

INSTITUTO POLITECNICO NACIONAL
ESCUELA NACIONAL DE CIENCIAS BIOLÓGICAS

Unidad de Investigación de Enfermedades Infecciosas y Parasitarias
Hospital de Pediatría – Centro Médico Nacional Siglo XXI

Resolución de huecos encontrados durante la
secuenciación de una cepa mexicana de
Helicobacter pylori aislada de un paciente con
cáncer gástrico

Proyecto de Investigación

Que como uno de los requisitos para obtener el título de:
Químico Bacteriólogo Parasitólogo

Presenta:

Muñoz Mateo Luis Fernando

Directoras de tesis:

M. en C. Hilda Alicia Valdez Salazar

Dra. Violeta Larios Serrato



México D.F.

Año: 2016

El presente trabajo se realizó Laboratorio de Bacteriología de la Unidad de Investigación Médica de Enfermedades Infecciosas y Parasitarias (UIMEIP) del Hospital de Pediatría del Centro Médico Nacional (CMN) Siglo XXI y del laboratorio de Biotecnología y Bioinformática Genómica, departamento de Bioquímica, Escuela Nacional de Ciencias Biológicas, IPN, bajo la dirección de la M. en C. Hilda Alicia Valdez Salazar y la Dra. en C. Violeta Larios Serrato.

Agradecimientos

Este trabajo no hubiera sido capaz sin aquellas personas que estuvieron conmigo a lo largo del camino la M. en C. Hilda, la Dra. Violeta y el Dr. Avilés, quienes me compartieron todos sus conocimientos, experiencias y el material necesario para desarrollar los experimentos, y llevar a por buen camino proyecto. Nunca podré terminar de agradecer su confianza para permitirme trabajar con ustedes.

También el pilar de mi vida que sin su apoyo esto no hubiera sido posible, para que pudiera concluir esta meta, “mi familia”. Mi madre Patricia por haberme apoyado en todo momento, por sus consejos, sus valores, por la motivación constante que me ha permitido ser una persona de bien, pero más que nada, por su amor. A mi padre Felipe por los ejemplos de responsabilidad y constancia que lo caracterizan y que me ha infundado siempre, por el valor mostrado para salir adelante. A mi hermano Gustavo que muy a su manera me reta a continuar y ser mejor. Este trabajado está dedicado a ellos

A la familia Torres Barrera, que han sido fuente de apoyo y motivación llegando a ser un gran ejemplo de vida y cómo quiero llegar a ser, Emmanuel amigo de una vida, Sra. Anita consejera.

A mis compañeros que son grandes amigos Roberto que me has apoyado con tus conocimientos y contribución a este trabajo, Zilia que te has convertido en una gran amiga para mí y de gran apoyo en diversos momentos.

A mis amigos de la carrera que han estado para mí cuando más los he necesitado, David que has sido como un hermano para mí desde la prepa, también desde aquellos tiempos Vanessa, Irvin, Dulce, Brenda, Vicky. Mis amigos de la universidad; Rubén que eres... tu sabes que eres un gran amigo para mí, Emmita mi mejor amiga que has estado en las buenas y malas, Uriel que hemos compartido cada momento, Manalli que incontables veces fuiste mi salvadora. Personas que me eh encontrado en mis andadas como Jess gracias por ser mi amiga.

No puedo evitar escribir esto para ella, con la esperanza de que lo lea, solo quiero decirle que ella fue uno de los motivos más importantes que me llevo a decidir y animarme en estudiar esta gratificante carrera y que me llevo hasta esta situación.

Todos aquellos familiares y amigos que no recordé al momento de escribir esto. Ustedes saben quiénes son.

A todos ustedes, gracias infinitas.

“Mi salvación fue leer, leer los buenos libros, refugiarme en esos mundos donde vivir era exaltante, intenso, una aventura tras otra, donde podía sentirme libre y volvía a ser feliz”

— Mario Vargas Llosa

“Las ciencias aplicadas no existen, sólo las aplicaciones de la ciencia.”

— Louis Pasteur

“La ciencia no nos ha enseñado aún si la locura es o no lo más sublime de la inteligencia.”

— Edgar Allan Poe

“Los hombres de ciencia sospechan algo sobre ese mundo, pero lo ignoran casi todo. Los sabios interpretan los sueños, y los dioses se ríen.”

— Howard Phillips Lovecraft

INDICE GENERAL

Índice general.....	I
Índice de Tablas	II
Índice de Gráficas.....	II
Índice de Figuras.....	III
Lista de Abreviaturas.....	V
Resumen	VI
1. Introducción.....	1
1.1 Marco histórico.....	1
1.2 Epidemiología de gastritis asociadas con <i>H. pylori</i>	2
1.3 Transmisión.....	2
1.4 Asociación con enfermedades	3
1.5 Características microbiológicas.....	4
1.5.1 Morfología	4
1.5.2 Cultivo bacteriano	4
1.6 Virulencia y patogenicidad de <i>H. pylori</i>.....	5
1.7 Genoma.....	6
1.8 Secuenciación de DNA.....	7
1.8.1 Preparación del material genético.....	7
1.8.2 Secuenciación	8
1.8.3 Ensamblado genómico	8
1.8.4 Anotación genómica.....	9
2. Justificación	10
3. Hipótesis	11
4. Objetivo general	12
4.1 Objetivos particulares	12
5. Materiales y métodos.....	11
5.1 Extracción de DNA.....	14
5.2 Diseño de primers.....	15
5.2.1 Diseño de primers primera estrategia	15
5.2.2 Diseño de primers segunda estrategia empleando cepa de referencia <i>H. pylori</i> ELS3716	15
5.3 Reacción en cadena de la polimerasa (PCR)	17
5.4 Gel de agarosa y electroforesis	18
5.5 Purificación de productos de PCR a partir del gel de agarosa con el kit QIAquick gel Extraction (QIAGEN).	19
5.6 Secuenciación de los fragmentos.....	20

5.7 Ensamblado de los fragmentos secuenciados	20
5.8 Anotación genómica	21
6. Resultados	23
6.1 Diseño de primers específicos	23
6.2 Amplificación mediante PCR de los fragmentos que unan a las secuencias problema.	24
6.3 Purificación y secuenciación automática	26
6.4 Ensamblado de las secuencias.	28
6.5 Anotación genómica	30
7. Discusiones	34
8. Conclusiones	37
9. Bibliografía	38
10. Apéndice	42
10.1 Cálculos para preparar un gel de Agarosa	42

INDICE DE GRÁFICAS

	Página
1 Casos e incidencia de úlceras, gastritis y duodenitis México 2003-2008	2

INDICE DE TABLAS

	Página
1 Características genómicas comparativas de <i>H. pylori</i>	6
2 Lista de huecos con los contigs para llevar a cabo la secuencia, así como las proteínas usadas y a que cepa pertenece.	15
3 Mezcla maestra de PCR	17
4 Ciclo de amplificación de DNA	17
5 Mezcla maestra de PCR	20
6 Diseño de Primers de los contigs con la primera estrategia	23
7 Diseño de primers de los contigs con la segunda estrategia	24
8 Lista de contigs unidos mediante las técnicas empleadas	33

INDICE DE FIGURAS

	Página
1 <i>H. pylori</i>	4
2 Colonias de <i>H. pylori</i>	4
3 Secuenciación de DNA	9
4 Diagrama general de trabajo	13
5 Alineamiento entre los contigs seleccionados de la tabla 5 y la secuencia del gen que codifica para la proteína usada como referencia.	16
6 Formulario de información del organismo en un proyecto de anotación en el servidor RAST	21
7 Opciones para el procedimiento de anotación en servidor RAST	22
8 Productos de PCR <i>H. pylori</i> CG-IMSS-2012 de primeros primers	23
9 Productos de PCR <i>H. pylori</i> CG-IMSS-2012 de segundos primers	23
10 Secuencia de nucleótidos del fragmento 286-631-26F para el hueco entre los contigs 026 – 024	26
11 Secuencia de nucleótidos del fragmento 286-515-16F para el hueco entre los contigs 016 – 017	26
12 Secuencia de nucleótidos del fragmento 456-566-34R para el hueco entre los contigs 002 – 034	26
13 Secuencia de nucleótidos del fragmento 456-700-14F para el hueco entre los contigs 014 – 015	27
14 Secuencia de nucleótidos del fragmento 456-735-31F para el hueco entre los contigs 031 – 032	27
15 Secuencia de nucleótidos del fragmento 456-712-32F para el hueco entre los contigs 032 – 033	28
16 Secuencia de nucleótidos del fragmento 456-722-39F para el hueco entre los contigs 039 – 040	28
17 Ensamblado del fragmento 456-566-34R secuenciado con los contigs 002 – 034	28
18 Ensamblado del fragmento 286-515-16F secuenciado con los contigs 016 – 017	29
19 Ensamblado del fragmento 286-631-26F secuenciado con los contigs 026 – 024	29
20 Ensamblado del fragmento 456-700-14F secuenciado con los contigs 014 – 015	29
21 Ensamblado del fragmento 456-735-31F secuenciado con los contigs 031 – 032	29
22 Ensamblado del fragmento 456-712-32F secuenciado con los contigs 032 – 033	30
23 Ensamblado del fragmento 456-722-39F secuenciado con los contigs 039 – 040	30
24 Anotación genómica resultante de la unión de los contigs 14 – 15	31
25 Anotación genómica resultante de la unión de los contigs 31 – 32	31

26	Anotación genómica resultante de la unión de los contigs 32 – 33	32
27	Anotación genómica resultante de la unión de los contigs 39 – 40	32

Lista de Abreviaturas

Abreviatura: Significado

ATCC: American Type Control Collection

cagA: gen A asociado a la citotoxina

cagPAI: isla de patogenicidad

DNA: ácido desoxirribonucleico

H. pylori: *Helicobacter pylori*

Mb: Mega pares de bases

NCBI: National Center for Biotechnology Information

NGS: Secuenciación de nueva generación

pb: Pares de bases

PBS: Buffer fosfato salino

PCR: Reacción en cadena de la polimerasa

PE: Buffers Tris-HCl, Etanol

QG: Tiocianato de guanidina - Tris

RAST: Rapid Annotation using Subsystem Technology

SINAVE: Sistema Nacional de Vigilancia Epidemiológica

TAE: Tris, Ácido Acético, EDTA

vacA: Toxina vacuolizante

Resumen

Helicobacter pylori infecta a más del 50% de la población mundial, coloniza la mucosa gástrica, es la principal causa del desarrollo de gastritis crónica y es un factor de riesgo para la aparición de cáncer gástrico. *H. pylori* se caracteriza por su gran diversidad genética por lo que existen notables diferencias en la estructura y organización del genoma entre cepas provenientes de diferentes regiones geográficas.

Por lo anterior, el objetivo de este trabajo fue resolver el mayor número de huecos posibles encontrados durante el ensamblado del genoma de la cepa de *H. pylori* CG-IMSS-2012.

Previamente la cepa *H. pylori* CG-IMSS-2012 fue sometido a secuenciación masiva, y en el ensamblado se obtuvieron 45 contigs. A partir de un análisis bioinformático previo, se identificaron las proteínas cuya secuencia codificante era compartida por más de un segmento diferente lo que generó cinco candidatos. Se diseñaron primers mediante el programa Primer-BLAST y se establecieron las condiciones para amplificar los fragmentos mediante la técnica de PCR. Los productos obtenidos se sometieron a secuenciación automática. Una segunda estrategia es el alineamiento de los extremos de cada contig y la secuencia de la cepa *H. pylori* ELS37 filogenéticamente más cercana a la secuencia problema con el programa MegAlign DNASTar v5.03, con el objeto de identificar que extremos pudieran estar continuos en el genoma de la cepa de *H. pylori* CG-IMSS-2012; y se diseñaron los primers correspondientes para la amplificación de los fragmentos.

Se amplificaron 3 productos correspondientes con la primera estrategia para finalmente obtener su secuencia. Por otro lado, el alineamiento de los extremos de los contigs ha generado 5 candidatos para diseño de iniciadores y amplificaron 4 de los 5 segmentos que se pudieron secuenciar y ensamblar. Con los ensamblados correspondientes de la segunda estrategia se prosiguió a realizar su anotación genómica, descubriendo que proteínas se encontraban en los extremos y la existencia de regiones intergenicas.

Con las estrategias planteadas se ha resuelto la secuencia de 7 huecos dando el lugar a la formación de scaffolds y se redujo el número de contigs a 31 por unir acercándose al objetivo de completar el genoma. Conocer el genoma de una cepa Mexicana de *H. pylori* será de suma importancia en futuras investigaciones ya que podrá servir como referencia en diversos estudios en México. Representará un punto de referencia para conocer las características genéticas de las cepas que prevalecen en nuestro país.

1. Introducción

1.1 Marco histórico

El patólogo Robin Warren a principios de junio de 1979, observó por primera vez a *Helicobacter pylori* en una biopsia gástrica proveniente de un paciente con gastritis crónica activa. Sus observaciones en estudios histopatológicas, continuaron por un par de años, asociando la presencia del microorganismo con esta patología. En 1981, un médico gastroenterólogo Barry Marshall se une a la investigación realizada por Warren y confirma lo reportado por éste último (Warren and Marshall, 1983).

La bacteria morfológicamente semejaba a *Campylobacter* spp., razón por la cual fue llamada *Campylobacter pyloridis*, se emplearon los medios específicos de *Campylobacter* y las condiciones de su crecimiento para aislar a la bacteria de biopsias gástricas bajo las condiciones de crecimiento para *Campylobacter* spp (Marshall and Goodwin, 1987).

Sin embargo fue hasta 1982, que *H. pylori* fue aislada por primera vez, esto sucedió después de dejar el cultivo por más de 5 días debido a que hubo un día de asueto, que prolongó el fin de semana, dando el tiempo necesario para permitir el crecimiento bacteriano adecuado (Warren JR., 1997). En 1984, que se publicó en la revista *Lancet* la asociación de *H. pylori* con la gastritis crónica y por primera vez se sugirió que la úlcera péptica pudiera ser de etiología infecciosa. En 1994 se efectuó una conferencia consenso por los Institutos Nacionales de Salud, donde *H. pylori* es declarado la principal causa de úlcera péptica y es en éste mismo año que la Agencia Internacional para la Investigación del Cáncer de la Organización Mundial de la Salud, declaró que *H. pylori* es un cancerígeno en humanos (Blaser, M. J., 1992).

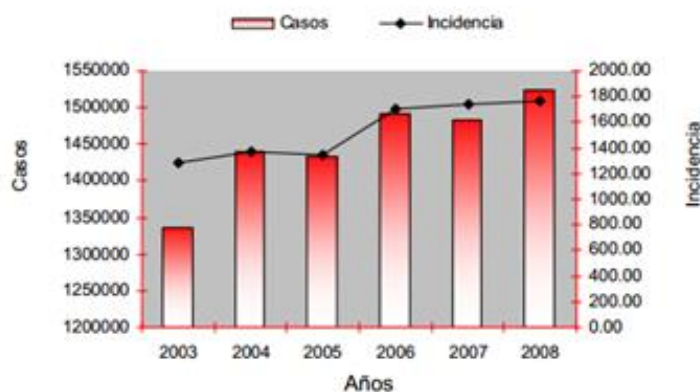
Esta bacteria coloniza el estómago del humano y como característica especial es la colonización en forma de parches en la mucosa gástrica, permaneciendo en la mucosa por años o décadas. Las personas infectadas con este organismo tienen una inflamación crónica superficial que involucra tanto al antro como al fondo gástrico; la asociación de *H. pylori* con diferentes patologías gástricas es altamente significativa (Miyazawa M, *et al.*, 2003); persiste generalmente de por vida, a no ser erradicado por la terapia antimicrobiana (Malfertheiner P. *et al.*, 2002).

“Resolución de huecos encontrados durante la secuenciación de una cepa mexicana de *Helicobacter pylori* aislada de un paciente con cáncer gástrico”

1.2 Epidemiología de gastritis asociadas con *H. pylori*

H. pylori se encuentra en el estómago del humano en todas las partes del mundo. En los países en desarrollo, del 70 al 90% de la población es portadora; casi todos ellos adquieren la infección antes de cumplir los 10 años. En los países desarrollados, la prevalencia de la infección es menor, entre 25 a 50%. (Eusebi, L. H. *et al.*, 2014).

H. pylori causa más del 90% de las úlceras duodenales y hasta en 80% de las gástricas o estomacales, ocupando el cuarto lugar como causa de morbilidad. En México durante el 2005 se notificó el menor número de casos de gastritis como se puede apreciar en el gráfico 1, teniéndose un total de 1, 433, 592 casos registrados en el Sistema Nacional de Vigilancia Epidemiológica (SINAVE) y se presentó una incidencia de 1, 347 casos nuevos por 100 000 habitantes. Durante el 2015 se reportaron 1, 428, 661 casos de úlceras, gastritis y duodenitis. Los grupos con mayor riesgo fueron el género femenino con 920, 036 casos mientras que el masculino tuvo un total de 508, 625 casos (Boletín epidemiológico. SINAVE, 2016).



*Incidencia por 100 000 habitantes
FUENTE: SUIVE/DGAE/Secretaría de Salud
GRÁFICA 1. Casos e incidencia de Úlceras, gastritis y duodenitis
México, 2003-2008.
FUENTE: SUIVE/DGAE/Secretaría de Salud

1.3 Transmisión

No hay nicho de *H. pylori* que no sea el estómago humano, la forma de transmisión de cómo es que se desplaza desde el estómago de una persona al de otra se desconoce, para la población en general

“Resolución de huecos encontrados durante la secuenciación de una cepa mexicana de *Helicobacter pylori* aislada de un paciente con cáncer gástrico”

el modo más probable de transmisión es de persona a persona, se han descrito tres rutas posibles de su transmisión:

- La primera y la menos común es iatrogénica, endoscopios, o especímenes en contacto con la mucosa gástrica de una persona se introducen a otra persona. (Morris B., 2000)
- La segunda ya sea por la vía oral-oral (a través de vómito o posiblemente saliva). La detección de DNA de *H. pylori* en el vómito, saliva, placa dental, jugo gástrico, y las heces, apoya esta teoría. (Morris B., 2000)
- La tercera vía fecal-oral. El modo de persona a persona de transmisión es apoyado por la incidencia mayor de la infección entre los niños y adultos. Transmisión por el agua, es probable debido a la contaminación fecal, especialmente en partes del mundo donde el agua no es tratada. (Morris B., 2000)

En resumen, a pesar de que *H. pylori* está presente en el estómago de aproximadamente la mitad de la población mundial, todavía no se entiende claramente cuál es su mecanismo para que se lleve a cabo la transmisión (Eusebi *et al.*, 2014).

1.4 Asociación con enfermedades

Los pacientes con infección por *H. pylori* desarrollan inflamación crónica gástrica, conocida como gastritis, pero esta condición generalmente es asintomática.

Antes la forma idiopática de la gastritis, era la úlcera péptica que representa el 60 a 95% de todos los casos; ahora sabemos que *H. pylori* es el agente causal de casi todos estos casos en adultos y que el tratamiento que la erradique conduce a curar las úlceras.

También está fuertemente asociado con el riesgo de desarrollo de gastritis crónica atrófica donde la mucosa se adelgazada, con disminución de glándulas; infiltración linfocitaria y plasmocitaria en todo el espesor de la lámina propia, acompañada de neutrófilos; puede haber también una metaplasia intestinal: el epitelio y de las glándulas está reemplazado principalmente por células caliciformes y células cilíndricas similares a las células de función absorción del intestino (enterocitos), que es una lesión precursora del cáncer gástrico (Cao, Q., *et al.*, 2007).

1.5 Características microbiológicas

1.5.1 Morfología

H. pylori es un microorganismo Gram negativo, curvo, posee múltiples flagelos en uno de sus polos de 5 a 6 lo que lo hace altamente móvil. Es un microorganismo de crecimiento lento. Su característica bioquímica más sobresaliente es la abundante producción de la enzima ureasa, que cataliza la hidrólisis de la urea en amonio y bióxido de carbono; la producción de amonio es un mecanismo importante para la sobrevivencia de la bacteria en un ambiente con pH tan ácido debido al jugo gástrico (Syuro L. K. *et al.*, 2011).



Figura 1. *H. pylori*

Bacilo espiral, Gram negativo, flagelos polares, tamaño 3µm. Fotografía obtenida a un aumento de 1, 000, 000x por microscopía electrónica (URL 1)

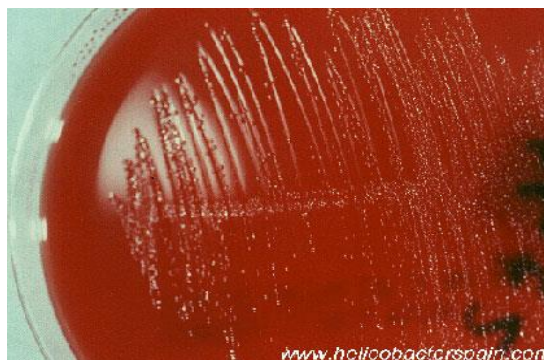


Figura 2. Colonias de *H. pylori*

Medio geloso sangre + suero de carnero, presentan una morfología puntiforme con aspecto brillante y transparente de 1 a 2 mm de diámetro. (URL 1)

1.5.2 Cultivo bacteriano

Las cepas se siembran en medios de gelosa sangre de carnero. Las placas se incuban en condiciones de microaerofilia (CO₂ 9%, a 37°C) durante 3 a 7 días.

La identificación del microorganismo se hace a través de su morfología colonial, tinción de Gram y características bioquímicas (catalasa positiva, oxidasa positiva y ureasa positiva) (Bayona R, M., *et al.*, 2013).

1.6 Virulencia y patogenicidad de *H. pylori*

Las enzimas metabólicas que posee pueden ser utilizadas por *H. pylori* para producir daño en los tejidos y para defenderse de las condiciones adversas del ambiente en el que debe sobrevivir. Pero además, *H. pylori* posee otras características que le permiten colonizar la mucosa gástrica del estómago, inducir daños en los tejidos o liberarse de los mecanismos de defensa del huésped. Entre las características de virulencia podríamos destacar:

La estructura espiral. La propia estructura de la bacteria le permite introducirse a través de la capa de moco gástrico actuando de forma similar a un sacacorchos y favoreciendo por lo tanto el acercamiento a las células epiteliales gástricas.

La movilidad. *H. pylori* posee de 4 a 6 flagelos polares que le confieren una gran movilidad y le permiten llegar a la mucosa y no ser eliminado por los mecanismos defensivos del huésped.

Las adhesinas. Posee una gran variedad de adhesinas que reconocen de forma específica a los receptores de la mucosa gástrica y se unen a ellos comenzando la colonización bacteriana (Magalhães and Reis, 2010)

La toxina vacuolizante. Se ha descrito la presencia de una toxina que produce la formación de grandes vacuolas en las células eucarióticas. Este efecto lo producen más de la mitad de los aislamientos clínicos de *H. pylori* y aquellos que son inducen la vacualización se han asociado con cuadros más graves de enfermedad. La toxina está codificada por un gen denominado *vacA* que está presente en todos los aislamientos. (Atherton, 1995)

El gen A asociado a la citotoxina (*cagA*). Se encuentra localizado en una región del cromosoma que se conoce como Isla de patogenicidad (*cagPAI*).

La proteína CagA. Se ha descrito la presencia de una proteína codificada por el gen *cagA*. La presencia de esta proteína induce la respuesta inflamatoria y aumento de la expresión de interleucina-8. Algunos autores han observado una clara diferencia en cuanto al proceso de enfermedad que se produce cuando se trata de una cepa *cagA+* o *cagA-* (Hatakeyama and Higashi, 2005).

La producción y actividad de la enzima ureasa. Esta enzima hidroliza la urea generando una gran cantidad de iones amonio los cuales modifican el pH del medio ambiente que rodea a la bacteria, facilitando su permanencia y diseminación (Harris *et al.*, 1996).

“Resolución de huecos encontrados durante la secuenciación de una cepa mexicana de *Helicobacter pylori* aislada de un paciente con cáncer gástrico”

1.7 Genoma

En agosto de 1997 se publicó la secuencia completa del genoma de *H. pylori* ATCC 26695 que fue aislada en el Reino Unido de un paciente con gastritis y secuenciado; el cromosoma circular de la cepa contiene 1, 667, 867 pb. En contraste, *H. pylori* J99, que se secuenció en un esfuerzo de colaboración entre Astra AB (ahora AstraZeneca PLC) y Genome Therapeutics Corporation, aislada en los Estados Unidos en Enero de 1999 a partir de un paciente con úlcera duodenal su cromosoma circular contiene 1, 643, 831 pb. Aunque el genoma de la 26695 fue más grande con respecto de la J99, tanto la J99 como la 26695 poseen un contenido de G+C del 39% (Jean-F. Tomb, *et al.*, 1997; Richard A. Alm, *et al.*, 1999).

El genoma de la cepa 26695 muestra 117 genes que no están presentes en la cepa J99, mientras que la cepa J99 contiene 89 genes que no están presentes en la cepa 26695 (Tabla 1). Se piensa que la diversidad genética es a causa de un elevado índice de mutaciones puntuales, cambios en los marcos de lectura, re-arreglos del genoma, transferencias horizontales y recombinación entre cepas (Hacker *et al.*, 1997). Algunos estudios indican que *H. pylori* destaca como la especie bacteriana con la más alta tasa de recombinación poblacional, la cual se lleva a cabo cuando se producen infecciones crónicas mixtas que persisten por muchos años desde su establecimiento en la mucosa gástrica (Suerbaum and Josenhans, 2007). Otro factor que puede contribuir es la presión selectiva que ejercen las condiciones gástricas (Kusters *et al.*, 2006).

Tabla 1. Características genómicas comparativas de *H. pylori*

Genome feature	J99	26695
Tamaño (pb)	1,643,831	1,667,867
(G+C)%	39	39
Regiones con diferente contenido en (G+C)%	9	8
Marcos abiertos de lectura		
Porcentaje del genoma codificantes	90.8	91.0
Secuencias codificantes	1,496	1,590
Con función descrita	877	899
Sin función descrita	361	375
Específicos de <i>H. pylori</i>	258	279

“Resolución de huecos encontrados durante la secuenciación de una cepa mexicana de *Helicobacter pylori* aislada de un paciente con cáncer gástrico”

Al igual que con la mayoría de los proyectos de secuenciación de genoma microbiano hasta la fecha, J99 y 26695 fueron secuenciados utilizando bibliotecas de fragmentos cromosómicos creados al azar mediante clonación (Richard A. Alm, *et al.*, 1999).

En los últimos años, ha habido una explosión de secuencias genómicas de datos generados a partir de una amplia gama de organismos. Con 75 cepas de *H. pylori* que ya se han completado los genomas microbianos ahora publicados y más de 100 proyectos de genómica en curso.

H. pylori cepa CG-IMSS-2012 se secuenció utilizando la plataforma 454 FLX (Roche, Alemania), la generación de una biblioteca que contiene 309, 265 lecturas simples con una longitud media de 209 pb y una cobertura promedio de 35 veces. Esta estrategia proporciona 47 contigs grandes (45 de ellos con más de 200 bp). Este proyecto de secuenciación del genoma de *H. pylori* CG-IMSS-2012 tiene un tamaño de aproximadamente 1, 599, 050 pb y un contenido de G + C de 39.0%. El genoma reveló un 91.5% de las regiones codificantes y 1, 613 genes (Méndez-Tenorio, A. *et al.*, 2014).

1.8 Secuenciación de DNA

Son los procedimientos destinados a determinar la secuencia nucleotídica, en una cadena de DNA. Se han utilizado dos técnicas para la secuenciación de DNA, uno el método de Alan Maxam y Walter Gilbert de hidrólisis química específica y, por otra, la interrupción controlada con didesoxinucleótidos de Frederic Sanger (Teijón, 2006).

La secuenciación de nueva generación (NGS) es una tecnología que ha desplazado a la ahora llamada tecnología de primera generación, tal como la secuenciación automatizada de Sanger (Martin *et al.*, 1985; Sanger and Coulson, 1975; Zimmermann *et al.*, 1988). La tecnología NGS combina diferentes estrategias, como la preparación del material genético que se desea analizar, la secuenciación como tal y el ensamblado genómico; y ofrece la posibilidad de generar una gran cantidad de datos (en algunos casos de hasta mil millones de lecturas por corrida) en menor costo y tiempo. Por lo que más allá de solo determinar el orden de una secuencia, esta tecnología tiene el potencial de permitir la comparación de genomas completos o de mejorar nuestro entendimiento de cómo las diferencias genéticas afectan la salud y la enfermedad. (Mardis, 2008; Metzker, 2010)

1.8.1 Preparación del material genético

El primer paso en un proyecto de NGS es la fragmentación del DNA genómico en tamaños más pequeños que, dependiendo de la plataforma elegida, pueden ser usados como molde en una fase

“Resolución de huecos encontrados durante la secuenciación de una cepa mexicana de
Helicobacter pylori aislada de un paciente con cáncer gástrico”

de amplificación. Por lo general, las secuencias molde son incluidas en una superficie o soporte sólido que las mantiene inmobilizadas y separadas entre sí. Esto permite que miles o miles de millones de reacciones de secuenciación se realicen simultáneamente. (Metzker, 2010)

1.8.2 Secuenciación

Posterior a la preparación de la muestra se efectúa la reacción de secuenciación. Existen cuatro principales metodologías de secuenciación: secuenciación por terminación reversible cíclica, secuenciación por ligación, secuenciación en tiempo real y secuenciación por adición de un solo nucleótido.

1.8.3 Ensamblado genómico

El ensamblado genómico es un proceso que se realiza posteriormente a la secuenciación, y puede definirse como la superposición de las lecturas de secuenciación para reconstruir el genoma original. El ensamblado es necesario dada la imposibilidad de los secuenciadores para determinar la identidad de un genoma completo en una sola corrida. (Brown, 2013)

El ensamblado genómico es un proceso jerárquico en el que se pueden distinguir los siguientes elementos:

- **Lectura de secuenciación.** Es el elemento más sencillo en la jerarquía y se obtiene directamente de los datos de la secuenciación.
- **Contig.** Es el siguiente nivel de la jerarquía y lo podemos definir como una secuencia consenso obtenida a partir del alineamiento de múltiples lecturas de secuenciación. La principal característica de esta secuencia es que no presenta un orden definido en el genoma.
- **Scaffold.** Este nivel comprende la suma de dos o más contigs, que podrán denominarse scaffolds siempre que presenten un orden y orientación coherente con la estructura del genoma original. En este nivel aún pueden presentarse huecos.

“Resolución de huecos encontrados durante la secuenciación de una cepa mexicana de *Helicobacter pylori* aislada de un paciente con cáncer gástrico”

- **Cromosoma.** Es el último nivel jerárquico del ensamblado de genomas, y se llega a él cuando se logran ordenar y orientar todos los datos de secuenciación y se logran cerrar los huecos.

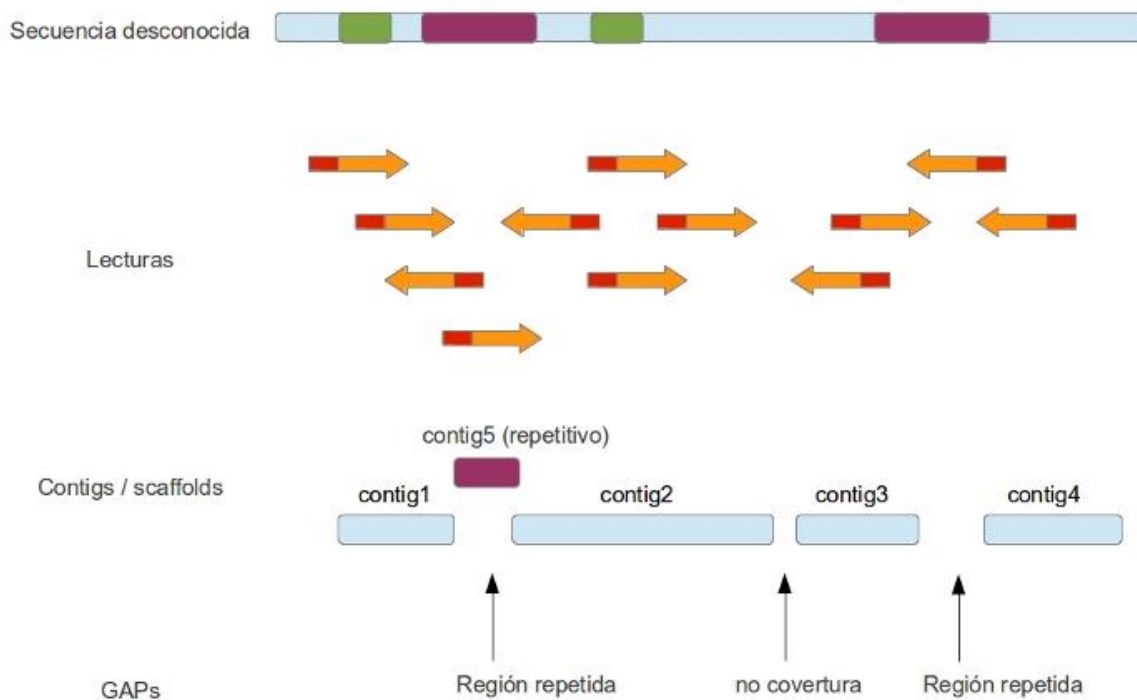


Figura 3. Secuenciación de DNA

Todas las tecnologías de secuenciación producen fragmentos más o menos cortos y a partir de estos se debe recomponer la secuencia del fragmento analizado. Este proceso se denomina proyecto de secuenciación. En muchos de los casos el resultado no será la secuencia completa del fragmento y será necesario completar los huecos entre los diferentes conjuntos de secuencias obtenidos (URL 2).

1.8.4 Anotación genómica

El ensamblado está relacionado con la anotación genómica, por lo que puede incluirse como un paso posterior a un proyecto de NGS.

La anotación puede definirse como un proceso para estructurar el ensamblado genómico, identificar los modelos de genes y características genómicas, además de proveer una conexión entre los elementos del genoma y su significado biológico (Brown, 2013).

2. Justificación

La infección con *Helicobacter pylori* es la causa principal del desarrollo de gastritis crónica, enfermedad ulcero-péptica y es un factor de riesgo para el desarrollo de cáncer gástrico. Conocer el genoma de *H. pylori* aislada de un paciente Mexicano representa un punto de referencia para conocer las características genéticas de las cepas que prevalecen en nuestro país, por lo que es importante obtener el menor número de huecos en la secuencia obtenida.

“Resolución de huecos encontrados durante la secuenciación de una cepa mexicana de *Helicobacter pylori* aislada de un paciente con cáncer gástrico”

3. Hipótesis

Mediante un análisis bioinformático de los 45 huecos, se identificaron que genes correspondían a los extremos y el posible orden de los contigs lo que nos permitió tratar de completar el genoma de *Helicobacter pylori*. Si el análisis bioinformático es correcto, entonces los huecos encontrados podrían resolverse mediante reacciones de PCR y secuenciación automática, lo que nos acercaría más a completar su genoma.

“Resolución de huecos encontrados durante la secuenciación de una cepa mexicana de *Helicobacter pylori* aislada de un paciente con cáncer gástrico”

4. Objetivo general

Resolver el mayor número de huecos posibles encontrados durante el ensamblado del genoma de la cepa de *Helicobacter pylori* CG-IMSS-2012.

4.1 Objetivos particulares

1. Diseñar primers específicos para amplificar los fragmentos que sirvan para cerrar los huecos mediante su secuenciación.
2. Amplificar mediante PCR los fragmentos que unan a las secuencias problema.
3. Purificación de los fragmentos amplificados mediante PCR y someterlos a una secuenciación automática.
4. Ensamblado de las secuencias resultantes con los contigs correspondientes.
5. Realizar la anotación funcional de secuencias genómicas mediante un servidor.

5. Materiales y métodos

El diagrama de general de trabajo desarrollado en la presente investigación se muestra en la figura 4.

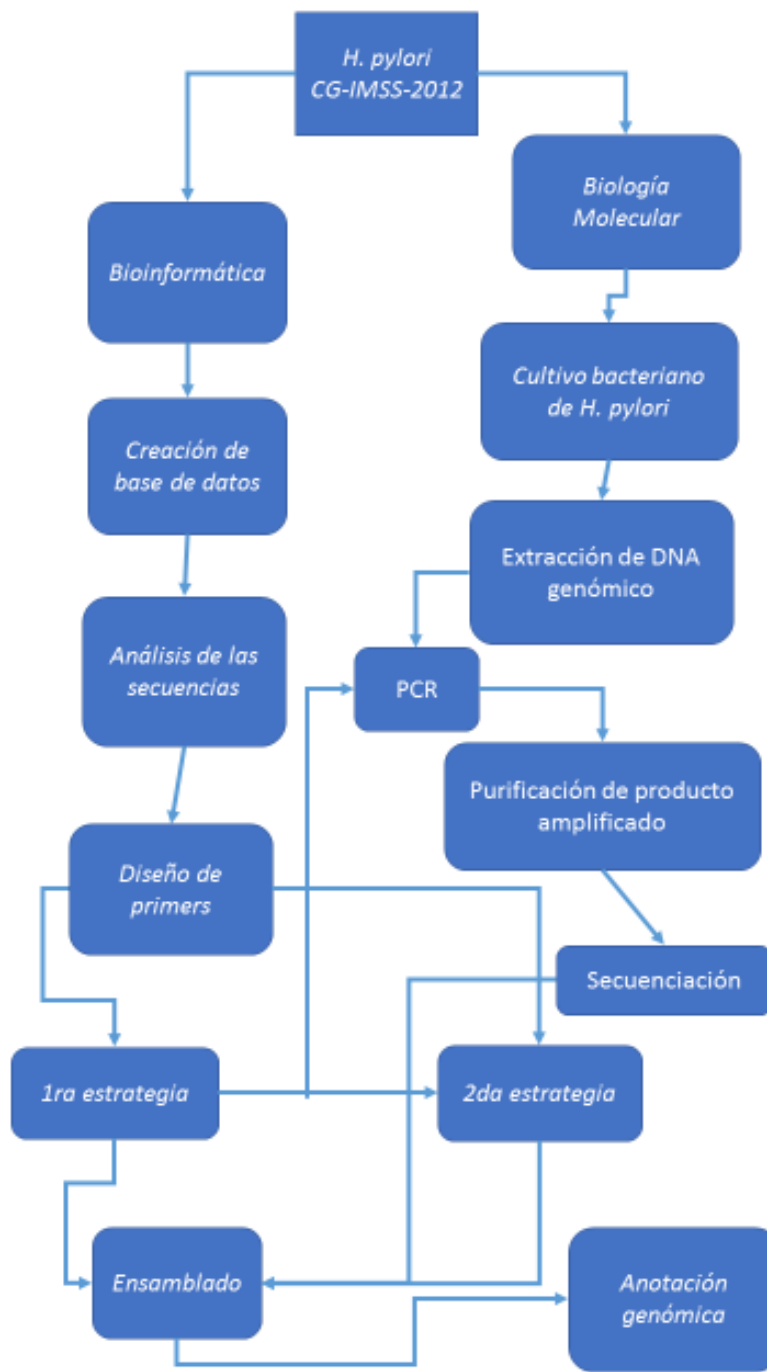


Figura 4. Diagrama general de trabajo

“Resolución de huecos encontrados durante la secuenciación de una cepa mexicana de
Helicobacter pylori aislada de un paciente con cáncer gástrico”

5.1 Extracción de DNA

Con un hisopo estéril se cosecharon las colonias de *H. pylori*, el cultivo se incubo en un periodo de 4 días a 37°C en condiciones de microaerofilia, se cosechó en 1 ml PBS donde se le dio a las células un tiempo de lavado.

El lavado se realizó al centrifugar a 28300 g / 5 min, después se retiró el PBS, y se le adiciono 500 µL del reactivo GES a la pastilla, se re suspendió usando la micropipeta. Se incubó a 5 min a temperatura ambiente.

Posteriormente se adicionó un volumen de 250 µL acetato de amonio 7.5 M frío (mantiene las condiciones osmóticas, favorece la precipitación de la fracción proteica), se incubo por 5 min en hielo.

Después del tiempo se adicionó un volumen de 650 µL de cloroformo y se agitó suavemente y se continuó con una centrifugación a 28300 g / 5 min. El cloroformo tiene la función de desnaturalizar las proteínas y dar una separación de dos fases una acuosa (se encuentra nuestro DNA) y la segunda una orgánica (proteínas y restos celulares).

La fase superior (acuosa) se recuperó con cuidado se pasó a un tubo Eppendorf de 1 ml donde se le adicionó un volumen de 400 µL de isopropanol, se mezclaron suavemente y dejo precipitando toda la noche a -20 °C.

Transcurrido este tiempo se realizó nuevamente una centrifugación a 28300 g / 5 min, para continuar con un lavado con etanol, se le adicionó un volumen de 800 µL de etanol al 70% que fue preparado al momento. Se centrifugó por 5 min a máxima velocidad y se decantó. El etanol al 70 % permite una mayor purificación o elución del DNA de la sal residual que va junto con el isopropanol así como los residuos de este.

El DNA finalmente se re suspendió en 100 µL de agua bidestilada libre de DNAsas.

“Resolución de huecos encontrados durante la secuenciación de una cepa mexicana de *Helicobacter pylori* aislada de un paciente con cáncer gástrico”

5.2 Diseño de primers

De la de la página del Centro Nacional para la Información Biotecnológica (National Center for Biotechnology Information; NCBI) se descargó los 45 contigs en archivos fasta, lo cuales contienen la secuencia de *H. pylori* CG-IMSS-2012, mediante la unión entre los extremos de los contigs se buscó completar el genoma, esto mediante el llenado de los huecos que existen entre ellos.

5.2.1 Diseño de primers primera estrategia

Mediante un análisis bioinformático *in silico* que previamente se llevó a cabo un estudiante de maestría obtuvo una lista de 5 genes que codifican para diferentes proteínas que comparten los extremos de algunos de los contigs.

En la tabla 2 se muestra el resultado del análisis, los contigs y las proteínas con las que se alinean en los extremos de cada uno de ellos, así como número de acceso de la NCBI de las proteínas y a la cepa de *H. pylori* a la cual pertenece.

Tabla 2. Lista de huecos con los contigs para llevar a cabo la secuencia, así como las proteínas usadas y a la cepa pertenece

Contig 3'	Contig 5'	No. Acceso	Proteína	Cepa
contig002_3'	contig034_5' rc	NP_222885	Enzima de Restricción	J99
contig016_3'	contig017_5'	YP_005424778	α 1, 3-fucosiltransferasa	ELS37
contig018_3'_RC	contig023_5'	NP_207343	cag26	26695
contig026_3'	contig024_5'	YP_005424959	Topoisomerasa I	ELS37
contig025_5' rc	contig020_5'	YP_005424915	cag7	ELS37
contig020_5'	contig023_3'	YP_005424915	cag7	ELS37

Con el programa DNA STAR – SEQ MAN v5.03; se alinearon las secuencias de los genes de las proteínas con los extremos de los contigs, se revisó el tamaño de los huecos. En la figura 5 observamos el alineamiento entre los extremos de los contigs con la secuencia del gen, lo cual nos permitió determinar el tamaño del hueco que existe entre los contigs.

“Resolución de huecos encontrados durante la secuenciación de una cepa mexicana de *Helicobacter pylori* aislada de un paciente con cáncer gástrico”

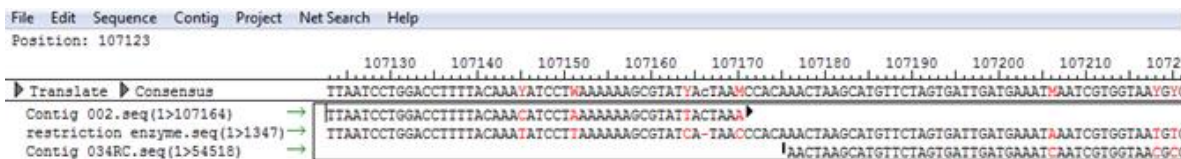


Figura 5. Alineamiento entre los contigs seleccionados de la tabla 5 y la secuencia del gen que codifica para la proteína usada como referencia.

Se muestra el alineamiento que resulto entre los extremos de los contigs con la secuencia que codifica para las proteínas que se mencionan en la tabla 5 y se aprecia el espacio que se buscó completar el cual debe quedar en medio al momento del diseño

Con los alineamientos obtenidos, se procedió a diseñar los juegos de primers, el alineamiento se cargó en la página de la NCBI, en la sección de Data & Software – PRIMER BLAST. Se hicieron los ajustes necesarios para la amplificación de la zona de interés: como que el forward y el reverse quedaran en posiciones estratégicas para que amplificaran de la zona de interés quedara en una posición central, se escogieron primers que tuvieran como resultado un producto de un tamaño aproximado 600 y 750 pares de bases (pb).

5.2.2 Diseño de primers segunda estrategia empleando cepa de referencia *H. pylori* ELS37

En una segunda metodología para el diseño de nuevos primers donde se trató de tener nuevos huecos que secuenciar y tratar de completar el genoma de *H. pylori* CG – IMSS – 2012.

Para lo cual se realizaron alineamientos de los contigs de la *H. pylori* CG–IMSS–2012 con el genoma de *H. pylori* ELS37 la cual es una cepa del Salvador la cual tiene un acercamiento filogenético de nuestra cepa mexicana.

Se determinó el orden de los contigs con el alineamiento de nuestra cepa con la otra cepa de *H. pylori* ELS37. Para realizar los alineamientos se utilizó el programa de BLAST Assembled Genomes de la página de la NCBI y con un segundo programa SEQ MAN v5.03 – DNASTAR se realizó otro alineamiento para corroborar el primero.

De esta manera se escogieron 5 nuevos candidatos a los cuales se les diseñaron un juego de primers específicos para cada uno de los huecos encontrados. El diseño se realizó como se describe en el punto anterior.

“Resolución de huecos encontrados durante la secuenciación de una cepa mexicana de *Helicobacter pylori* aislada de un paciente con cáncer gástrico”

5.3 Reacción en cadena de la polimerasa (PCR)

Los reactivos para la PCR (excepto el DNA) se mezclaron en un solo tubo de acuerdo a la Tabla 3, a un volumen de 15 μL por una prueba de un juego de primers; cada primers se probó por triplicado para corroborar que amplificaran correctamente.

Tabla 3. Mezcla maestra de PCR

Reactivos	[Inicial]	[Final]	1X μL
H ₂ O			6.725
dNTP's	10 mM	0.2 mM	0.3
Buffer PCR	5X	1X	3
MgCl ₂	25mM	1.5 mM	0.9
Primers F	50 μM	5 μM	1.5
Primers R	50 μM	5 μM	1.5
Taq polimerasa	5 u/ μL	0.025 u/ μL	0.075
DNA			1
Volumen Final			15 μL

Se desarrolló el programa de amplificación con las temperaturas adecuadas para cada juego de primers y tiempos para los ciclos en el termociclador.

Tabla 4. Ciclo de amplificación de DNA

Ciclos	Temp/ Tiempo	Pasos PCR
40	94 °C / 5 min	Desnaturalización del DNA
	94 °C / 1 min	
	50 °C / 30 s	Alineamiento
	72 °C / 1 min	Extensión
72 °C / 15 min		

“Resolución de huecos encontrados durante la secuenciación de una cepa mexicana de
Helicobacter pylori aislada de un paciente con cáncer gástrico”

5.4 Gel de agarosa y electroforesis

Se preparó geles de agarosa al 1.5% en 30 ml de agua. Se pesó la agarosa, y después se le adicionó un volumen de agua, esto se calentó para disolver la agarosa, se dejó enfriar un momento y se le adicionó regulador TAE a una concentración final de 1x en 600 μ L y GelRed a una concentración final de 1X en 1 μ L.

Nota: Revisar el apéndice para ejemplo de calculos de un gel.

Se agitó suavemente para evitar la formación de burbujas de aire y se procedió a vaciar sobre la cámara electroforesis ya montada con las barreras y el peine adecuado, donde se dejó enfriar para que se polimerice. Luego se le adiciono buffer TAE al 1X de corrimiento en la cámara electroforética.

Ya que se llenó la cámara electroforética con el buffer, se cargaron los productos de PCR en los pozos y se anotó el orden en que corresponde a cada uno de ellos; además se cargó un marcador de peso molecular y los controles correspondientes tanto negativos como positivos.

Electroforesis

Es una técnica para la separación de moléculas según la movilidad de éstas en un campo eléctrico. En la fuente de poder se establecieron las condiciones de la electroforesis fue un voltaje de 70 a 90 volts (V) durante un tiempo de 45 - 60 min, dependiendo del tamaño de los fragmentos.

Se dejó correr el tiempo adecuado, después se capturó la imagen empleando el software Quantum Cap que es el documentador que se empleó para su visualización bajo luz UV y hacer una fotografía de los productos.

“Resolución de huecos encontrados durante la secuenciación de una cepa mexicana de *Helicobacter pylori* aislada de un paciente con cáncer gástrico”

5.5 Purificación de productos de PCR a partir del gel de agarosa con el kit QIAquick gel Extraction (QIAGEN).

Este protocolo está diseñado de acuerdo a las indicaciones del fabricante para extraer y purificar el DNA de 70 pb a 10 kb de geles de agarosa en regulador TAE.

Se cortó la banda del gel que contenía nuestros productos de PCR usando el transiluminador para la visualización de las bandas. Se colocó en un tubo Eppendorf de 1.5 ml, y después se pesó el segmento de agarosa (tarando antes con un tubo vacío). El peso del segmento de agarosa se expresó en mg, esta cantidad de agarosa equivale a un volumen de los reactivos del Kit QIAquick.

Posteriormente se adicionaron 3 volúmenes de buffer QG y se incubó a 50 °C /10 min y mezclando en intervalos de 2 min para disolver el gel de agarosa.

Finalmente se verificó que el gel se haya disuelto completamente, se comprobó que el pH sea de 7.5 (color amarillo).

Para precipitar el DNA, se adicionó un volumen de isopropanol y se mezcló. Se centrifugó a 11900 g / 1 min.

La solución se transfirió a la columna de separación proporcionada por el kit de QIAquick y se centrifugó a 11900 g / 1 min.

Para eliminar las trazas remanentes de agarosa, se adicionó nuevamente un volumen de 500 µL buffer QG y se centrifugó a 11900 g / 1 min.

Finalmente se le adicionó un volumen de 750 µl de regulador PE, y se incubó por 5 min / T° A y se continuó con una centrifugación 11900 g / 1 min. El PE es un regulador de lavado para eliminar sales presentes en DNA.

Se prosiguió nuevamente con una centrifugación a 11900 g / 1 min para eliminar el exceso de etanol que contiene el buffer PE.

Por último se obtuvo el DNA al adicionar 40 - 50 µl de agua inyectable, se incubó 5 min / T °A y se terminó con una centrifugación a 11900 g /1 min.

“Resolución de huecos encontrados durante la secuenciación de una cepa mexicana de *Helicobacter pylori* aislada de un paciente con cáncer gástrico”

5.6 Secuenciación de los fragmentos

Se procedió a realizar una nueva mezcla maestra de PCR para un volumen final de 50 μL por reacción de los primers que tuvieron una amplificación exitosa, la mezcla se hizo como se indica en la siguiente tabla 5:

Tabla 5. Mezcla maestra de PCR

	[Stock]	[Final]	1X μL	1X VF 50 μL
H ₂ O			9.425	31.35
dNTP's	10 mM	0.2 mM	0.3	1
Buffer PCR	5X	1X	3.0	10
MgCl ₂	25mM	1.5 mM	0.9	3
Taq	5 u/ μL	0.05 u / μL	0.15	0.25
PrimerF	5 μM	0.5 μM	1.5	0.5
Primer R	5 μM	0.5 μM	1.5	0.5
DNA			1	3.4
Volumen Final				50 μL

Se realizó de la misma forma para cada juego de primers, se observó que la amplificación del fragmento de PCR fuera exitosa.

Cada fragmento, se corrió en un gel de agarosa y se purifico directamente del gel.

Una vez purificado, se realizó un gel de los fragmentos purificados, se capturo la imagen y se envió al Instituto de Fisiología Celular (IFC), UNAM, donde se llevó a cabo la secuenciación.

5.7 Ensamblado de los fragmentos secuenciados

El resultado de la secuenciación se obtiene en formato *ab1* y posteriormente se cambió a formato *seq.*, cada secuencia se identificó con el nombre de la cepa: la cepa – tamaño – contigs en un formato *.ab1*; a partir de este archivo se cambió su formato a *.seq* con el programa (Editseq de DNASTar v5.03) para hacer el ensamblado del fragmento correspondiente con los extremos de los contigs de los huecos a llenar el cual se hizo con el programa (SeqMan de DNASTar v5.03) se guardó una imagen y un archivo en formato *.sqd* y FASTA, para tener el alineamiento de la unión de los contigs.

“Resolución de huecos encontrados durante la secuenciación de una cepa mexicana de *Helicobacter pylori* aislada de un paciente con cáncer gástrico”

5.8 Anotación genómica

Se realizó una anotación genómica de las secuencias ensambladas, mediante el servidor Rapid Annotation using Subsystem Technology (RAST).

Para anotar el genoma en el servidor RAST se cargó el archivo *.fasta* de la secuencia ensamblada en la página <http://rast.nmpdr.org/>, posteriormente se completó la información solicitada acerca del organismo, la primera parte se describe: identificación taxonómica, dominio, especie, cepa, y código genético.

We have analyzed your upload and have computed the following information.

Contig statistics

Statistic	As uploaded	After splitting into scaffolds
Sequence size	1599050	1599050
Number of contigs	45	45
GC content (%)	39.0	39.0
Shortest contig size	243	243
Median sequence size	30576	30576
Mean sequence size	35534.4	35534.4
Longest contig size	107164	107164

Please enter or verify the following information about this organism:

- RAST bases its genome identifiers on NCBI taxonomy-IDs.
- If you provide a valid taxonomy-ID, RAST will attempt to fill in the genome metadata for you.
- If you leave the taxonomy-ID field blank, RAST will assign a meaningless taxonomy-ID, and you will need to fill in the
- If you plan on submitting this genome to [PATRIC](#) you will need to provide the most descriptive NCBI taxonomic group identifier and the genome will not be suitable for submission to PATRIC. We discuss the motivation and process for such
- You may search for the taxonomy-ID of your organism using the search facilities at the [NCBI taxonomy browser](#).

Genome information:

Taxonomy ID: [Look up taxonomy ID at NCBI.](#)

Taxonomy string:

Domain: Bacteria Archaea Virus

Genus:

Species:

Strain:

Genetic Code: 11 (Archaea, most Bacteria, most Virii, and some Mitochondria)
 4 (Mycoplasmata, Spiroplasmata, Ureoplasmata, and Fungal Mitochondria)

[Use this data and go to step 3](#)

Figura 6. Formulario de información del organismo en un proyecto de anotación en el servidor RAST

En la segunda sección se activan las opciones de: corrección automática de errores y relleno de huecos, así como las instrucciones de depuración, y se dejó activada la opción de corrección de marcos de lectura y la ausencia de una o más bases.

“Resolución de huecos encontrados durante la secuenciación de una cepa mexicana de *Helicobacter pylori* aislada de un paciente con cáncer gástrico”

Please consider the following options for the RAST annotation pipeline:

RAST Annotation Settings:		
Choose RAST annotation scheme	<input type="text" value="Classic RAST"/>	Choose "Classic RAST" for the current production RAST, or "RASTtk" for the new modular RAST pipeline currently in testing.
Select gene caller	<input type="text" value="RAST"/>	Please select which type of gene calling you would like RAST to perform. Note that using GLIMMER-3 will disable automatic error fixing, frameshift correction and the backfilling of gaps.
Select FIGfam version for this run	<input type="text" value="Release70"/>	Choose the version of FIGfams to be used to process this genome.
Automatically fix errors?	<input checked="" type="checkbox"/> Yes	The automatic annotation process may run into problems, such as gene candidates overlapping RNAs, or genes embedded inside other genes. To automatically resolve these problems (even if that requires deleting some gene candidates), please check this box.
Fix frameshifts?	<input type="checkbox"/> Yes	If you wish for the pipeline to fix frameshifts, check this option. Otherwise frameshifts will not be corrected.
Build metabolic model?	<input type="checkbox"/> Yes	If you wish RAST to build a metabolic model for this genome, check this option.
Backfill gaps?	<input checked="" type="checkbox"/> Yes	If you wish for the pipeline to blast large gaps for missing genes, check this option.
Turn on debug?	<input type="checkbox"/> Yes	If you wish debug statements to be printed for this job, check this box.
Set verbose level	<input type="text" value="0"/>	Set this to the verbosity level of choice for error messages.
Disable replication	<input type="checkbox"/> Yes	Even if this job is identical to a previous job, run it from scratch.

Figura 7. Opciones para el procedimiento de anotación en servidor RAST

“Resolución de huecos encontrados durante la secuenciación de una cepa mexicana de *Helicobacter pylori* aislada de un paciente con cáncer gástrico”

6. Resultados

6.1 Diseño de primers específicos.

Después de los alineamientos con el programa SEQ MAN v5.03, se diseñaron los primers específicos con el programa Primer-BLAST de la NCBI, para cada hueco se buscó variar el tamaño de los fragmentos para una identificación más practica; en la siguiente tabla 6 se muestra los primeros primers diseñados para los contigs de la tabla 5.

Tabla 6. Diseño de primers de los contigs con la primera estrategia

Num.	Contigs	Primer	Secuencia (5' → 3')	Longitud	Inicio	Final	Temp. media	GC%
1	016	Forward	TCACGCTCTCGGCTTTATGG	23	275	297	59.57	47.83
	017	Reverse	TTAAAACAGCACGCCACGC	21	840	820	58.62	47.62
	Tamaño (pb)		515					
2	002	Forward	CGGCTTAATCAGTAATAGCAGGC	23	275	297	59.57	47.83
	034	Reverse	CGCTCGTATTCATTGTAGCGA	21	840	820	58.62	47.62
	Tamaño (pb)		566					
3	026	Forward	AGGCTTTACCTGAAAATGAGAAAGT	25	228	252	58.93	36.00
	024	Reverse	AAGAGCAAGTCTTGTTTGCATAA	24	858	835	57.06	33.33
	Tamaño (pb)		631					
4	018	Forward	CCCTTTTGATACAGCACGCCT	21	324	344	60.95	52.38
	023	Reverse	ATCATTGGCTAAAGCAACGGG	21	886	866	59.25	47.62
	Tamaño (pb)		563					
5	025-020	Forward	ATGGTATAGAGTTAATGAAATTGCAGAA	25	228	252	58.93	36.00
	020-023	Reverse	CTT GCG GAT CGT TGC TAT CTT TTA	24	858	835	57.06	33.33
	Tamaño (pb)		5788					

Con la segunda estrategia de alineamiento entre la cepa de *H. pylori* CG-IMSS-2012 con la ELS37, se obtuvieron 5 candidatos en el orden de los contigs que se pueden observar en la tabla 7, estos sirvieron para el diseño de nuevos juegos de primers.

"Resolución de huecos encontrados durante la secuenciación de una cepa mexicana de *Helicobacter pylori* aislada de un paciente con cáncer gástrico"

Tabla 7. Diseño de primers de los contigs con la segunda estrategia

Núm.	Contigs	Primer	Secuencia (5' → 3')	Longitud	Inicio	Final	Temp. media	GC%
1	Forward	014	ATGCGTGGCTTGGTCTATGA	20	3474	3493	59.46	50.00
	Reverse	015	CACCCAGCCTAAAGACAGGAAA	22	4173	4152	60.22	50.00
	Tamaño (pb)		700					
2	Forward	031	TGGATGATACCCCTATTGAAGAAA	24	3076	3099	57.16	37.50
	Reverse	032	GCGCCGTTACAATATAGACATCC	23	3810	3788	59.63	47.83
	Tamaño (pb)		735					
3	Forward	032	CCTTGATTTTTTCGCGCTCCC	21	3072	3092	60.20	52.38
	Reverse	033	GAGCGGTGCTTGAATTGGTG	20	3783	3764	60.11	55.00
	Tamaño (pb)		712					
4	Forward	033	CCTTTCCTTAAATCAAGCTCTGTG	24	3016	3039	58.05	41.67
	Reverse	034	GAGAACCTGGTATTCGCTGCT	21	3745	3725	60.13	52.38
	Tamaño (pb)		730					
5	FORWARD	039	TATAACCCAGGGGTGAGCAAC	21	3139	3159	59.44	52.38
	REVERSE	040	GCACCCGTGATTGCACTAGG	20	3860	3841	61.37	60.00
	TAMAÑO (PB)		722					

6.2 Amplificación mediante PCR de los fragmentos que unan a las secuencias problema.

Se hizo una amplificación de los huecos con los primers mediante la técnica de PCR. En las siguientes figuras se puede apreciar los geles de agarosa y la visualización de las bandas únicas con ayuda del transiluminador Quantum Cap.

En la figura 8 podemos observar la imagen obtenida por el transiluminador QUANTUM CAP, del primer grupo de primers solo logro la amplificación de 3 de los 5 productos esperados, cada uno con una banda única y correspondientes al tamaño esperado.

“Resolución de huesos encontrados durante la secuenciación de una cepa mexicana de *Helicobacter pylori* aislada de un paciente con cáncer gástrico”

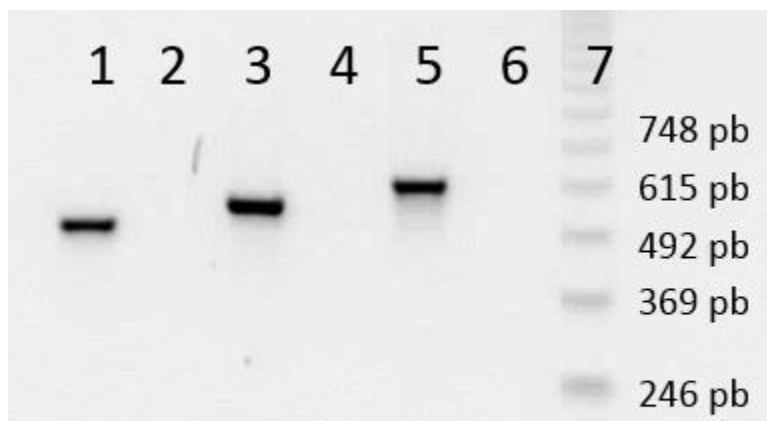


Figura 8. Productos de PCR *H. pylori* CG-IMSS-2012 de primeros primers

(1) *H. pylori* 515 pb contigs 016-017 (3) *H. pylori* 566 pb contigs 002-034 (5) *H. pylori* 631 pb contigs 026-024 (7) Marcador 123 pb

En la figura 9 obtenida por el transiluminador QUANTUM CAP, se muestra el resultado de los primers realizados con la segunda estrategia de diseño, se observa que los productos obtenidos por la PCR presentan una banda única y correspondientes al tamaño esperado, en este caso solo amplificaron 4 de los 5 primers.

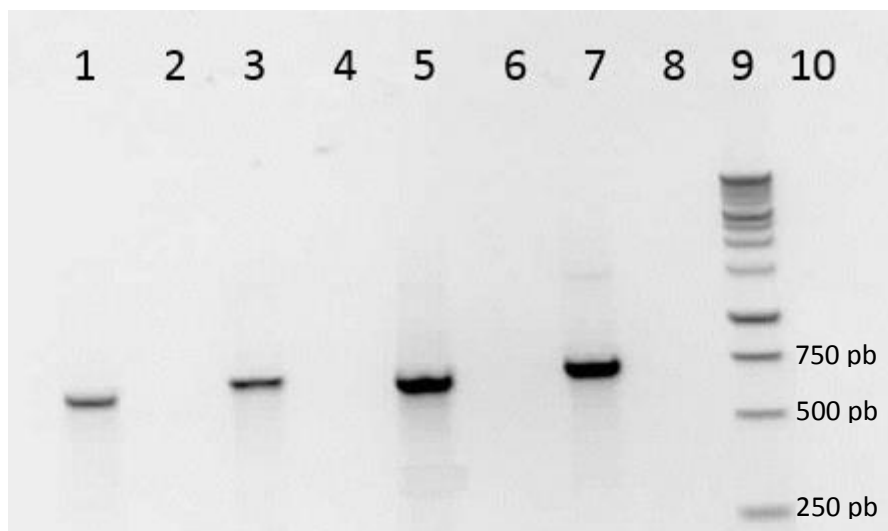


Figura 9. Productos de PCR *H. pylori* CG-IMSS-2012 de segundos primers

Productos amplificados por PCR de
1. *H. pylori* 700 pb contigs 014-015 3. *H. pylori* 735 pb contigs 031-032
5. *H. pylori* 712 pb contigs 032-033 7. *H. pylori* 722 pb contigs 039-040
9. Marcador 1 kb

“Resolución de huecos encontrados durante la secuenciación de una cepa mexicana de *Helicobacter pylori* aislada de un paciente con cáncer gástrico”

6.3 Purificación y secuenciación automática.

Los productos obtenidos mediante la PCR se cortaron del gel de agarosa y se purificaron para obtener solamente el DNA, el cual se mandó a secuenciar en el IFC donde se registraron con un nombre descrito que fuera característico para cada uno donde se utilizó: la cepa – tamaño de producto – núm. de contig, como por ejemplo: 286 – 631 – 26F, los productos fueron sometidos a una secuenciación automática. El resultado de cada secuenciación se envió un archivo correspondiente en formato .ab1 con el nombre que se les dio. En las figuras siguientes 10, 11 y 12 se muestra el archivo de las secuencias obtenidas con los primeros primers diseñados.

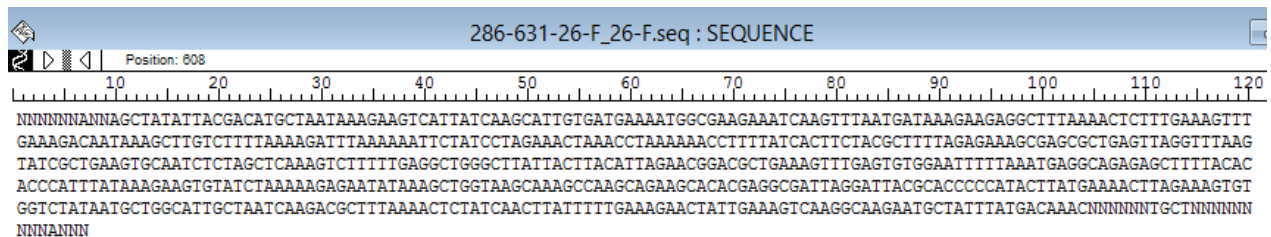


Figura 10. Secuencia de nucleótidos del fragmento 286-631-26F para el hueco entre los contigs 026 - 024.

Secuencia obtenida a través del diseño de primers específicos y amplificación mediante PCR.

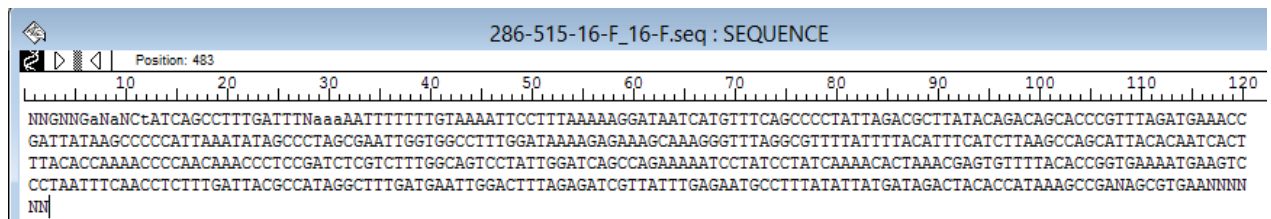


Figura 11. Secuencia de nucleótidos del fragmento 286-515-16F para el hueco entre los contigs 016 - 017.

Secuencia obtenida a través del diseño de primers específicos y amplificación mediante PCR.

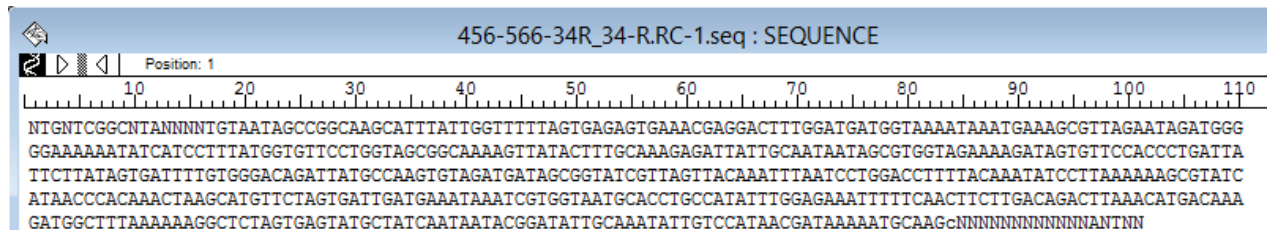


Figura 12. Secuencia de nucleótidos del fragmento 456-566-34R para el hueco entre los contigs 002 - 034.

Secuencia obtenida a través del diseño de primers específicos y amplificación mediante PCR. 26

“Resolución de huecos encontrados durante la secuenciación de una cepa mexicana de *Helicobacter pylori* aislada de un paciente con cáncer gástrico”

Los primers que fueron diseñados con la segunda estrategia se amplificaron mediante PCR, los cuales se cortaron del gel de agarosa para ser purificados y mandarlos secuenciar en el IFC donde se registraron con un nombre descrito anteriormente, en las siguientes figuras 13, 14, 15 y 16 se muestran las secuencias obtenidas.

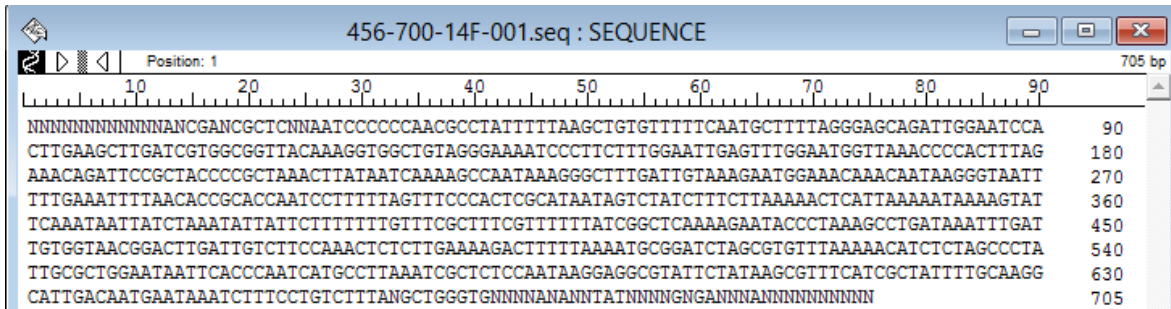


Figura 13. Secuencia de nucleótidos del fragmento 456-700-14F para el hueco entre los contigs 014 - 015.

Secuencia obtenida a través del diseño de primers específicos y amplificación mediante PCR con un tamaño de 705 pb.

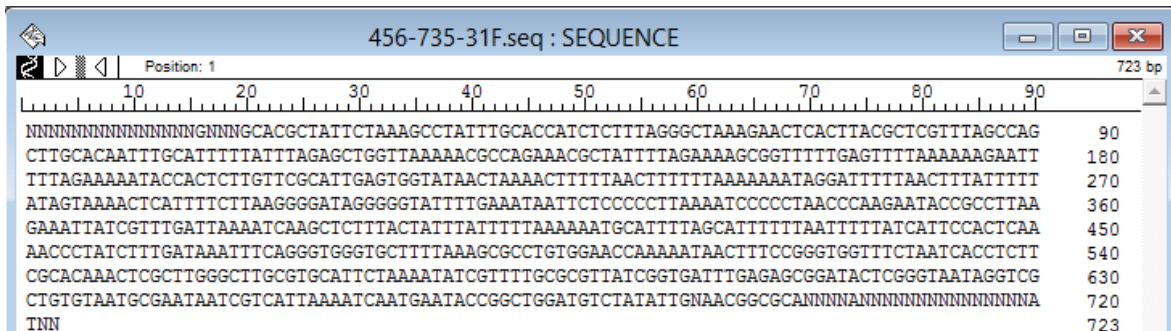


Figura 14. Secuencia de nucleótidos del fragmento 456-735-31F para el hueco entre los contigs 031- 032.

Secuencia obtenida a través del diseño de primers específicos y amplificación mediante PCR con un tamaño de 723 pb.

“Resolución de huecos encontrados durante la secuenciación de una cepa mexicana de *Helicobacter pylori* aislada de un paciente con cáncer gástrico”

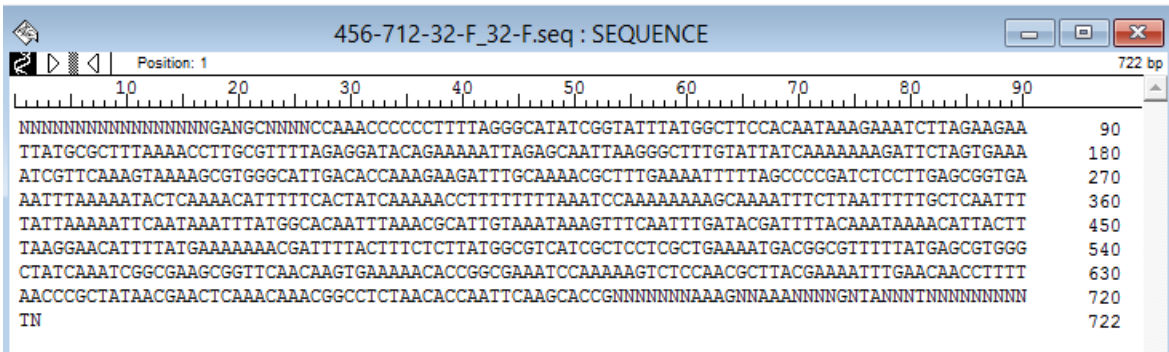


Figura 15. Secuencia de nucleótidos del fragmento 456-712-32F para el hueco entre los contigs 032 - 033.

Secuencia obtenida a través del diseño de primers específicos y amplificación mediante PCR con un tamaño de 722 pb.

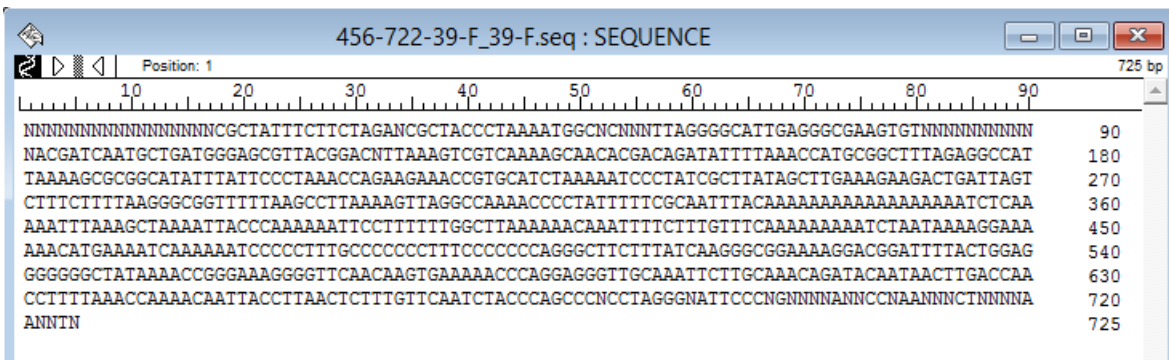


Figura 16. Secuencia de nucleótidos del fragmento 456-722-39F para el hueco entre los contigs 039 - 040.

Secuencia obtenida a través del diseño de primers específicos y amplificación mediante PCR con un tamaño de 725.

6.4 Ensamblado de las secuencias.

Con las secuencias obtenidas se realizó el alineamiento con los contigs correspondientes para completar los huecos que hacían falta y de esta manera llenar el espacio existente entre los contigs y así obtener 7 scaffold. En las siguientes figuras se observan dichos ensamblajes.

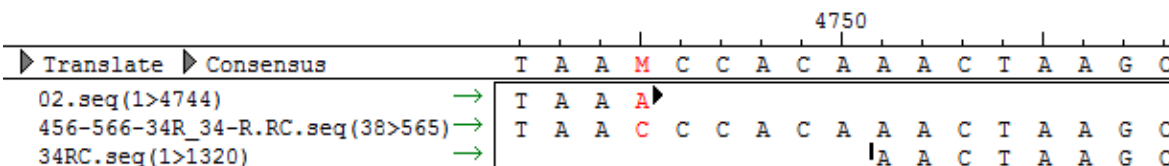


Figura 17. Ensamblado del fragmento 456-566-34R secuenciado con los contigs 002 - 034.

Se muestra el alineamiento entre los contigs 002 – 034 y el fragmento secuenciado 456-566-34R para cerrar el espacio que existe entre ellos conociendo la secuencia de bases correspondientes.

“Resolución de huecos encontrados durante la secuenciación de una cepa mexicana de *Helicobacter pylori* aislada de un paciente con cáncer gástrico”

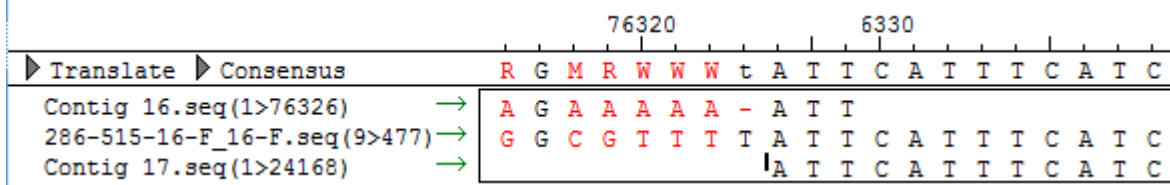


Figura 18. Ensamblado del fragmento 286-515-16F secuenciado con los contigs 016 - 017.
Se muestra el alineamiento entre los contigs 016 - 017 y el fragmento secuenciado 286-515-16F para observar en este alineamiento como corrobora que estos Contigs se traslapan por lo que se puede unir.

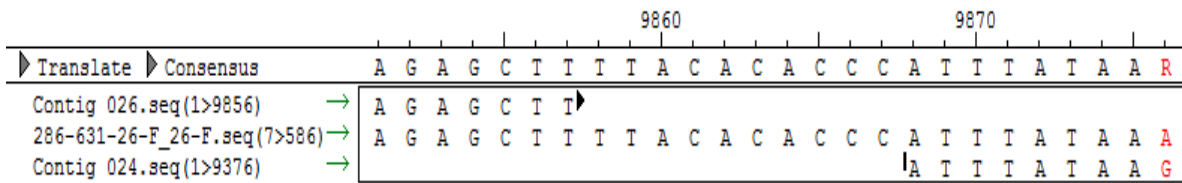


Figura 19. Ensamblado del fragmento 286-631-26F secuenciado con los contigs 026 - 024.
Se muestra el alineamiento entre los contigs 026 – 024 y el fragmento secuenciado 286-631-26F para cerrar el espacio que existe entre ellos conociendo la secuencia de bases correspondientes.

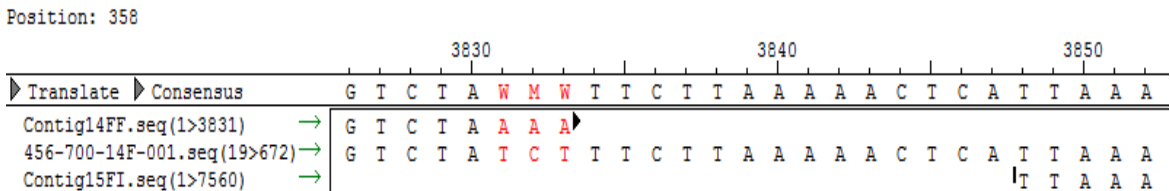


Figura 20. Ensamblado del fragmento 456-700-14F secuenciado con los contigs 014 - 015.
Se muestra el alineamiento entre los contigs 026 – 024 y el fragmento secuenciado 456-700-14F para cerrar el espacio que existe entre ellos conociendo la secuencia de bases correspondientes.

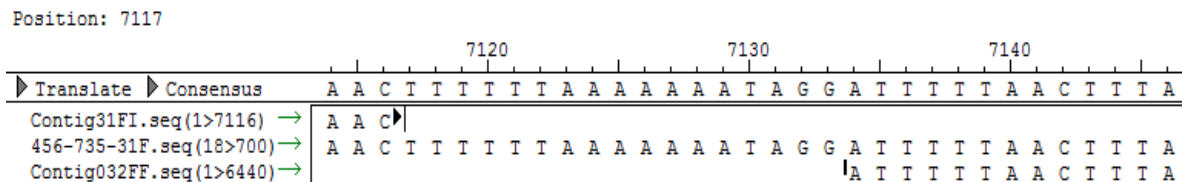


Figura 21. Ensamblado del fragmento 456-735-31F secuenciado con los contigs 031 - 032.
Se muestra el alineamiento entre los contigs 031 – 032 y el fragmento secuenciado 456-735-31F para cerrar el espacio que existe entre ellos conociendo la secuencia de bases correspondientes.

“Resolución de huecos encontrados durante la secuenciación de una cepa mexicana de *Helicobacter pylori* aislada de un paciente con cáncer gástrico”

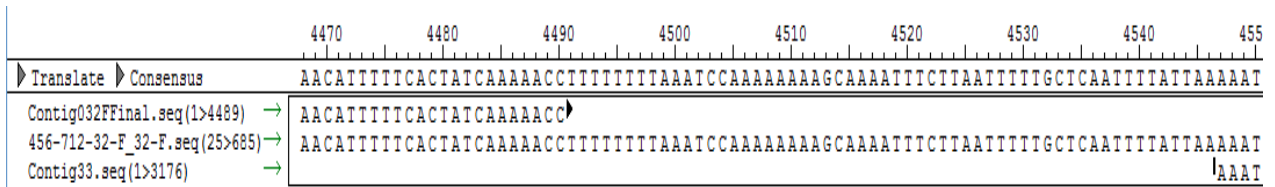


Figura 22. Ensamblado del fragmento 456-712-32F secuenciado con los contigs 032 - 033.

Se muestra el alineamiento entre los contigs 032 – 033 y el fragmento secuenciado 456-712-32F para cerrar el espacio que existe entre ellos conociendo la secuencia de bases correspondientes.

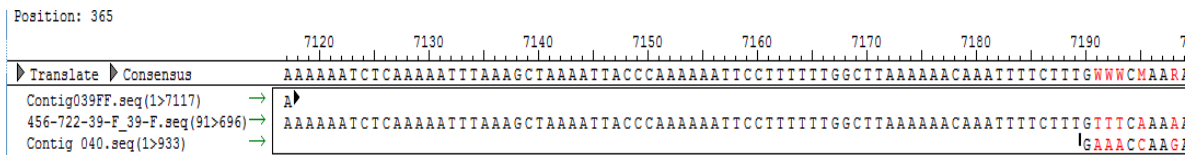


Figura 23. Ensamblado del fragmento 456-722-39F secuenciado con los contigs 039 - 040.

Se muestra el alineamiento entre los contigs 039 – 040 y el fragmento secuenciado 456-722-39F para cerrar el espacio que existe entre ellos conociendo la secuencia de bases correspondientes.

6.5 Anotación genómica

RAST está diseñado para anotar los genes de un genoma completo o casi completo de bacterias. La información obtenida del servidor RAST dio la anotación de los genes de algunos scaffolds obtenidos, el detalle es que solo una parte del gen corresponde a alguno de los extremos de los contigs, mientras que el espacio existente entre el extremo del otro contig no tiene una función aún descrita.

En las imágenes resultados de la anotación podemos observar diferentes barras la primera se puede observar el tamaño del genoma, la segunda flechas azules se describe el funcionamiento del gen, luego sigue una barra gris que es la fuente, siguiendo con los aminoácidos codificados según la secuencia del genoma descrito.

“Resolución de huecos encontrados durante la secuenciación de una cepa mexicana de *Helicobacter pylori* aislada de un paciente con cáncer gástrico”

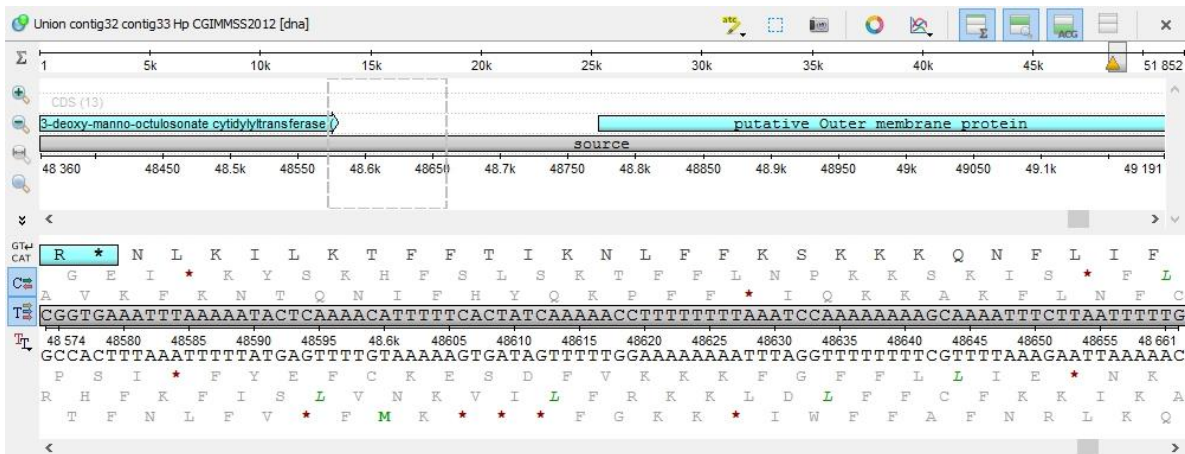


Figura 26. Anotación genómica resultante de la unión de los contigs 32-33.

La información obtenida en la búsqueda del gen nos dio la secuencia para la enzima 3-dioximano-octulosonato-citidiltransferasa



Figura 27. Anotación genómica resultante de la unión de los contigs 39-40.

La información obtenida en la búsqueda del gen nos dio la secuencia para el sistema de transporte sideróforo férrico, en banda periplásmica.

“Resolución de huecos encontrados durante la secuenciación de una cepa mexicana de *Helicobacter pylori* aislada de un paciente con cáncer gástrico”

En la siguiente tabla 8 se pueden observar la lista de los scaffolds generados durante esta marcha de trabajo, donde de manera general fue el resultado de la unión de los contigs llevada a cabo con la secuencia obtenida a base de la amplificación de la zona que faltaba mediante PCR con el diseño de primers específicos para dichas zonas. Como se puede apreciar en la tabla 8 se enumeran los scaffolds siguiendo con el número de contigs con los que se unieron, el hueco que existía entre cada uno; además del tamaño del scaffold en pb, así como el gen al que pertenece.

Tabla 8. Lista de contigs unidos mediante las técnicas empleadas

Scaffold	Contig 1	Contig 2	Tamaño de hueco	Tamaño del scaffold (pb)	Proteína
Primer estrategia					
1	02	34	5 pb	161687	Enzima de Restricción
2	16	17	3 pb	100494	α 1, 3-fucosiltransferasa
3	26	24	10 pb	19233	Topoisomerasa
Segunda estrategia					
4	14	15	14 pb	69582	Factor precursor de ensamblaje de proteínas de membrana externa YaeT – Proteína hipotética FIG00638667
5	31	32	17 pb	90329	ARNt guanina transglicosilasa – Región Intergenica
6	32	33	56 pb	51852	3 dioxi-mano-octulosonato-citidiltransferasa – Región Intergenica
7	39	40	72 pb	49705	Sistema de transporte sideroforo férrico, banda periplásmica – Región Intergenica

7. Discusiones

El DNA es una de las moléculas más importantes para la vida. Consiste de una doble hélice donde cada una de las cadenas es un polímero integrado por miles o incluso millones de nucleótidos (Stansfield, 1992). El orden que tienen los nucleótidos en el DNA es lo que se denomina secuencia, las técnicas y métodos que se utilizan para conocer esa secuencia son llamados de secuenciación (de Necochea y Canul 2004).

Sin embargo, los avances científicos y tecnológicos han permitido tener hoy en día una gran cantidad de secuencias de DNA con información de importancia médica, ecológica, fisiológica y evolutiva. En 1977 surge el método que revolucionó a la biología molecular, la secuenciación enzimática de Sanger (Sanger *et al.* 1977).

El ensamblado genómico es un proceso que se realiza posteriormente a la secuenciación, es la superposición de las lecturas de secuenciación para reconstruir el genoma original. El ensamblado es necesario dada la imposibilidad de los secuenciadores para determinar la identidad de un genoma completo en una sola corrida (Brown, 2013).

En 1995, año que representa un parteaguas en las ciencias genómicas, se secuenció el primer genoma completo de un organismo, de la bacteria oportunista *Haemophilus influenzae*, que causa múltiples infecciones en humano como meningitis, otitis, conjuntivitis y sinusitis (Fleischmann *et al.* 1995). El genoma de esta especie es de alrededor de 1, 800, 000 pb con más de 100 genes asociados con el proceso de infección (Fleischmann *et al.* 1995). A partir de ese año se comenzó la secuenciación de genomas, principalmente de microorganismos de importancia médica.

En agosto de 1997 se secuenció el genoma de *H. pylori* ATCC 26695. *H. pylori* J99, que se secuenció en 1999. Aunque el genoma de la 26695 fue más grande con respecto de la J99, tanto la J99 como la 26695 poseen un contenido de G+C del 39% (Jean-F. Tomb, *et al.*, 1997; Richard A. Alm, *et al.*, 1999).

El genoma de la cepa 26695 muestra 117 genes que no están presentes en la cepa J99, mientras que la cepa J99 contiene 89 genes que no están presentes en la cepa 26695. Se piensa que la diversidad genética es a causa de un elevado índice de mutaciones puntuales, cambios en los marcos de lectura, re-arreglos del genoma, transferencias horizontales y recombinación entre cepas (Hacker *et al.*, 1997). Algunos estudios indican que *H. pylori* destaca como la especie bacteriana con la más

“Resolución de huecos encontrados durante la secuenciación de una cepa mexicana de *Helicobacter pylori* aislada de un paciente con cáncer gástrico”

alta tasa de recombinación poblacional, la cual se lleva a cabo cuando se producen infecciones crónicas mixtas que persisten por muchos años desde su establecimiento en la mucosa gástrica (Suerbaum y Josenhans, 2007). Otro factor que puede contribuir es la presión selectiva que ejercen las condiciones gástricas (Kusters *et al.*, 2006)

En 2014 se publicó el borrador de la secuencia del genoma de *H. pylori* cepa CG-IMSS-2012 aislado de un hombre mexicano. Este genoma se sumó a las escasas secuencias reportadas a partir de aislados clínicos y proporcionará información valiosa para una mejor comprensión de que genes predominan (Méndez *et al.*, 2014). Lo anterior sugiere la necesidad de implementar un trabajo donde se buscó completar el genoma que quedo a nivel de contigs, siguiendo la formación del mayor número posibles de scaffolds que es el resultado de la unión de dos o más contigs (URL 3).

Para ello se ideó una primera estrategia que a partir de un trabajo de maestría se obtuvo un análisis de genes que compartían los extremos de los contigs (Torres-López, 2014). De esta manera se lograron detectar cinco probables candidatos que pudieran unir los extremos de los contigs, para ello se llevó a cabo el diseño de los primers, el diseño correcto de primers es un punto clave para que fueran estrictamente específicos para la zona de interés. Si el diseño de los iniciadores fue el correcto, la estabilidad y especificidad del complejo será eficiente (Tamay de Dios, *et al.*, 2013). Los iniciadores utilizados en esta investigación cumplían con la condición de dar como resultado un alineamiento específico, aunque experimentalmente mediante la técnica de PCR únicamente tres tuvieron una amplificación exitosa.

La principal razón por la que no pudieron detectarse más uniones mediante este procedimiento puede ser explicada dado que las secuencias de los genes ubicados en los extremos de la mayoría de los contigs no se encuentran fragmentados y por tal, no existe una secuencia complementaria. Por otro lado, también es importante considerar que los extremos de algunos contigs presentan una baja cobertura de las bases ensambladas, lo que implica que alguna lectura de secuenciación se incluyera en esa posición de forma incorrecta (Torres López, 2014).

Es por esto que se planteó una segunda estrategia donde se pretendió encontrar nuevas probables uniones, para ello se realizaron alineamientos de los contigs con el genoma de la cepa de *H. pylori* ELS37 la cual proviene del Salvador la cual tiene un acercamiento filogenético con nuestra cepa mexicana, de este análisis se obtuvo una lista de cinco probables uniones, mediante un análisis *in silico* se demostró que estos se alineaban de manera correcta y no en algún otro sitio que no fuera

“Resolución de huecos encontrados durante la secuenciación de una cepa mexicana de *Helicobacter pylori* aislada de un paciente con cáncer gástrico”

el deseado; se prosiguió a probarlos experimentalmente de las cuales cuatro fueron comprobadas experimentalmente de igual manera con la técnica de PCR (URL 4, URL 5).

Debido al buen diseño de los primers se lograron amplificar mediante PCR un total de siete fragmentos con ambas estrategias, estos se llevaron al IFC donde fueron secuenciados, la secuencia se recibió en un archivo. Con los archivos de la secuenciación, se realizaron ensamblados de los fragmentos correspondientes con los extremos de los contigs con los que se buscó completar los huecos y de manera exitosa se completó el espacio entre ellos, dando como resultado la unión de siete productos formando nuevos scaffolds y una lista de 31 contigs (Torres López, 2014).

Con los scaffolds obtenidos de la segunda estrategia del diseño de primers se continuó con la anotación genómica, utilizando el servidor RAST. La información obtenida reveló que proteína estaba en uno de los extremos de los contigs, pero al buscar que proteína continuaba, se encontró que en el scaffolds obtenido del ensamblado de los contigs 14-15 en uno de los extremos se encuentra una proteína hipotética, mientras que en los otros tres scaffolds obtenidos a lado de la proteína identificada, se encontrada una región intergenica, que son aquellas regiones genéticas que no codifican o que son de función desconocida (URL 6).

Eso explicaría porque con la primera metodología no se logró identificar genes que se compartieran entre los extremos, ya que según lo descrito en la metodología solo se tomaron 100 pb de los extremos para realizar el análisis; además de sumar las causas antes descritas. Se ha hecho comparación entre dos herramientas de anotación genómica con la cepa de *H. pylori* CG-IMSS-2012, en la anotación realizada con el NCBI Prokaryotic Annotation Pipeline se encontraron 1,613 genes, mientras que la anotación por RAST reveló 1,678 y como resultado se observa la discrepancia en cuanto al número de genes que pueden estar presente en la misma cepa, entre diferentes herramientas de anotación genómica, esto tendría significado en que no se encontrara gen alguno en estos sitios (Torres López, 2014).

Aún queda un total de 31 contigs de los cuales se desconoce su orden y la secuencia para poder unirlos, y así completar el genoma de *H. pylori* CG-IMSS-2012; por lo que se debe buscar o plantear una nueva estrategia que nos permita completar el genoma de *H. pylori*.

8. Conclusiones

Se logró el diseño de primers específicos para amplificar los fragmentos que sirvieran para unir a las secuencias. Se logró amplificar las zonas deseadas mediante PCR.

Con las secuencias obtenidas el ensamblado con los contigs correspondientes fue posible logrado cerrar 7 de los huecos existentes entre los contigs.

Se llevó acabo la anotación funcional de la secuencias genómicas.

Con las estrategias planteadas se obtuvo la formación de 7 scaffolds y se redujo el número de contigs de 45 a 31 por unir, acercándose al objetivo de completar el genoma.

Conocer el genoma de una cepa mexicana de *H. pylori* representará un punto de referencia para conocer las características genéticas de las cepas que prevalecen en nuestro país.

“Resolución de huecos encontrados durante la secuenciación de una cepa mexicana de *Helicobacter pylori* aislada de un paciente con cáncer gástrico”

9. Bibliografía

Alm, R. *et al.* (1999) **Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori***. *Nature*, 397, 176–180.

Atherton, J. (1998) ***Helicobacter pylori* virulence factors**. *British Medical Bulletin*, 54(1), 105–20.

Aziz, R. *et al.* (2008) **The RAST Server: Rapid Annotations using Subsystems Technology**. *BMC Genomics*, 9, 1–15.

Bayona R., Bact., Esp., M. Sc (2013) **Microbiological conditions for culturing *Helicobacter pylori*** Asociaciones Colombianas de Gastroenterología, Endoscopia digestiva, Coloproctología y Hepatología

Blaser, M. J. (1992) **Hypotheses on the pathogenesis and natural history of *Helicobacter pylori*-induced inflammation**. *Gastroenterology* 102:720–727

Boletín epidemiológico. Sistema Nacional de Vigilancia Epidemiológica; Sistema Único de la Información Número 11 | Volumen 33 | Semana 11 | Del 13 al 19 de marzo del 2016

Brown, S. M. (2013). **"Next-Generation DNA Sequencing Informatics"** (p. 241). New York: Cold Spring Harbor Laboratory Press.

Camacho, C. *et al.* (2009). **BLAST+: architecture and applications**. *BMC Bioinformatics*, 10, 1–9.

Cao, Q., *et al.*, (2007). **Screening of atrophic gastritis and gastric cancer by serum pepsinogen, gastrin-17 and *Helicobacter pylori* immunoglobulin G antibodies**. *Journal of digestive diseases*, 8(1), 15-22.

De Necochea R. y J. C. Canul. (2004) **Métodos fisicoquímicos en biotecnología: Secuenciación de ácidos nucleicos**. Instituto de Biotecnología, Universidad Nacional Autónoma de México.

Eusebi, L. H. *et al.* (2014) **Epidemiology of *Helicobacter pylori* Infection**. *Helicobacter*, 19 Suppl 1, 1–5.

Fleischmann R.D., M.D. Adams , O. White, R.A. Clayton, E.F. Kirkness, A.R. Kerlavage, C.J. Bult , J.

“Resolución de huecos encontrados durante la secuenciación de una cepa mexicana de
Helicobacter pylori aislada de un paciente con cáncer gástrico”

- Tomb, B.A. Dougherty y otros autores. (1995) **Whole-genome random sequencing and assembly of *Haemophilus influenzae*** Rd. Science 269: 496-512. Frankham R., J.D. Ballou y D.A. Briscoe. 2002. Introduction to Conservation Genetics. Cambridge University Press, United Kingdom
- Garcia, a. (2016) *HelicoBacterSpain.com*. [online] Helicobacterspain.com. Available at: <http://www.helicobacterspain.com> [Accessed 30 Mar. 2016].
- Hacker, J. *et al.* (1997). "**Pathogenicity islands of virulent bacteria: structure, function and impact on microbial evolution.**" Molecular Microbiology, 23(6), 1089–97.
- Harris, P. *et al.* (1996) ***Helicobacter pylori* Urease is a Potent Stimulus of Mononuclear Phagocyte Activation and Inflammatory Cytokine Production.** *Gastroenterology*, 111, 419–425.
- Hatakeyama, M. and Higashi, H. (2005) ***Helicobacter pylori* CagA: a new paradigm for bacterial carcinogenesis.** Cancer Science, 96(12), 835–43.
- Jean-F. Tomb, *et al.* (1997) **The complete genome sequence of the gastric pathogen *Helicobacter pylori*** Nature **388**, 539-547
- Kusters, J. *et al.* (2006). "**Pathogenesis of *Helicobacter pylori* infection.**" Clinical Microbiology Reviews, 19(3), 449–90.
- Linda Morris Brown (2000). ***Helicobacter pylori*: Epidemiology and Routes of Transmission** Epidemiol Rev; 22(2):283-97
- Magalhães, A. and Reis, C. (2010) ***Helicobacter pylori* adhesion to gastric epithelial cells is mediated by glycan receptors.** Brazilian Journal of Medical and Biological Research, 43(7), 611–618.
- Malfertheiner P. *et al* (2002) **European *Helicobacter pylori* Study Group (EHPSG), author Current concepts in the management of *Helicobacter pylori* infection - the Maastricht 2-2000 Consensus Report.** Aliment Pharmacol Therap. 16:167–180.
- Mardis, E. (2008). "**Next-generation DNA sequencing methods**". Annual Review of Genomics and Human Genetics, 9, 387–402.
- Marshall, B. and Goodwin, C. (1987). **Revised Nomenclature of *Campylobacter pyloridis*.** *International Journal of Systematic Bacteriology*, 37(1), 68.

“Resolución de huecos encontrados durante la secuenciación de una cepa mexicana de *Helicobacter pylori* aislada de un paciente con cáncer gástrico”

Martin, W. J. *et al.* (1985). **"Automation of DNA Sequencing: A System to Perform the Sanger Dideoxysequencing Reactions"**. *Bio/Technology*, 3, 911–915.

Méndez-Tenorio, A. *et al.* (2014) **Genome Sequence of a *Helicobacter pylori* Strain Isolated from a Mexican Patient with Intestinal Gastric Cancer** *Genome Announc.* 2(1):e01214-13. doi:10.1128/genomeA.01214-13.

Metzker, M. L. (2010). **"Sequencing technologies - the next generation"** *Nature Reviews Genetics*, 11(1), 31–46.

Miyazawa M, *et al.* (2003) **Suppressed apoptosis in the inflamed gastric mucosa of *Helicobacter pylori*-colonized iNOS-knockout mice.** *Free Radic Biol Med* 34:1621–1630.

Parker P.G., A.A. Snow, M.D. Schug, G.C. Booton y P.A. Fuerst. (1998) **What molecules can tell us about populations: choosing and using a molecular marker.** *Ecology* 79:361-382.

Pop M, Phillippy A, Delcher AL, Salzberg SL. (2004) **Comparative genome assembly.** *Brief. Bioinform.* 5:237–248. <http://dx.doi.org/10.1093/bib/5.3.237>.

Richard A. Alm, *et al.* (1999) **Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*** *Nature* 397, 176-180

Sanger, F. and Coulson, A. R. (1975). **"A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase."** *Journal of Molecular Biology*, 94, 441– 448.

Sistema Único de Información para la Vigilancia Epidemiológica (SUIVE) 2008, Dirección General de Epidemiología (DGE),

Suerbaum, S. and Josenhans, C. (2007). **"*Helicobacter pylori* evolution and phenotypic diversification in a changing host."** *Nature Reviews. Microbiology*, 5(6), 441–52.

Sycuro, L. K. *et al.*, (2011). **"Relaxation of peptidoglycan cross-linking promotes *Helicobacter pylori*'s helical shape and stomach colonization"**. *NIH-PA Author Manuscript*, 141(5), 822–833.

Tamay de Dios, L., Ibarra, C., & Velasquillo, C. (2013). **Fundamentos de la reacción en cadena de la polimerasa (PCR) y de la PCR en tiempo real.** *Investigación en discapacidad*, 2(2), 70-78.

“Resolución de huecos encontrados durante la secuenciación de una cepa mexicana de *Helicobacter pylori* aislada de un paciente con cáncer gástrico”

Warren JR. (1997) **The Discovery and Pathology of *Helicobacter pylori***. In Xth Int. Workshop on **Gastrointestinal Pathology and *Helicobacter pylori*: The Logic of Eradication**. P.7-13, Lisboa-Portugal

Warren, J. and Marshall, B. (1983). "**Unidentified curve bacilli on gastric epithelium in active chronic gastritis**". *The Lancet*, 1273–1275.

Zimmermann, J. *et al.* (1988). "**Automated Sanger dideoxy sequencing reaction protocol**". *FEBS Letters*, 233(2), 432–436.

Referencias electrónicas URL's

URL No.

1. www.helicobacterspain.com
2. https://bioinf.comav.upv.es/courses/intro_bioinf/ensamblaje.html
3. <http://genome.jgi.doe.gov/help/scaffolds.html>
4. <http://personales.upv.es/jcanizar/bioinformatica/blast.html>
5. <http://personales.upv.es/jcanizar/bioinformatica/alineamientos.html>
6. http://depa.fquim.unam.mx/amyd/archivero/08OrganizacionGenoma_19401.pdf

Libros

- Teijón José María (2006) **Fundamentos de bioquímica estructural** Pp 281-285
- Richard A. Alm and Brian Noonan (2001) ***Helicobacter pylori: Physiology and Genetics***. Capitulo El genoma
- Stansfield W.D. (1992) **Genética**. Mc Graw Hill. México

Tesis

- Torres-Lopez RC "**Anotación y análisis genómico de una cepa de *Helicobacter pylori* aislada de un paciente Mexicano con diagnóstico de cáncer gástrico**" Tesis de Maestría. ENCB-IPN. México DF, 2014.

10. Apéndice

10.1 Cálculos para preparar un gel de Agarosa

En base a nuestra cámara es la cantidad de agarosa a preparar, ej. Una “cámara chica” basta con 30 ml de Agarosa al X%

$$\frac{30(X)}{100} = Xg \text{ de Agarosa}$$

Se pesó 0.6g de agarosa lo necesario para preparar un gel al 2% y se pesaba directo en el frasco, después se le añadía un volumen de agua 29.6 ml; Se calentó aproximadamente 1 min, Se sacaba y agitaba para asegurarnos que no quedaran cristales de agarosa para asegurarnos que no se volvía a calentar otros 30 seg se dejaba que se enfriara un poco lo suficiente para que se le pudiera adicionar una solución amortiguadora de TAE: $\frac{1}{50}(30) = 0.6 \text{ mL} = 600 \mu\text{L}$ era el volumen que se le adicionaba además de Red Gel $\frac{1}{10000x}(30) = 0.003 \text{ mL} = 3\mu\text{L} \sim 1\mu\text{L}$