 genes necesarios para la síntesis de dihydrofolato. Al parecer ese incremento en la dotación génica le confirió a este patógeno resistencia a sulfonamida y a Trimetoprim por dos mecanismos diferentes. La resistencia a sulfonamida fue atribuida a los altos niveles de un gen único en la secuencia, cuyos productos proteicos son el objetivo de la droga. Por el contrario, la resistencia a Trimetoprim apareció como resultado de un nuevo mecanismo a través del cual la alta expresión de toda la secuencia aumentó los niveles intracelulares de dihydrofolato. Al ser el Trimetoprim un inhibidor competitivo de la enzima que actúa sobre el dihydrofolato, las cantidades aumentadas de este sustrato metabólico podrían ser suficientes para vencer los efectos inhibitorios de la droga (87).

Estudios en el organismo *Bacilus Subtilis* han reportado que la amplificación del locus tmrA resulta en resistencia a tunicamycin e hiperproducción de alfa amilasa. Así mismo se demostró que la multiplicación del locus del gen de resistencia tetB es uno de los factores que confiere resistencia a *B. Subtilis* en altas concentraciones de tetraciclina (88).

El aumento de la resistencia a ampicilina en *Yersinia Enterolítica* , la resistencia a múltiples antibióticos en *Salmonella Entérica* y el aumento de resistencia a la tetraciclina en *Enterococcus Faecalis* (85) son otros ejemplos de amplificación que afectan la resistencia a fármacos.

Por todo lo anterior, se puede decir que la duplicación y la amplificación son fenómenos no despreciables que pueden aportar un mejor entendimiento acerca de los mecanismos evolutivos involucrados en la emergencia y difusión de la resistencia.

TRANSFERENCIA HORIZONTAL DE GENES Y DUPLICACIÓN GÉNICA

Es sabido el papel fundamental de la transferencia horizontal en la innovación genómica de los procariotas. Este mecanismo ha sido el muy estudiado, fundamentalmente debido a su frecuencia y a sus implicancias en la bacteriología médica, por ejemplo en relación a la resistencia antibiótica. La transferencia horizontal es un proceso de gran importancia evolutiva, en el cual bacterias intercambian material genético y sufren recombinaciones posteriores. Existen tres mecanismos clásicos de transferencia génica: transformación, transducción y conjugación. La transformación es el proceso por el cual ciertas bacterias competentes son

capaces de incorporar ADN exógeno libre en el medio, proveniente de otras bacterias. El proceso de transducción implica la transferencia de dicho ADN mediante un intermediario, un bacteriófago. Finalmente, la conjugación; se basa en el intercambio unidireccional de información genética desde una bacteria donante a otra receptora, mediante un contacto real.

Un artículo publicado en 2003 sostiene que la pérdida del gen es el factor más importante en la conformación de contenido del genoma siendo hasta tres veces más frecuente que la transferencia horizontal, seguido en frecuencia por la duplicación génica, que podría contribuir hasta dos veces más en número de genes. Los análisis realizados mostraron que, a pesar de la importancia de la transferencia de genes, la mayoría de las familias de proteínas han sido amplificadas por duplicación génica (2).

Sin embargo, posteriormente el artículo “*Horizontal Transfer, Not Duplication, Drives the Expansion of Protein Families in Prokaryotes*” publicado en el año 2011 demuestra que la mayoría de las expansiones de familias de proteínas surgen por transferencia de genes entre bacterias, atribuyendo aproximadamente un 88% a este proceso.

Mediante el análisis realizado se sugiere que los Xenólogos permanecen por un periodo de tiempo más largo en el genoma en comparación al gen duplicado; se encontró que los Xenólogos tienen una edad media de introducción al genoma que es el doble al de los Parálogos (1). Estos genes transferidos horizontalmente tienden a situarse en distinta posición a la del homólogo nativo, fenómeno que explicaría la mayor persistencia de los mismos, ya que son menos propensos a la eliminación genética según los hallazgos de los autores. Por otro lado la persistencia de los genes duplicados tiende a ser más transitoria. Estos resultados permitieron a los autores proponer una función particular de cada proceso, es decir, sugirieron que la transferencia horizontal permite la adquisición de nuevas funciones mientras que la duplicación génica conduce a un mayor número de genes, los cuales podrán o no diferenciarse en cuanto a su funcionalidad (1).

Un estudio publicado en 2006 se centró en analizar la relación entre estos procesos de variabilidad genómica y ciertas familias de proteínas encargadas de la señalización, las cuales desempeñan un papel importante en la adaptación de las bacterias a entornos cambiantes permitiendo detectar y responder a señales externas. Particularmente se estudiaron las proteínas *histidin kinasas* (HPKs), las cuales históricamente se han considerado como el principal mecanismo de transducción de señales en bacterias. Observaron que muchos genomas constan con un gran repertorio de genes de señalización recientemente evolucionados, lo que puede reflejar la presión selectiva para adaptarse a las nuevas condiciones ambientales. El análisis realizado permitió dilucidar qué tanto la transferencia horizontal como la duplicación génica

tienen un rol importante en la introducción de nuevas HPKs al genoma, sin embargo se observó que estas fuerzas evolutivas actúan de manera distinta: los genes transferidos horizontalmente son propensos a retener su funcionalidad mientras que los Parálogos son fuente de diversidad genética permitiendo descubrir nuevas funciones (92). Esta investigación pone de manifiesto que existen claras preferencias específicas de las especies respecto al proceso evolutivo que predomina; en algunos genomas, tales como *E. Coli* y *Ralstonia Solanacearum*, la transferencia horizontal es responsable de casi la totalidad de las HPKs recientemente adquiridas. Para otros, tales como *Desulfovibrio Vulgaris* y *Geobacter Sulfurreducens*, la duplicación génica es el proceso predominante. La explicación de la preferencia en el proceso de adquisición de nuevas proteínas de señalización no es evidente, pero se encontró que bacterias con grandes cantidades de HPKs, inusuales en relación al tamaño del genoma, tienden a haber acumulado dichas proteínas por medio de duplicación (92).

Este estudio encontró que existen también diferencias cualitativas entre las principales estrategias para la adquisición de nuevas HPKs. Genes importados a través de transferencia horizontal son más propensos a retener su funcionalidad original, mientras que la duplicación como fenómeno de expansión parece ser una fuente importante de diversidad genética (91).

Otro estudio realizado por Sean D. Hooper y Otto G. Berg y publicado en el año 2003 muestra como la transferencia horizontal puede acelerar el proceso evolutivo de la duplicación. Se observó que la adquisición de nuevos genes duplicados ocurre con significativa frecuencia sobre genes previamente transferidos en comparación con genes nativos del genoma. Los resultados se discuten en el contexto de genes con funciones débiles principalmente y no con genes con funciones establecidas e importantes. En este trabajo se concluye que la importancia de la duplicación dependerá de la disponibilidad y variabilidad de genes con función débil, los cuales se adquieren por transferencia horizontal, para que nuevas funciones puedan ser descubiertas. Dicho fenómeno permite la supervivencia del microorganismo, ya que si el gen tiene funciones bien establecidas o, por ejemplo forma parte de un delicado equilibrio de proteínas, a continuación, la amplificación puede ser contraproducente (93). Según los autores al menos una cuarta parte de todos los genes que se duplican en una bacteria fueron adquiridos previamente por el proceso de transferencia horizontal, por lo que este último mecanismo puede tener una contribución considerable en la innovación de nuevo material genético en este contexto; además de su rol para transferir genes estables y con funciones específicas importantes para el proceso adaptativo de las bacterias a los entornos cambiantes y su consiguiente supervivencia (93).

En particular la ganancia por duplicación genética y la pérdida de genes contribuye a la adaptación del organismos, al permitir a regiones de ADN expandirse y contraerse como un acordeón. La amplificación genética es usualmente transitoria, variable y difícil de detectar, es

por esto que se refieren a ella como un modelo en forma de acordeón (17, 29, 30). La misma es observada contribuyendo a la resistencia antibiótica (como ya hemos mencionado), a la tolerancia de metales pesados, el consumo sub-óptimo de fuentes de carbono y la sobrevida en condiciones variables o extremas.

ANÁLISIS DE INPARÁLOGOS EN GENOMAS CERRADOS DE SALMONELLA

OBJETIVOS

- Encontrar duplicaciones de genes para lo cual se usa un criterio de inclusión el cual es de 95% de similitud de dichos genes.
- Localizar la cantidad de genes que están duplicados e identificar la función de los mismos.
- Identificar que regiones bacterianas (plásmidos, cromosoma, fagos) son las responsables de estos procesos.
- Luego de haber identificado con precisión a qué lugar corresponde cada duplicación, analizar en orden de frecuencia donde se produce más dicho evento si en plásmidos, cromosoma, fagos.
- Analizar si estas duplicaciones están presentes en todos los organismos estudiados.

METODOLOGIA

Mediante la utilización del programa [get_homologous.pl](#) se agruparon genomas en base a criterios de identidad y cobertura génica, con un mínimo de 95% y 75% de similitud, respectivamente. Estos parámetros permitieron detectar duplicaciones de genes muy recientes.

Se analizaron 36 genomas completos cerrados de la especie *Salmonella entérica* que presentaban duplicaciones recientes en diversos genes. Dentro de las especies se incluyeron 11 pertenecen al serotipo Typhimurium, 4 al serotipo Typhi, 4 a Paratyphi, 3 a Heidelberg, 2 a Agona, 2 a Gallinarum-Pullorum, 2 a Newport, 1 a Arizonae, 1 a Dublin, 1 a Gallinarum, 1 a Pullorum, 1 a Thompson, 1 a Schwarzengrund, 1 a Choleraesuis, 1 a Bareilly y 1 genoma sin identificación de serotipo.

RESULTADOS Y ANALISIS

Al realizar el análisis se encontraron 174 posibles genes duplicados. De estos, 10 representan probablemente genes que se encuentran duplicados basalmente en el grupo, es decir, que son familias génicas en las cuales sus dos copias homologas están conservadas en casi todos los

organismos analizados. Por lo tanto no fueron considerados en el análisis por ser posibles pérdidas secundarias en linajes específicos.

En promedio los genomas presentaron 88 genes recientemente duplicados. El genoma que presento más duplicaciones recientes fue el de Typhimurium 14028S con 160 duplicaciones mientras que el menor fue Gallinarum Pullorum RKS5078 con 40. (Tabla 1)

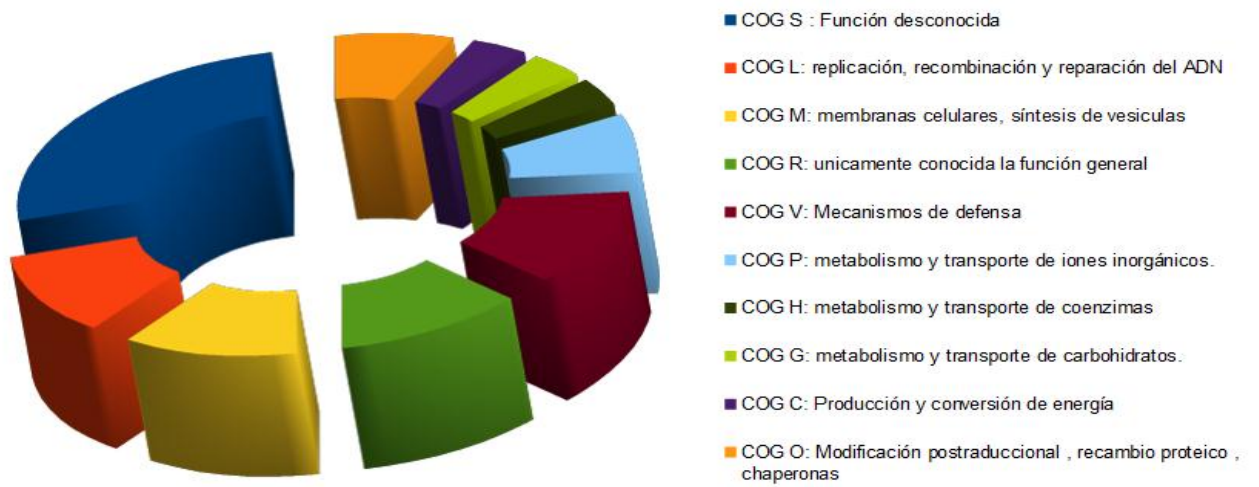
El 96,8% de las duplicaciones fueron localizadas en cromosomas mientras que el restante 3,2% fueron localizados en plásmidos. En total se identificaron 20 clusters duplicados que representan 106 genes duplicados recientemente.

En cuanto a las proteínas codificadas por los genes analizados encontramos que el 25,0% de las duplicaciones recientes corresponden a proteínas de fagos y el 16,5% pertenecen a transposasas y genes relacionados a elementos móviles. El resto tiene una distribución relativamente heterogénea.

Por otra parte el 15,9% de los genes analizados están relacionados con resistencia a antibióticos y otros compuestos químicos. Se identificaron duplicaciones recientes en los genes que codifican proteínas de resistencia a aminoglucósidos, proteínas de flujo multidroga, betalactamasas, proteínas de resistencia (amonió), biosíntesis de antígeno, proteínas de virulencia, proteínas de replicación del ADN. (Figura 1)

Estos genes de resistencia fueron localizados en cepas pertenecientes a los serotipos: Schwarzengrund, Typhi CT18, Typhimurium T000240, Typhimurium U288.

Figura 1: Frecuencia de Categorías COG
Análisis de genes vinculados a la resistencia



NOMBRE DEL GENOMA	TOTAL DE GENES CROMOSOMA	TOTAL DE GENES PLÁSMIDOS	TOTAL DUPLICACIONES
Arizonae_serovar_62_z4_z23	4467		57
Salmonella_enterica_serovar_4_5_12_i_08_1736	4489		113
Agona_24249	4342		57
Agona_SL483	4580	52	66
Bareilly_CFSAN000189	4363	78	69
Choleraesuis_SC_B67	4385	170	146
Dublin_CT_02021853	4530	103	73
Gallinarum_287_91	3965		43
Gallinarum_Pullorum_CDC1983_67	4331		41
Gallinarum_Pullorum_RKS5078	4325		40
Heidelberg_B182	4282	44	51
Heidelberg_CFSAN002069	4447		54
Heidelberg_SL476	4666		71
Newport_SL254	4605	3	106
Newport_USMARC_S3124_1	4828		87
Paratyphi_A_AKU_12601	4079		42
Paratyphi_A_ATCC_9150	4096		46
Paratyphi_C_RKS4594	4577	62	90
Pullorum_S06004	4052		46
Schwarzengrund_CVM19633	4538		101
Thompson_RM6836	4357		73

Typhi_CT18	4111		105
Typhimurium_14028S	5315	101	160
Typhimurium_798	4318	104	97
Typhimurium_D23580	4446		102
Typhimurium_DT2	4321		84
Typhimurium_LT2	4451	102	108
Typhimurium_SL1344	4446		105
Typhimurium_ST4_74	4625		111
Typhimurium_T000240	4713		155
Typhimurium_U288	4569		119
Typhimurium_UK_1	4451		111
Typhimurium_var_5_CFSAN001921	4561		129
Typhi_P_stx_12	4690		108
Typhi_Ty21a	4494		97
Typhi_Ty2	4352		94

TABLA 1

CONCLUSIONES GENERALES

En este trabajo se revisó el fenómeno de la duplicación reciente en Enterobacterias, es decir la presencia de Imparálogos. El mismo no es tan frecuente como la transferencia horizontal, principal mecanismo de adquisición de genes en procariotas. Sin embargo, es un proceso esencial para la conservación y evolución del genoma y también juega un rol en la determinación de las características fenotípicas de los microorganismos. Esto queda evidenciado en una serie de trabajos experimentales que permiten decir que las bacterias realizan duplicaciones altamente relevantes para su supervivencia, las cuales están vinculadas a diferentes grupos, cada grupo con una función específica, vinculada a la resistencia, la

virulencia o la adaptación al medio. Lo que es apoyado además por varios ejemplos que ponen en evidencia que la duplicación y amplificación genética permiten al organismo habituarse al cambio ambiental modificando su resistencia antibiótica, la virulencia de los patógenos, la eficacia de las vacunas y la biosíntesis de los antibióticos.

En paralelo, se analizaron un total de 36 genomas completos cerrados de la especie *S. Entérica*, en los cuales, con un criterio de identidad mínimo de 95% y cobertura de 75%, se encontraron 174 duplicaciones recientes. A este número se le resta la cantidad de fagos y transposasas encontrados. La transposasa es una proteína que le permite al fago (un virus bacteriófago que infecta a la bacteria) insertarse aleatoriamente en el cromosoma.

El número de elementos móviles del genoma fue mayor al que esperábamos encontrar, de todos modos consideramos importante detectarlos y excluirllos. Cabe mencionar que de esos 174 también se excluyeron las duplicaciones basales.

Una vez realizado lo anterior, se llegó a la conclusión de que el 15,9% de las duplicaciones recientes se dan en genes vinculados con la resistencia antibiótica. Los mismos están altamente conservados en todas las serovariedades y duplicados en alguna de ellas, lo cual habla de la relevancia que estos tienen para las bacterias.

Se vio también que las duplicaciones son más frecuentes en cromosomas que en plásmidos, esto resalta la trascendencia (envergadura, importancia) del fenómeno de la duplicación en bacterias, el cual es un hecho para nada despreciable que deberá seguir siendo estudiado, y que con seguridad realizara grandes aportes a la medicina.

REFERENCIAS

1. Treangen TJ, Rocha EP. Horizontal transfer, not duplication, drives the expansion of protein families in prokaryotes. *PLoS Genet.* 2011;7(1):e1001284.
2. Kunin V, Ouzounis CA. The balance of driving forces during genome evolution in prokaryotes. *Genome Res.* 2003;13(7):1589-94.
3. Gevers D, Vandepoele K, Simillon C, Van de Peer Y. Gene duplication and biased functional retention of paralogs in bacterial genomes. *Trends Microbiol.* 2004;12(4):148-54.
4. Mira A, Pushker R, Rodriguez-Valera F. The Neolithic revolution of bacterial genomes. *Trends Microbiol.* 2006;14(5):200-6.
5. Blattner FR, Plunkett G, 3rd, Bloch CA, Perna NT, Burland V, Riley M, et al. The complete genome sequence of *Escherichia coli* K-12. *Science.* 1997;277(5331):1453-62.
6. Sonnhammer EL, Koonin EV. Orthology, paralogy and proposed classification for paralog subtypes. *Trends Genet.* 2002;18(12):619-20.
7. Bratlie MS, Johansen J, Sherman BT, Huang da W, Lempicki RA, Drablos F. Gene duplications in prokaryotes can be associated with environmental adaptation. *BMC Genomics.* 2010;11:588.

8. Anderson RP, Roth JR. Tandem genetic duplications in phage and bacteria. *Annu Rev Microbiol.* 1977;31:473-505.
9. Lupski JR, Roth JR, Weinstock GM. Chromosomal duplications in bacteria, fruit flies, and humans. *Am J Hum Genet.* 1996;58(1):21-7.
10. Bergthorsson U, Andersson DI, Roth JR. Ohno's dilemma: evolution of new genes under continuous selection. *Proc Natl Acad Sci U S A.* 2007;104(43):17004-9.
11. Romero D, Palacios R. Gene amplification and genomic plasticity in prokaryotes. *Annu Rev Genet.* 1997;31:91-111.
12. Beckmann JS, Estivill X, Antonarakis SE. Copy number variants and genetic traits: closer to the resolution of phenotypic to genotypic variability. *Nat Rev Genet.* 2007;8(8):639-46.
13. Conrad B, Antonarakis SE. Gene duplication: a drive for phenotypic diversity and cause of human disease. *Annu Rev Genomics Hum Genet.* 2007;8:17-35.
14. Sandegren L, Andersson DI. Bacterial gene amplification: implications for the evolution of antibiotic resistance. *Nat Rev Microbiol.* 2009;7(8):578-88.
15. Lee C, Lee S, Shin SG, Hwang S. Real-time PCR determination of rRNA gene copy number: absolute and relative quantification assays with *Escherichia coli*. *Appl Microbiol Biotechnol.* 2008;78(2):371-6.
16. Anderson P, Roth J. Spontaneous tandem genetic duplications in *Salmonella typhimurium* arise by unequal recombination between rRNA (*rrn*) cistrons. *Proc Natl Acad Sci U S A.* 1981;78(5):3113-7.
17. Nilsson AI, Zorzet A, Kanth A, Dahlstrom S, Berg OG, Andersson DI. Reducing the fitness cost of antibiotic resistance by amplification of initiator tRNA genes. *Proc Natl Acad Sci U S A.* 2006;103(18):6976-81.
18. Andersson DI, Slechta ES, Roth JR. Evidence that gene amplification underlies adaptive mutability of the bacterial lac operon. *Science.* 1998;282(5391):1133-5.
19. Straus DS, D'Ari Straus L. Large overlapping tandem genetic duplications in *Salmonella typhimurium*. *J Mol Biol.* 1976;103(1):143-53.
20. Schaechter M, View From Here G. *Escherichia coli* and *Salmonella* 2000: the view from here. *Microbiol Mol Biol Rev.* 2001;65(1):119-30.
21. Lehner AF, Hill CW. Involvement of ribosomal ribonucleic acid operons in *Salmonella typhimurium* chromosomal rearrangements. *J Bacteriol.* 1980;143(1):492-8.
22. Bertini A, Poirel L, Bernabeu S, Fortini D, Villa L, Nordmann P, et al. Multicopy blaOXA-58 gene as a source of high-level resistance to carbapenems in *Acinetobacter baumannii*. *Antimicrob Agents Chemother.* 2007;51(7):2324-8.
23. Szczepanowski R, Krahn I, Puhler A, Schluter A. Different molecular rearrangements in the integron of the IncP-1 beta resistance plasmid pB10 isolated from a wastewater treatment plant result in elevated beta-lactam resistance levels. *Arch Microbiol.* 2004;182(6):429-35.
24. Haack KR, Roth JR. Recombination between chromosomal IS200 elements supports frequent duplication formation in *Salmonella typhimurium*. *Genetics.* 1995;141(4):1245-52.
25. Raymond M, Chevillon C, Guillemaud T, Lenormand T, Pasteur N. An overview of the evolution of overproduced esterases in the mosquito *Culex pipiens*. *Philos Trans R Soc Lond B Biol Sci.* 1998;353(1376):1707-11.
26. Shen P, Huang HV. Homologous recombination in *Escherichia coli*: dependence on substrate length and homology. *Genetics.* 1986;112(3):441-57.
27. Watt VM, Ingles CJ, Urdea MS, Rutter WJ. Homology requirements for recombination in *Escherichia coli*. *Proc Natl Acad Sci U S A.* 1985;82(14):4768-72.
28. Petit MA, Mesas JM, Noirot P, Morel-Deville F, Ehrlich SD. Induction of DNA amplification in the *Bacillus subtilis* chromosome. *EMBO J.* 1992;11(4):1317-26.
29. Galitski T, Roth JR. Pathways for homologous recombination between chromosomal direct repeats in *Salmonella typhimurium*. *Genetics.* 1997;146(3):751-67.
30. Jessop AP, Clugston C. Amplification of the ArgF region in strain HfrP4X of *E. coli* K-12. *Mol Gen Genet.* 1985;201(2):347-50.

31. Tlsty TD, Albertini AM, Miller JH. Gene amplification in the lac region of *E. coli*. *Cell*. 1984;37(1):217-24.
32. Whoriskey SK, Nghiem VH, Leong PM, Masson JM, Miller JH. Genetic rearrangements and gene amplification in *Escherichia coli*: DNA sequences at the junctures of amplified gene fusions. *Genes Dev*. 1987;1(3):227-37.
33. Pettersson ME, Sun S, Andersson DI, Berg OG. Evolution of new gene functions: simulation and analysis of the amplification model. *Genetica*. 2009;135(3):309-24.
34. Kugelberg E, Kofoed E, Reams AB, Andersson DI, Roth JR. Multiple pathways of selected gene amplification during adaptive mutation. *Proc Natl Acad Sci U S A*. 2006;103(46):17319-24.
35. Nicoloff H, Perreten V, McMurry LM, Levy SB. Role for tandem duplication and lon protease in AcrAB-TolC- dependent multiple antibiotic resistance (Mar) in an *Escherichia coli* mutant without mutations in marRAB or acrRAB. *J Bacteriol*. 2006;188(12):4413-23.
36. Rownd R, Mickel S. Dissociation and reassociation of RTF and r-determinants of the R-factor NR1 in *Proteus mirabilis*. *Nat New Biol*. 1971;234(45):40-3.
37. Edlund T, Normark S. Recombination between short DNA homologies causes tandem duplication. *Nature*. 1981;292(5820):269-71.
38. Brochet M, Couve E, Zouine M, Poyart C, Glaser P. A naturally occurring gene amplification leading to sulfonamide and trimethoprim resistance in *Streptococcus agalactiae*. *J Bacteriol*. 2008;190(2):672-80.
39. Seoane A, Sanchez E, Garcia-Lobo JM. Tandem amplification of a 28-kilobase region from the *Yersinia enterocolitica* chromosome containing the blaA gene. *Antimicrob Agents Chemother*. 2003;47(2):682-8.
40. Roth JR, Kugelberg E, Reams AB, Kofoed E, Andersson DI. Origin of mutations under selection: the adaptive mutation controversy. *Annu Rev Microbiol*. 2006;60:477-501.
41. Serres MH, Kerr AR, McCormack TJ, Riley M. Evolution by leaps: gene duplication in bacteria. *Biol Direct*. 2009;4:46.
42. Labedan B, Riley M. Gene products of *Escherichia coli*: sequence comparisons and common ancestries. *Mol Biol Evol*. 1995;12(6):980-7.
43. Labedan B, Riley M. Widespread protein sequence similarities: origins of *Escherichia coli* genes. *J Bacteriol*. 1995;177(6):1585-8.
44. Liang P, Labedan B, Riley M. Physiological genomics of *Escherichia coli* protein families. *Physiol Genomics*. 2002;9(1):15-26.
45. Serres MH, Riley M. Gene fusions and gene duplications: relevance to genomic annotation and functional analysis. *BMC Genomics*. 2005;6:33.
46. Hooper SD, Berg OG. On the nature of gene innovation: duplication patterns in microbial genomes. *Mol Biol Evol*. 2003;20(6):945-54.
47. Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res*. 2009;37(1):1-13.
48. Marstrand TT, Frellsen J, Moltke I, Thiim M, Valen E, Retelska D, et al. Asap: a framework for over-representation statistics for transcription factor binding sites. *PLoS One*. 2008;3(2):e1623.
49. Barkow S, Bleuler S, Prelic A, Zimmermann P, Zitzler E. BicAT: a biclustering analysis toolbox. *Bioinformatics*. 2006;22(10):1282-3.
50. Domazet-Lošo M, Haubold B. Efficient estimation of pairwise distances between genomes. *Bioinformatics*. 2009;25(24):3221-7.
51. Ridge PG, Zhang Y, Gladyshev VN. Comparative genomic analyses of copper transporters and cuproproteomes reveal evolutionary dynamics of copper utilization and its link to oxygen. *PLoS One*. 2008;3(1):e1378.
52. Hobman JL. MerR family transcription activators: similar designs, different specificities. *Mol Microbiol*. 2007;63(5):1275-8.

53. Peariso K, Huffman DL, Penner-Hahn JE, O'Halloran TV. The PcoC copper resistance protein coordinates Cu(I) via novel S-methionine interactions. *J Am Chem Soc.* 2003;125(2):342-3.
54. Anderson JK, Smith TG, Hoover TR. Sense and sensibility: flagellum-mediated gene regulation. *Trends Microbiol.* 2010;18(1):30-7.
55. Shen Y, Ronald P. Molecular determinants of disease and resistance in interactions of *Xanthomonas oryzae* pv. *oryzae* and rice. *Microbes Infect.* 2002;4(13):1361-7.
56. Shen Y, Chern M, Silva FG, Ronald P. Isolation of a *Xanthomonas oryzae* pv. *oryzae* flagellar operon region and molecular characterization of flhF. *Mol Plant Microbe Interact.* 2001;14(2):204-13.
57. Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res.* 2009;37(Database issue):D233-8.
58. Kataeva IA, Yang SJ, Dam P, Poole FL, 2nd, Yin Y, Zhou F, et al. Genome sequence of the anaerobic, thermophilic, and cellulolytic bacterium "Anaerocellum thermophilum" DSM 6725. *J Bacteriol.* 2009;191(11):3760-1.
59. Bentley SD, Chater KF, Cerdeno-Tarraga AM, Challis GL, Thomson NR, James KD, et al. Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature.* 2002;417(6885):141-7.
60. Perna NT, Plunkett G, 3rd, Burland V, Mau B, Glasner JD, Rose DJ, et al. Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. *Nature.* 2001;409(6819):529-33.
61. Toh H, Weiss BL, Perkin SA, Yamashita A, Oshima K, Hattori M, et al. Massive genome erosion and functional adaptations provide insights into the symbiotic lifestyle of *Sodalis glossinidius* in the tsetse host. *Genome Res.* 2006;16(2):149-56.
62. Starr CR, Engleberg NC. Role of hyaluronidase in subcutaneous spread and growth of group A streptococcus. *Infect Immun.* 2006;74(1):40-8.
63. Collier J. Epigenetic regulation of the bacterial cell cycle. *Curr Opin Microbiol.* 2009;12(6):722-9.
64. Heusipp G, Falker S, Schmidt MA. DNA adenine methylation and bacterial pathogenesis. *Int J Med Microbiol.* 2007;297(1):1-7.
65. Low DA, Weyand NJ, Mahan MJ. Roles of DNA adenine methylation in regulating bacterial gene expression and virulence. *Infect Immun.* 2001;69(12):7197-204.
66. Goris J, De Vos P, Coenye T, Hoste B, Janssens D, Brim H, et al. Classification of metal-resistant bacteria from industrial biotopes as *Ralstonia campinensis* sp. nov., *Ralstonia metallidurans* sp. nov. and *Ralstonia basileensis* Steinle et al. 1998 emend. *Int J Syst Evol Microbiol.* 2001;51(Pt 5):1773-82.
67. Kandler O, Hippe H. Lack of peptidoglycan in the cell walls of *Methanosarcina barkeri*. *Arch Microbiol.* 1977;113(1-2):57-60.
68. Jarvis GN, Strompl C, Burgess DM, Skillman LC, Moore ER, Joblin KN. Isolation and identification of ruminal methanogens from grazing cattle. *Curr Microbiol.* 2000;40(5):327-32.
69. Flo TH, Smith KD, Sato S, Rodriguez DJ, Holmes MA, Strong RK, et al. Lipocalin 2 mediates an innate immune response to bacterial infection by sequestering iron. *Nature.* 2004;432(7019):917-21.
70. Braun V, Killmann H. Bacterial solutions to the iron-supply problem. *Trends Biochem Sci.* 1999;24(3):104-9.
71. Hille R. Molybdenum and tungsten in biology. *Trends Biochem Sci.* 2002;27(7):360-7.
72. Macara IG, Lounsbury KM, Richards SA, McKiernan C, Bar-Sagi D. The Ras superfamily of GTPases. *FASEB J.* 1996;10(5):625-30.
73. Gogarten JP, Olendzenski L. Orthologs, paralogs and genome comparisons. *Curr Opin Genet Dev.* 1999;9(6):630-6.
74. Mongodin EF, Nelson KE, Daugherty S, Deboy RT, Wister J, Khouri H, et al. The genome of *Salinibacter ruber*: convergence and gene exchange among hyperhalophilic bacteria and archaea. *Proc Natl Acad Sci U S A.* 2005;102(50):18147-52.

75. Sanchez-Perez G, Mira A, Nyiro G, Pasic L, Rodriguez-Valera F. Adapting to environmental changes using specialized paralogs. *Trends Genet.* 2008;24(4):154-8.
76. Ihmels J, Collins SR, Schuldiner M, Krogan NJ, Weissman JS. Backup without redundancy: genetic interactions reveal the cost of duplicate gene loss. *Mol Syst Biol.* 2007;3:86.
77. Krakauer DC, Plotkin JB. Redundancy, antiredundancy, and the robustness of genomes. *Proc Natl Acad Sci U S A.* 2002;99(3):1405-9.
78. de Visser JA, Hermisson J, Wagner GP, Ancel Meyers L, Bagheri-Chaichian H, Blanchard JL, et al. Perspective: Evolution and detection of genetic robustness. *Evolution.* 2003;57(9):1959-72.
79. Francino MP. An adaptive radiation model for the origin of new gene functions. *Nat Genet.* 2005;37(6):573-7.
80. Pegueroles C, Laurie S, Alba MM. Accelerated evolution after gene duplication: a time-dependent process affecting just one copy. *Mol Biol Evol.* 2013;30(8):1830-42.
81. Nasvall J, Sun L, Roth JR, Andersson DI. Real-time evolution of new genes by innovation, amplification, and divergence. *Science.* 2012;338(6105):384-7.
82. Harbottle H, Thakur S, Zhao S, White DG. Genetics of antimicrobial resistance. *Anim Biotechnol.* 2006;17(2):111-24.
83. Normark BH, Normark S. Evolution and spread of antibiotic resistance. *J Intern Med.* 2002;252(2):91-106.
84. Andersson DI, Hughes D. Gene amplification and adaptive evolution in bacteria. *Annu Rev Genet.* 2009;43:167-95.
85. Craven SH, Neidle EL. Double trouble: medical implications of genetic duplication and amplification in bacteria. *Future Microbiol.* 2007;2(3):309-21.
86. Fallis G, Hilditch J. A comparison of seasonal variation in birthweights between rural Zaire and Ontario. *Can J Public Health.* 1989;80(3):205-8.
87. Wang YD, Thompson JR, Goulstine DB, Rosenthal AR. A survey of the initial referral of children to an ophthalmology department. *Br J Ophthalmol.* 1990;74(11):650-3.
88. Wannarat W, Motoyama S, Masuda K, Kawamura F, Inaoka T. Tetracycline tolerance mediated by gene amplification in *Bacillus subtilis*. *Microbiology.* 2014;160(Pt 11):2474-80.
89. Massacesi L, Abbamondi AL, Giorgi C, Sarlo F, Lolli F, Amaducci L. Suppression of experimental allergic encephalomyelitis by retinoic acid. *J Neurol Sci.* 1987;80(1):55-64.
90. Corrado F, Doninelli M, Urbanetti S, Valli E. [Clinico-statistic study on 126 cases of benign ovarian neoplasms]. *Minerva Ginecol.* 1988;40(12):687-90.
91. Elliott KT, Cuff LE, Neidle EL. Copy number change: evolving views on gene amplification. *Future Microbiol.* 2013;8(7):887-99.
92. Alm E, Huang K, Arkin A. The evolution of two-component systems in bacteria reveals different strategies for niche adaptation. *PLoS Comput Biol.* 2006;2(11):e143.
93. Hooper SD, Berg OG. Duplication is more common among laterally transferred genes than among indigenous genes. *Genome Biol.* 2003;4(8):R48.