



Universidad de Chile
Facultad de Ciencias Sociales
Departamento de Antropología

Ancestría versus Selección:

Infección con *Helicobacter pylori* en Población Chilena.

Memoria para optar al título profesional de Antropóloga Física

Autora: Liesbeth Frías Villarroel
Profesor Guía: Dr. Sergio Flores C.

Santiago, Chile.
Junio 2010.

"Eine neue wissenschaftliche Wahrheit
pflegt sich nicht in der Weise durchzusetzen,
daß ihre Gegner überzeugt werden und sich als belehrt erklären,
sondern vielmehr dadurch, daß ihre Gegner allmählich aussterben
und daß die heranwachsende Generation von vornherein
mit der Wahrheit vertraut gemacht ist"

--Max Planck--

"The most exciting phrase to hear in science,
the one that heralds new discoveries,
is not 'Eureka!' but 'That's funny...'"

--Isaac Asimov--

"An ocean traveler
has even more vividly the impression
that the ocean is made of waves
than that it is made of water"

--Arthur S. Eddington--

ÍNDICE

I. Antecedentes	
Microbios como marcadores genéticos alternativos.....	Pp. 5-6
Genoma bacteriano.....	Pp. 6-9
Transmisión intra-familiar versus transmisión horizontal.....	Pp. 9-10
Recombinación.....	Pp. 10-13
<i>H. pylori</i> y el éxodo africano.....	Pp. 13-18
Composición genética de la población chilena: fundamentos para el estudio de <i>H. pylori</i>	Pp. 18-19
II. Problema de investigación	
Problema de investigación.....	Pp. 20
III. Hipótesis, predicciones y objetivos	
Hipótesis y predicciones.....	Pp. 20-21
Objetivos.....	Pp. 21
IV. Materiales y métodos	
Muestras.....	Pp. 22
Análisis de PCR.....	Pp. 22-23
Análisis de secuencias.....	Pp. 23-27
V. Resultados	
DNA bacteriano.....	Pp. 28
Análisis filogenético de los genes estudiados.....	Pp. 28-35
Análisis comparativo de los árboles filogenéticos en cepas de población chilena y en cuatro estructuras poblacionales (Falush <i>et al.</i> 2003).....	Pp. 36-38

Análisis comparativo de árboles filogenéticos en cepas de población chilena y de otros estudios.....	Pp. 39-42
Test de selección.....	Pp. 43
Estimación de diferenciación nucleotídica.....	Pp. 44-45
Reconstrucción de relaciones entre cepas.....	Pp. 45-46
VI. Discusión	Pp. 47-49
VII. Conclusiones	Pp. 50-51
Direcciones futuras.....	Pp. 51-52
VIII. Agradecimientos	Pp. 53
IX. Glosario	Pp. 53- 58
X. Anexo	Pp. 59-60
XI. Referencias	Pp. 61-72

ANTECEDENTES

Microbios como marcadores genéticos alternativos

Tradicionalmente, los marcadores genéticos polimórficos han sido utilizados en los múltiples intentos por reconstruir la historia evolutiva de las poblaciones humanas y sus migraciones (Cann 2001; Rosenberg *et al.* 2002; Underhill *et al.* 2000; Zaho *et al.* 2000). El cuello de botella sufrido durante la salida de África, *ca.* 50.000-70.000 años atrás (Ingman *et al.* 2001; Liu *et al.* 2006; Mellars 2006; Underhill *et al.* 2000; Zhivotovsky *et al.* 2003), redujo drásticamente la diversidad de los humanos modernos en Asia y Europa, de tal manera que la diferenciación de las poblaciones modernas requiere de marcadores genéticos con tasas evolutivas rápidas (Wirth *et al.* 2004).

En los inicios de los años setenta, la diversidad nucleotídica dentro de los segmentos hipervariables del D-loop mitocondrial (HVS1 y HVS2) fue introducida como una herramienta para resolver diferencias entre individuos cercanamente emparentados (Cavalli-Sforza y Feldman 2003; Pakendorf y Stoneking 2005). Sin embargo, se han identificado una serie de problemas con respecto a los datos obtenidos del D-loop mitocondrial; sitios que ya habían sufrido sustitución están volviendo a su estado original (sustituciones recurrentes que llevan a homoplasias), algunas mutaciones están ocurriendo en un mismo sitio en linajes independientes y las tasas de mutación de ciertos sitios presentan grandes diferencias al compararlas con otros sitios en la misma región, evidenciando la presencia de *hot spots* mutacionales (Ingman *et al.* 2000). Además de esto, los fragmentos hipervariables de DNA mitocondrial presentan sólo unos pocos cientos de pares de bases, proporcionando un número limitado de polimorfismos de nucleótido simple (SNPs) informativos para reconocer haplotipos y haplogrupos (Wirth *et al.* 2004).

Debido a estas limitaciones en humanos se implementó el uso de microsatélites (Jorde *et al.* 1997; Rosenberg *et al.* 2002; Ruiz-Linares *et al.* 1999) y SNPs, tanto en el cromosoma Y (Karafet *et al.* 2008; Ke *et al.* 2001; Underhill *et al.* 2000) como en genes autosómicos (Li *et al.* 2008; Reich *et al.* 2001; Zhao *et al.* 2000). Cada uno de estos nuevos marcadores genéticos presenta limitaciones; por ejemplo, los sitios polimórficos en genes autosómicos y en el cromosoma Y son más sencillos de estandarizar, pero se encuentran en menor proporción que en los segmentos hipervariables del DNA mitocondrial, requiriendo el uso de un mayor número de SNPs independientes (Wirth *et al.* 2004).

En ausencia de marcadores universales para poblaciones humanas, el genoma de diversos microbios ha sido investigado como una alternativa posible (Disotell 2003; Wirth *et al.* 2005). El estudio de la estructura poblacional de microbios permite deducir importantes aspectos de la historia evolutiva de los hospederos con los que han coevolucionado (*Ibíd.*). Sin embargo, la eficacia de éstos como marcadores genéticos depende de varios factores, entre ellos, su modo de transmisión (vertical vs. horizontal), la fuerza de la relación hospedero-parásito (endosimbiosis, comensalismo, parasitismo) y su forma de propagación (epidémica vs. endémica) (Holmes 2004). Algunos de estos estudios se han enfocado en el uso de virus, como el virus JC que infecta el tejido renal (Agostini *et al.* 1997) y el virus de papiloma humano (Chan *et al.* 1992), virus RNA, como el virus humano T-linfotrópico (Miura *et al.* 1994), hongos (Fisher *et al.* 2001), ratas (Matisoo-Smith *et al.* 1998) y la bacteria *Helicobacter pylori* (de Thé 2007; Falush *et al.* 2003; Linz *et al.* 2007; Moodley *et al.* 2009; Wirth *et al.* 2004).

La diversidad genética de *Helicobacter pylori* es casi dos veces la de la región hipervariable HVS1 del DNA mitocondrial, siendo probablemente uniforme a lo largo de su genoma entero, al contrario de esta última, cuya diversidad se encuentra restringida a una región de algunos cientos de pares de bases (*Ibíd.* 2004). El uso de esta bacteria ha probado aportar mayor información, que los marcadores genéticos

clásicos, sobre la historia reciente de las poblaciones humanas, deduciendo incluso migraciones poblacionales en los últimos 5.000 años (Devi *et al.* 2007; Wirth *et al.* 2004) (Fig. 1).

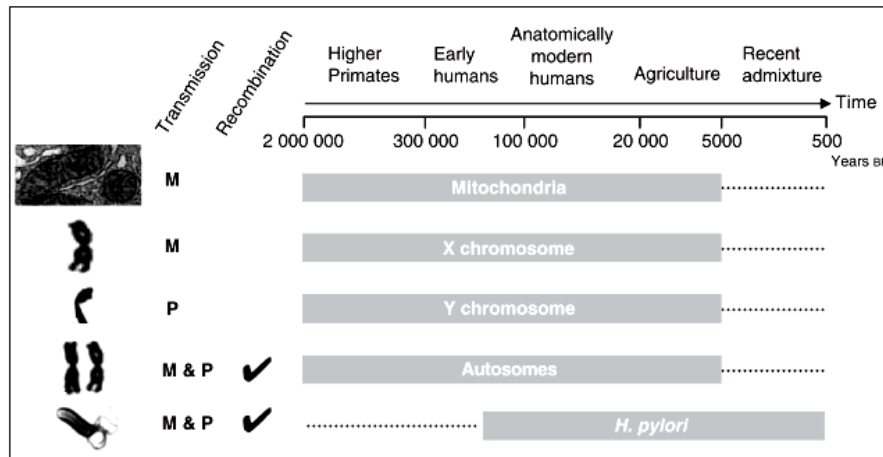


Fig. 1.- Características de los distintos marcadores genéticos para deducir migraciones humanas. Transmisión: M (materna), P (paterna) (Wirth *et al.* 2005).

Genoma bacteriano

Helicobacter pylori (*H. pylori*) es un organismo gram-negativo, con forma de espiral y flagelos (fig. 2), micro-aerófilico y de crecimiento lento, que infecta el estómago de más de la mitad de la población humana (Cover y Blaser 1996), causando gastritis crónica, úlceras peptídicas y gástricas, linfoma MALT y carcinoma gástrico (Fischbach *et al.* 2009). Sin embargo, los síntomas clínicos sólo ocurren en una minoría de la población infectada; en la mayoría de los humanos, *H. pylori* se comporta como un comensal que coloniza la mucosa gástrica (Wirth *et al.* 2004).

El genoma de *H. pylori* consiste en un cromosoma circular de un tamaño de 1.667.867 pares de bases (pb) (fig. 2) y alrededor de 1.500 genes (Tomb *et al.* 1997). Al presente se han identificado varios genes asociados a la patogenicidad en el genoma de *H. pylori*: *vacA*, *cagA* e *iceA*, de los cuales únicamente se hará mención a los dos primeros

por ser los más usados como marcadores de diversidad genética entre poblaciones (Devi 2006, 2007; Ito *et al.* 1997).

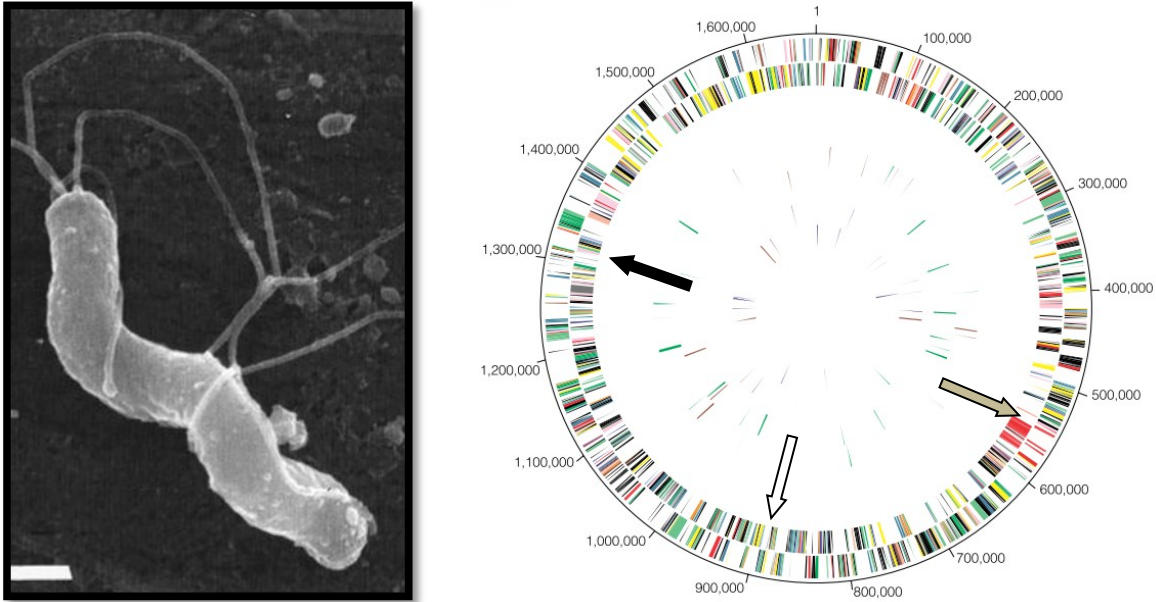


Fig. 2.- (Izq.) Morfología de *Helicobacter*. SEM, bar = 0.5 μ m. (O' Rourke y Bode 2001). (Der.) Representación circular del cromosoma de *H. pylori* 26695 (Tomb 1997): gen *cagA* (543605-547108) destacado con una flecha gris, gen *vacA* (900011-903877) con una flecha blanca y gen *trpC* (1334732-1333374) con una flecha negra.

Gen *vacA*. El gen *vacA* se encuentra presente en todas las cepas¹ de *H. pylori* y codifica una citotoxina que daña las células epiteliales del estómago al inducir la formación de vacuolas (van Doorn *et al.* 2009). Este gen contiene dos partes variables, la región de señal (s) y la región media (m). Existen al menos cuatro genotipos para la región de señal (s1a, s1b, s1c y s2) y tres para la región media (m1, m2a y m2b) (Atherton *et al.* 1995; van Doorn *et al.* 1998, 1999, 2009), los que producen proteínas con diferentes

¹ A lo largo del trabajo se hará referencia al concepto microbiológico de cepa, por el cual ésta se define como un aislado distinguible de otros del mismo género y especie por medio de características fenotípicas, genotípicas o ambas (Dijkshoorn *et al.* 2000).

grados de actividad citotóxica, donde la mayor actividad es expresada por las cepas de tipo s1/m1, siguiéndolas las de tipo s1/m2, y estando ausente en las de tipo s2/m2 (Alarcón *et al.* 1999; van Doorn *et al.* 2009).

Gen *cagA* e isla de patogenicidad *cag* (*cagPAI*). El gen *cagA* es indicador de una isla de patogenicidad (*cagPAI*), de aproximadamente 40 kpb en el genoma bacteriano (Alm *et al.* 1999; Tomb *et al.* 1997; Xiabo *et al.* 2001), que no se encuentra presente en todas las cepas de *H. pylori*. Esta isla contiene un número de genes cuyos productos están asociados a un aumento de patogenicidad en la bacteria (Akopyants *et al.* 1998; Covacci *et al.* 1997, 1999), situación que ha sido considerada como posible causa de que las cepas que presentan el gen (*cagA*-positivas) sean más prevalentes en pacientes con úlceras peptídicas, en comparación con pacientes que padecen únicamente de gastritis (*Ibíd.* 1997, 1999). La presencia del gen *cagA* se encuentra además fuertemente asociada a la del alelo s1 del gen *vacA* (van Doorn *et al.* 2009).

Genes housekeeping. Los genes *housekeeping* son genes constitutivos requeridos para la mantención de funciones celulares básicas. Los estudios realizados en base a este tipo de genes en *H. pylori* consisten en la medición de las variaciones en las secuencias nucleotídicas de un set de éstos (*atpA*, *efp*, *mutY*, *ppa*, *trpC*, *ureI*, *yphC*) y la subsiguiente caracterización de cepas mediante un perfil alélico único (MLST), lo que ha permitido deducir e identificar la estructura poblacional de cuatro poblaciones modernas (denominadas *hpAfrica1*, *hpAfrica2*, *hpEastAsia* y *hpEurope*) en base a sus distribuciones actuales (Falush *et al.* 2003).

Transmisión vertical o intra-familiar versus transmisión horizontal

La diversidad genética de *H. pylori* es alta entre cepas del mismo origen geográfico y es aún más alta a una escala global (Achtman *et al.* 1999). Estas diferencias en diversidad genética, asociadas con la biogeografía histórica entre los organismos, reflejarían por tanto las migraciones ancestrales y las modernas, a diferencia de lo que

se observa para otros patógenos, donde la transmisión horizontal es tan eficiente que estas asociaciones son raras (Roumagnac *et al.* 2006).

La transmisión para *H. pylori* ha sido descrita principalmente como vertical o intra-familiar (Kivi *et al.* 2003; Tindberg *et al.* 2001), pudiendo ocurrir entre padres e hijos y entre hermanos, donde han sido reportadas tanto cepas altamente relacionadas (Kersulyte *et al.* 1999; Li *et al.* 2002; Suerbaum *et al.* 1998) como también altamente diversas (Feldman 2001; Raymond *et al.* 2004). El período de mayor probabilidad de infección fue registrado entre el año y los cinco años de vida (Konno *et al.* 2005), período en el cual la madre es el miembro de la familia con un contacto más cercano al niño y donde el uso en común de cucharas, chupetes e incluso el masticar o probar la comida del niño, consisten en posibles vectores de transmisión (*Ibíd.*).

A pesar de que las vías de transmisión de *H. pylori* no han sido resueltas (Goosen *et al.* 2002), otras vías potenciales incluyen la feco-oral, oro-oral e incluso se ha publicado evidencia indirecta para una transmisión mediante el consumo de agua (Bunn *et al.* 2002; Delpont y van der Merwe 2007; Magalhaes Queiroz y Luzzza 2006) y de vegetales crudos regados con aguas servidas (Hopkins *et al.* 1993).

Otras complicaciones, en la inferencia de rutas de transmisión, podrían estar basadas en las interacciones hospedero-patógeno, ya que pareciera que todos los individuos coinfectados portarían una cepa dominante (Israel *et al.* 2001; Kraft *et al.* 2006; Raymond *et al.* 2004).

Recombinación

La estructuración filogenética de los genotipos globales de *H. pylori* habría surgido a través del aislamiento de la población humana antes de la migración inicial y mezcla. Las recombinaciones subsecuentes entre cepas habrían homogeneizado la estructura poblacional ancestral, pero entre cepas emparentadas habrían tendido a retenerla

como resultado de mecanismos genómicos selectivos que actuarían para limitar la recombinación entre diferentes grupos poblacionales ancestrales (Delport *et al.* 2006).

El concepto de recombinación, prevalente en la genética evolutiva, está basado en la meiosis de organismos eucariontes, donde ésta consiste en un proceso molecular complejo por el cual fragmentos de DNA son intercambiados recíprocamente entre cromosomas homólogos (*crossing-over*). Los organismos procariontes, por otro lado, presentan varias vías de recombinación (conjugación, transformación y transducción) que involucran reemplazos no recíprocos o adición de secuencias, en vez de su intercambio, tanto homólogas como no homólogas (Posada *et al.* 2002).

Pocos son los estudios que han tratado con los efectos de la recombinación en un contexto filogenético (Schierup y Hein 2000; Wiens 1998). Los métodos tradicionales de estimación filogenética, como los de máxima parsimonia (MP), mínima evolución (ME) y máxima verosimilitud (ML), asumen que una única historia evolutiva subyace a la muestra de secuencias bajo estudio. Sin embargo, distintas regiones de un alineamiento pueden tener diferentes historias evolutivas debido a procesos de *crossing-over*, conversión génica, transferencia horizontal e hibridación, que en el caso de bacterias es conocida como recombinación (Posada y Crandall 2002). Uno de los efectos que la recombinación podría causar sobre una filogenia, es la estimación de árboles que no reflejen una historia real y, por tanto, que lleven a errores en la inferencia filogenética. Posada y Crandall (2002) han demostrado, a través de una serie de simulaciones computacionales, que el efecto de la recombinación es evidente cuando las taxas involucradas son divergentes. No obstante, cuando los eventos de recombinación son ancestrales o involucran secuencias relacionadas, la historia común es frecuentemente recuperada.

Recombinación y mutación en *H. pylori*. *H. pylori* recombina frecuentemente y posee una estructura poblacional panmíctica (Go *et al.* 1996; Suerbaum *et al.* 1998), en

donde todas las cepas adquiridas por un individuo tienen la capacidad de intercambiar fragmentos de DNA entre sí. El intercambio genético ocurre por transformación de DNA entre varias cepas que colonizan a un solo individuo (Kersulyte *et al.* 1999), resultando en un conjunto de cepas relacionadas entre sí, pero que difieren en genes y alelos que han sido adquiridos por recombinación. Usando estadística bayesiana, el tamaño promedio de fragmentos recombinantes en *H. pylori* fue estimado en 500 pb con una tasa de sustitución nucleotídica de $\leq 4 \times 10^5$ bases por sitio por año (Wirth *et al.* 2005). También se extrapoló que la tasa de recombinación en *H. pylori* es una de las más altas, siendo 50 veces mayor que en *Escherichia coli* (Falush *et al.* 2001) (Fig. 3).

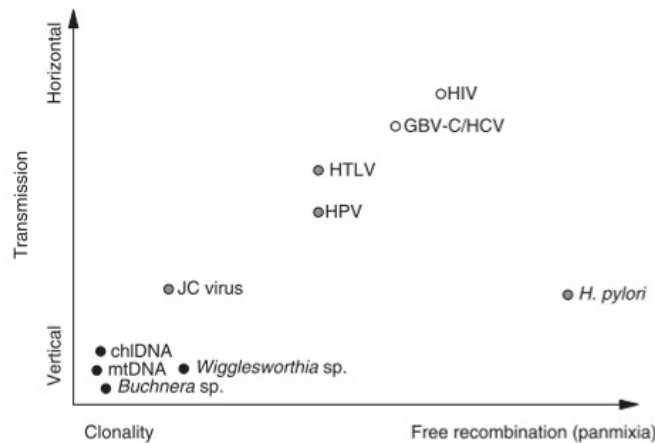


Fig. 3.- Comparación del modo y tasa de recombinación entre diferentes bacterias y especies virales usadas para inferir la historia de sus hospederos. Blanco: tasas de mutación altas; gris: intermedias; negro: bajas (*chlDNA*: DNA cloroplástico; *mtDNA*: DNA mitocondrial; HPV: virus del papiloma humano; HTLV: virus T-linfotrópico humano; GBV-C/HCV: virus de hepatitis G y C; HIV: virus de inmunodeficiencia humana) (Wirth *et al.* 2005).

Por otro lado, el tamaño promedio de fragmentos importados en *H. pylori*, provenientes de organismos no relacionados, es inusualmente pequeño para una bacteria, lo que podría estar reflejando la digestión de DNA en el ambiente gástrico y, probablemente, la presencia de múltiples endonucleasas de restricción en el genoma

bacteriano (Tomb *et al.* 1997), entre otros factores. Tanto la alta frecuencia de intercambios como el bajo número de fragmentos importados, sugieren la ausencia de mecanismos génicos propios de la bacteria que restrinjan el intercambio de DNA entre organismos no relacionados; mecanismos posiblemente innecesarios en *H. pylori*, debido a su ubicación protegida bajo una capa de mucosa en un ambiente aislado, estéril y sin competencia microbiana (Schreiber *et al.* 2003).

***H. pylori* y el éxodo africano**

En el año 2003, Falush y su equipo asignaron 380 cepas de *H. pylori*, provenientes de 27 agrupaciones geográficas, étnicas y/o lingüísticas, a cuatro poblaciones modernas designadas: hpAfrica1, hpAfrica2, hpEastAsia y hpEurope. Análisis posteriores dividieron a hpEastAsia en las sub-poblaciones hspAmerind, hspEAsia y hspMaori, y a hpAfrica1 en hspWestAfrica y hspSAfrica. Estos resultados confirmaron y extendieron datos previos que mostraban subdivisiones geográficas para esta bacteria (Achtman *et al.* 1999; Ghose *et al.* 2002; Yamaoka *et al.* 2002) (Fig. 4).

El equipo de Linz, en el año 2007, aumentó la muestra a un set de datos de 769 cepas provenientes de 51 orígenes étnicos distintos, asignándolas a las seis poblaciones previamente descritas y designando además dos poblaciones nuevas, hpAsia2 y hpNEAfrica. A partir del trabajo realizado, fue concluido que los patrones de distribución de la diversidad genética de *H. pylori* imitan a aquellos de su hospedero y que, tanto en *H. pylori* como en humanos, se observa una continua pérdida de diversidad genética con el aumento de la distancia geográfica del este africano (IBD) (Fig. 5). Finalmente, y mediante simulaciones, fue señalado que la bacteria se habría extendido desde el este africano alrededor de la misma escala temporal que los humanos anatómicamente modernos (Eppinger *et al.* 2006; Linz *et al.* 2007).

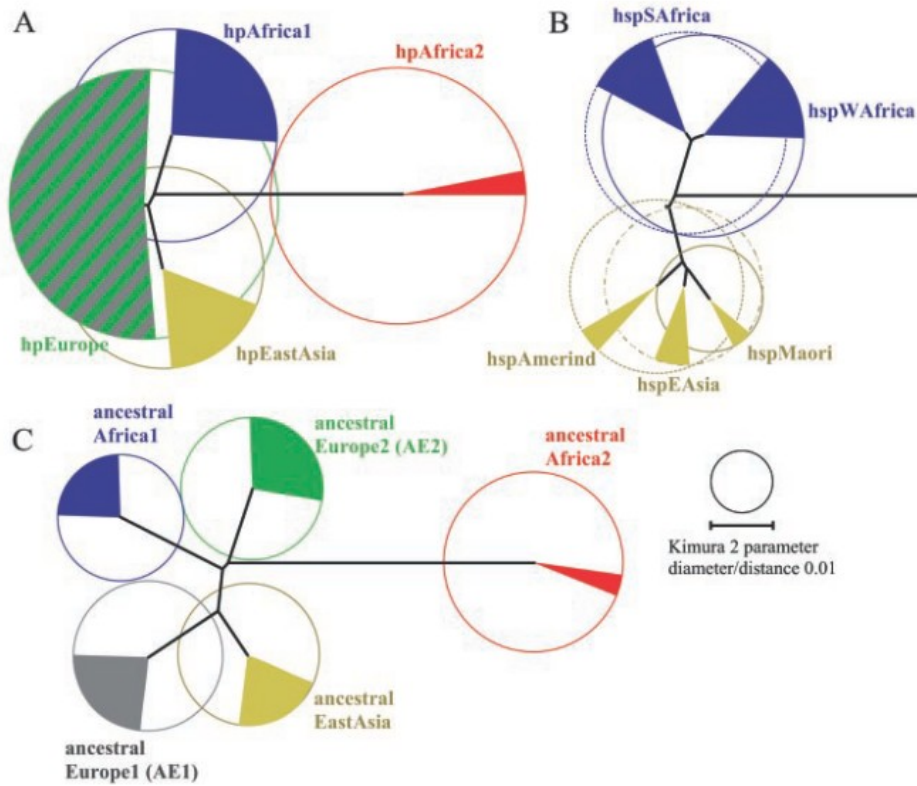


Fig. 4.- Relaciones entre poblaciones modernas (A), subpoblaciones modernas (B) y ancestrales (C) de *H. pylori*. Las líneas negras muestran árboles de Neighbor-Joining medidos por la distancia nucleotídica neta entre poblaciones. El diámetro de los círculos indica la diversidad genética, medida como la distancia genética promedio entre pares de individuos aleatorios. Los círculos más grandes reflejan los efectos de la mezcla entre poblaciones ancestrales y los arcos llenos reflejan el número de cepas en cada población (Falush et al. 2003).

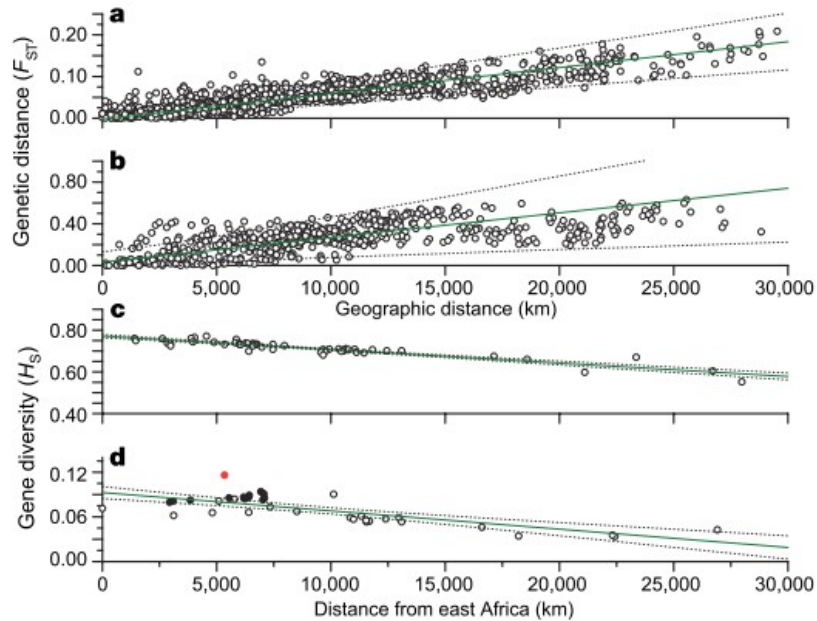


Fig. 5.- Patrones geográficos paralelos de la diversidad genética en humanos y *H. pylori*. **(a)** Distancia genética en humanos **(b)** y en *H. pylori* entre pares de poblaciones geográficas (F_{st}) versus la distancia geográfica entre las dos poblaciones. **(c)** Diversidad genética promedio en humanos **(d)** y en *H. pylori* entre poblaciones geográficas (H_s) versus la distancia geográfica desde el este de África. $R^2 = 0.77$ para **a**; 0.47 para **b**; 0.85 para **c**; y 0.59 para **d**. En **d**, las muestras compuestas predominantemente por cepas del grupo *hpEurope* están indicadas por círculos rellenos, mientras que los círculos rojos identifican la muestra de Sudáfrica. Los intervalos de confianza están indicados por líneas punteadas (Linz et al. 2007).

Estudios realizados en poblaciones extintas. Se han realizado únicamente dos estudios de este tipo en todo el continente. El primero fue llevado a cabo por Pelayo Correa y colaboradores en 1998, donde muestras de coprolitos de momias precolombinas del norte de Chile (3000 A.P.) dieron positivo para la presencia de antígenos de *H. pylori*. El segundo estudio fue realizado diez años después por Castillo-Rojas y colaboradores, donde se detectó la presencia de un fragmento de 109 pb del gen 16S rRNA en restos gástricos de una momia proveniente del norte de México (950-550 A.P.). De estos estudios se desprende que la presencia de *H. pylori* en

América es precolombina, entrando la bacteria muy probablemente hace 13.000 años por el estrecho de Bering.

Estudios realizados en poblaciones vivas. A pesar de las evidencias de recombinación frecuente y panmixis aparente, *H. pylori* presenta una estructura filogeográfica que sugiere que los polimorfismos bacterianos reflejan la filogeografía humana y sus migraciones (Falush *et al.* 2003; Kersulyte *et al.* 2000; Linz *et al.* 2007; Moodley *et al.* 2009; Wirth *et al.* 2004) (Tabla 1). Uno de los primeros trabajos en describir las diferencias genotípicas de la bacteria en distintas poblaciones humanas fue el de Kersulyte *et al.* (2000), que identificó cinco motivos distintos de delección, inserción y sustitución en el extremo derecho del *cagPAI*. De los tres motivos más comunes, el tipo I predominaba en españoles, peruanos nativos, ladinos guatemaltecos, africanos nativos y residentes de los Estados Unidos; el tipo II en chinos y japoneses; y el III en indios de Calcuta. Las secuencias analizadas para el gen *cagA* y el alelo m1 del gen *vacA*, de cepas provenientes de peruanos nativos, resultaron ser más similares a las españolas, a pesar de la relación más cercana entre amerindios y asiáticos. Esto los llevó a sugerir que la bacteria habría sido traída al Nuevo Mundo por los conquistadores europeos hace 500 años.

Robinson *et al.* (2002) confirmó que la prevalencia de la infección con *H. pylori* en tribus indígenas de Sudamérica era alta y por lo tanto precolombina, a pesar de no contar con cepas de *H. pylori* para análisis genéticos. El mismo año se realizaron un par de estudios en base al análisis molecular de genes bacterianos, donde Yamaoka y colaboradores (2002) presentaron un estudio epidemiológico-molecular basado en el análisis de los mismos genes que Kersulyte *et al.* (2000), comprobando que el 71% de las cepas provenientes de población nativa tenían características genotípicas de cepas no-asiáticas, a pesar de haber tenido una influencia occidental mínima. Al mismo tiempo, el equipo de Ghose (2002) analizó tres genes que habían demostrado previa afinidad filogeográfica (*vacA*, *babB* y *HP0638*) en dos poblaciones venezolanas con distinto origen étnico. En el grupo de sujetos amerindios del Amazonas se observaron

genotipos asiáticos para cada *loci* estudiado, mientras que en el grupo mestizo de Caracas éstos no fueron encontrados. Este hallazgo conformó nuevas evidencias de que la bacteria ha estado presente en humanos al menos desde que los ancestros de los amerindios migraron desde Asia hace más de 11.000 años atrás y de que las cepas europeas puedan probablemente desplazar a las nativas.

El grupo de Devi *et al.* (2006) estudió 7 genes *housekeeping*, los genes *babB*, *oipA* y el *cagPAI* de cepas provenientes de peruanos de ancestría amerindia, con el fin de compararlas con cepas de España y Japón. Se observó que mientras el genotipo europeo predominaba en las cepas peruanas, casi el 20% de éstas representaban una subpoblación con ancestría amerindia. Todas las cepas en cambio, sin importar su afiliación por ancestría, contenían un *cagPAI* de tipo occidental, indicando una posible adquisición de éste por parte de las cepas amerindias durante el período de colonización. El trabajo sugiere la presencia de *H. pylori* en peruanos amerindios que habrían logrado sobrevivir y competir con las cepas europeas, lo que habría sucedido posterior a la adquisición del *cagPAI*, ya sea por ganancia en cepas *cag*-negativas o por recombinación con cepas *cag*-positivas.

El trabajo más reciente consistió en el estudio de la prehistoria humana en el Pacífico mediante el análisis de la distribución de *H. pylori* (Moodley *et al.* 2009). En éste, los autores dedujeron que las poblaciones humanas habrían poblado el Pacífico a partir de dos olas de migración, una alcanzando Nueva Guinea y Australia y la otra extendiéndose a través de Melanesia hasta Polinesia. A partir de este trabajo fue identificada una nueva cepa de *H. pylori* llamada *hpSahul*, separada de la población asiática *ca.* 31.000-37.000 años atrás.

Tabla 1.- Resumen de los estudios sobre *H. pylori* realizados en poblaciones vivas.

Muestra	Proveniencia	Método	Referencia
DNA	Cepas del sur de Europa, Latinoamérica, África, Norteamérica, este Asiático, sur de Asia y noreste Europeo	Análisis molecular del extremo derecho del cagPAI, gen cagA y genotipos del gen vacA	Kersulyte <i>et al.</i> 2000
Serológica	Nativos del Amazonas brasileño	Ensayo serológico: ELISA y Western blot	Robinson <i>et al.</i> 2002
DNA	Nativos de Colombia y de Alaska	Análisis molecular de los genotipos de vacA, cagA y extremo derecho del cagPAI	Yamaoka <i>et al.</i> 2002
DNA	Pacientes de Caracas (ancestría europea y mixta) y Puerto Ayacucho (ancestría amerindia)	Análisis molecular de los genes vacA, HP0638, babB y cagA	Ghose <i>et al.</i> 2002
DNA	Cepas de España, Japón y Perú (ancestría amerindia)	FAFLP, MLST, análisis de secuencia de los genes babB y oipA, secuenciación del cagPAI	Devi <i>et al.</i> 2006
DNA	Cepas de Taiwan, Australia, Nueva Guinea, Melanesia y Polinesia	Análisis molecular de genes <i>housekeeping</i>	Moodley <i>et al.</i> 2009

Composición genética de la población chilena: fundamentos para el estudio de *H. pylori*.

La salida de África ocurrió *ca.* 50.000-70.000 años atrás (Ingman *et al.* 2001; Liu *et al.* 2006; Mellars 2006; Underhill *et al.* 2000; Zhivotovsky *et al.* 2003) hacia el Medioeste y a partir de entonces de forma independiente hacia Europa y Asia. Alrededor de 15.000 años atrás, América fue poblada por humanos de ancestría asiática que habrían cruzado el Estrecho de Bering. Estos primeros americanos sufrieron un cuello de botella genético (Bartolini *et al.* 2003; Reidla *et al.* 2003; Schurr *et al.* 2004) que

redujo drásticamente la diversidad genética en amerindios, evidenciada hoy en día en la absoluta dominancia del grupo sanguíneo O entre ellos, su baja heterocigocidad y el reducido número de haplogrupos mitocondriales (Liu *et al.* 2006; Wang *et al.* 2007). Sin embargo, la diversidad de los amerindios ha ido rápidamente en aumento a partir del influjo genético de europeos y africanos en los últimos 500 años, llevando a un aumento de la población mixta.

La composición actual de la población de Santiago se encuentra conformada por un 84% de haplogrupos mitocondriales indígenas, resultado significativamente superior al obtenido a partir de marcadores nucleares (40%) y del cromosoma Y (22%) (Moraga *et al.* 2004; Rocco *et al.* 2002). Lo anterior tiene una significancia importante, por cuanto el 84% de las mujeres que dieron origen a la población actual de la capital habrían sido indígenas, mientras que el aporte paterno habría sido predominantemente europeo, sin despreciar el aporte negroide a la constitución genética de la población (*Ibíd.* 2004).

En base a los antecedentes de transmisión predominantemente materna de *H. pylori* y a la intersección de dos procesos, por un lado, el origen predominantemente indígena de la población de Santiago, y por otro, el reciente mestizaje ocurrido en ella, surge la pregunta acerca de la distribución de la bacteria en población chilena, que busca contribuir con nuevas herramientas a la genética y evolución de esta población.

--II--

PROBLEMA DE INVESTIGACIÓN

Considerando la transmisión predominantemente vertical de *H. pylori* y la mayor patogenicidad de la cepa europea, se postula la siguiente pregunta:

*¿Qué procesos microevolutivos han modelado la distribución de *H. pylori*, desde inicios del poblamiento americano, en población chilena actual?*

--III--

HIPÓTESIS, RESULTADOS ESPERADOS Y OBJETIVOS

H1

En base a la evidencia de transmisión vertical de la bacteria y a la constitución predominantemente indígena de la línea materna durante el mestizaje, se postula transmisión principalmente matrilineal, desde inicios del poblamiento americano hasta el día de hoy, en esta población.

Predicción

La cepa asiática de *H. pylori* es más abundante, por lo tanto, si H1, el análisis filogenético concentrará las muestras de población chilena con las cepas asiáticas.

H2

En base a la evidencia de co-infección de cepas de *H. pylori* en un mismo hospedero, mayor patogenicidad de la cepa europea y de mestizaje reciente de la población chilena actual, se postula la predominancia de la cepa europea en esta población.

Predicción

La cepa europea es más abundante, por lo tanto, si H2, el análisis filogenético concentrará las muestras de población chilena con las cepas europeas.

H3

En base a la evidencia de transmisión horizontal de distintas cepas de *H. pylori* en un mismo individuo, se postula la predominancia de cepas recombinantes hpEuropa-hpEAsia en población chilena actual.

Predicción

Las cepas de población chilena actual son recombinantes, por lo tanto, si H3, éstas estarán constituidas por fragmentos de diversas estructuras geográficas.

Objetivo general

Discutir la evidencia en torno al origen de la infección con *H. pylori* en población chilena actual a través del análisis filogenético y de diversidad genética.

Objetivos específicos

- Desarrollar protocolos y métodos para la obtención y uso de marcadores moleculares específicos para *H. pylori*.
- Tipificar las variantes de los marcadores moleculares en cepas de población chilena actual.
- Inferir el origen geográfico de las cepas presentes en muestras de población chilena actual para cada gen.
- Identificar la presencia de recombinación en las cepas.

MATERIALES Y MÉTODOS

Muestras

El DNA genómico de 25 cepas de *H. pylori* fue obtenido de pacientes de Santiago. En el presente estudio, el contexto clínico de las cepas aisladas no fue tomado en consideración, como tampoco lo fue la información relacionada a los pacientes intervenidos, debido a que las muestras fueron extraídas en el marco de otro proyecto.

Análisis de PCR

Se seleccionaron dos marcadores moleculares para *H. pylori*, considerando uno que presentara variabilidad (*vacA*) y otro que se caracterizara por su conservación (*trpC*). El gen *vacA* fue seleccionado por codificar la única toxina conocida producida por *H. pylori* presente en todas las cepas. La elección del gen *trpC* (codificador de la proteína de la biosíntesis del triptofano), como representante de los demás genes *housekeeping*, fue realizada en base a la construcción de árboles filogenéticos sencillos para cada uno de estos genes con el fin de detectar el que exhibiera mayor claridad.

Partidores específicos para *H. pylori* fueron diseñados mediante el uso de Primer3 (Rozen y Skaletsky 1998) y Primer Blast (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) en base a los genomas completos de la bacteria: HP26695 y HPJ99 (acceso Genbank: NC_000915.1 y NC_000921.1, respectivamente), para identificar fragmentos de ~500 pares de bases (pb) en ambos genes (*tabla 2*).

La amplificación de los fragmentos de DNA bacteriano se realizó en un volumen final de 25 µl. El conjunto de reactivos de la mezcla se conformó con deoxinucleósidos trifosfatos en una concentración de 200 mM cada uno, solución de buffer (KCl 50 mM,

Tris-HCl 10 mM., Triton X-100 0,1%), MgCl₂ a 2,5 mM., 12,5 µmoles de cada uno de los partidores, una unidad de Taq DNA polimerasa y 1 µl de DNA. El proceso fue llevado a cabo en un termociclador con una desnaturalización inicial de 5 minutos a 95°C, 34 ciclos consistentes en 95°C por 45 segundos, 55°C por 45 segundos para el gen *vacA* (58°C para el gen *trpC*), 72°C por 45 segundos, y una elongación final a 72°C por 5 minutos. Los productos de PCR se resolvieron por electroforesis en gel de agarosa a una concentración del 2%, usando un estándar de peso molecular de 100 pb. La electroforesis se realizó en un buffer TAE 1x (40 mM Trisacetato y 1 mM EDTA) a un voltaje constante de 100 volts. Los productos de la amplificación para ambos genes fueron secuenciados haciendo uso de los partidores descritos en la *tabla 2*.

Tabla 2.- Características de los partidores diseñados para este estudio.

Gen	Primer	Secuencia 5'-->3'	T (°C)	Sitio (inicio-término)	%GC	Tamaño del producto (pb)
<i>vacA</i>	vacAL	ATGGAAATACAACAAACACAC	47.18	900011-900031	33.33	522*
	vacAR	AGAATTGCCATTGAATTGGC	49.39	900532-900513	40.00	
<i>trpC</i>	trpCL	AGGCTTGATGTATCGGGGCGA	57.75	490-510	57.14	494
	trpCR	TGGGGGCTAATGCGGTGCTT	58.68	983-964	60.00	

* El tamaño del producto incluye a los partidores.

Análisis de secuencias

Las secuencias obtenidas de las muestras fueron editadas y revisadas por comparación con sus respectivos cromatogramas. Posteriormente se buscó y descargó un set de 380 secuencias (Falush *et al.* 2003) para ambos genes depositadas en el sitio PubMLST (<http://pubmlst.org/helicobacter>), con un tamaño de 456 pb cada una. La

información acerca de su país de procedencia y estructura poblacional fue incluida en el registro (*tablas 3 y 4*).

Tabla 3.- Cepas por país y su asociación a estructuras poblacionales (Falush *et al.* 2003).

País	Nº	Estructura poblacional	Nº	País	Nº	Estructura poblacional	Nº	
Nueva Zelanda	28	hspMaori	23	Italia	6	hpEurope	10	
		hpEurope	3	Estados Unidos	31	hspWAfrica1	13	
		hspWAfrica1	2			hspAmerind	9	
Australia	3	hpEurope	3			hpEurope	5	
China	1	hspEAsia	1	Colombia	28	hpEurope	23	
Singapore	12	hspEAsia	10			hspAmerind	4	
		hpEurope	2			hspWAfrica1	1	
Japón	1	hspEAsia	1	Guatemala	1	hspWAfrica1	1	
Korea	11	hspEAsia	11	Perú	1	hspEAsia	1	
Hong Kong	1	hspEAsia	1	Canadá	6	hspAmerind	6	
India	17	hpEurope	17	Venezuela	4	hpEurope	2	
Tailandia	2	hspEAsia	1			hspAmerind	1	
		hpEurope	1			hspWAfrica1	1	
Israel	5	hpEurope	5	Burkina Faso	14	hspWAfrica1	14	
Estonia	11	hpEurope	11	Gambia	1	hspWAfrica1	1	
Finlandia	9	hpEurope	9	Senegal	5	hspWAfrica1	5	
Francia	1	hpEurope	1	Sudán	3	hpEurope	3	
Alemania	22	hpEurope	22	Sudáfrica	86	hspSAfrica1	52	
Holanda	1	hpEurope	1			hpEurope	17	
Lituania	1	hpEurope	1			hpAfrica2	10	
España	38	hpEurope	38			hspWAfrica1	7	
Reino Unido	30	hpEurope	28					
		hspEAsia	1					
		hspWAfrica1	1					

Tabla 4.- Cepas utilizadas para cada estructura poblacional (Falush *et al.* 2003)

Grupos poblacionales	Nº de cepas
hpEurope	202
hspEAsia	27
hspMaori	23
hspAmerind	20
hpAfrica2	10
hspWAfrica	46
hspSAfrica	52
TOTAL	380

El set total de 401 secuencias fue alineado con CLUSTAL W (Thompson *et al.* 1994), programa para alineamiento múltiple de secuencias de DNA o proteínas, y luego con DAMBE (Xia y Xie 2001), alineamiento que considera la distribución de codones mediante el alineamiento de secuencias nucleotídicas contra secuencias aminoacídicas.

La construcción de árboles filogenéticos para ambos genes fue realizada en MEGA 4 (Kumar *et al.* 2008; Tamura *et al.* 2007), utilizando el método de *Neighbor-Joining* con un *bootstrap* de 10.000, delección por pares y el modelo de parámetros de *Kimura-2*. Sobre una muestra de 91 cepas para cada gen, se seleccionó con MODELTEST (Posada *et al.* 1998) el modelo de sustitución nucleotídica que mejor encajara con los datos. A partir del modelo seleccionado, se utilizó MrBayes para crear árboles filogenéticos independientes a los obtenidos con MEGA 4.

Se estimó la verosimilitud (Foster 2001) para el set de 91 cepas de ambos genes, elaborando tres posibilidades de árboles (además del árbol inicial) de acuerdo a las hipótesis planteadas. Sobre el logaritmo de verosimilitud, entre sitios de los árboles candidatos para ambos genes, se utilizó CONSEL (Shimodaira y Hasegawa 2001), programa que evalúa la confianza de la selección de un árbol filogenético.

Se utilizaron dos enfoques para detectar, de forma preliminar, eventos de recombinación en el gen *vacA*. Para esto se implementó un análisis de *sliding windows* y reconstrucción filogenética para cada ventana, con una ventana de 20 nucleótidos, obteniendo 22 árboles. Se compararon las filogenias de cada ventana para detectar fragmentos con señal filogenética de estructura este-asiática para evaluar posibles eventos de recombinación con estas cepas. Además se hizo uso del software RAT para investigar eventos de recombinación en un set de secuencias nucleotídicas alineadas, mediante el método en base a distancias (Etherington *et al.* 2004).

Para la inferencia filogenética se utilizó el enfoque de supermatriz, haciendo uso directo y simultáneo de todos los *OTU's*, combinando en una matriz filogenética única a los genes *vacA* y *trpC* y siendo su principal ventaja la observación de una señal filogenética distinta de la observada en árboles por separado, presumiblemente debido a la superposición de la señal al ruido filogenético. Este enfoque ha demostrado además una mayor confianza (de Queiroz y Gatesy 2007; Gatesy *et al.* 2002). Previo a la concatenación de datos se realizó un test de homogeneidad (Farris *et al.* 1995) para evaluar, bajo un enfoque de parsimonia, si el nivel de incongruencia de la señal filogenética entre los dos genes era mayor que el esperado por azar.

Además, se implementó el uso de Network 4.5.1.6. (© 2009 Fluxus Technology Ltd. <http://www.fluxus-engineering.com/>) para la reconstrucción de las redes filogenéticas entre las distintas cepas en base al método de parsimonia. Se ocupó el algoritmo de *Median-Joining*, que permite el procesamiento de datos múltiples no binarios, sobre el set de secuencias de genes concatenados y secuencias completas de DNA mitocondrial.

Finalmente, se llevó a cabo un test de selección (Test Z), sobre los genes *vacA* y *trpC*, para evaluar la probabilidad de rechazar la hipótesis nula de estricta neutralidad ($dN=dS$, número de sustituciones no-sinónimas y sinónimas por sitio, respectivamente) y se estimó la diversidad nucleotídica (π) de *H. pylori*, calculando el

número promedio de sustituciones por sitio entre todos los pares de secuencias dentro de cada grupo.

--V--
RESULTADOS

DNA bacteriano

De un total de 25 muestras estudiadas para población chilena, sólo fue posible amplificar el DNA de 21 de ellas (84%). La amplificación por PCR reveló la presencia de los genes *trpC* y *vacA* en todas las muestras, siendo confirmada la calidad del DNA por medio de electroforesis en gel de agarosa.

Análisis filogenético de los genes estudiados

Gen *trpC*. Las relaciones filogenéticas observadas en la *fig. 6* revelan la formación de grupos para hpEAsia, hpAfrica1 y hpAfrica2, mientras que hpEurope se distribuye a lo largo de todo el árbol formando tres subdivisiones principales, una africana, otra asiática y una última europea, a la que pertenece el mayor número de muestras chilenas (n=16). El resto de ellas se encuentra distribuido en hpAfrica1 (n=5). El árbol filogenético de la *fig. 7*, construido en base a una muestra de 91 cepas, exhibe agrupaciones claras para hpEAsia, hpAfrica2, hpAfrica1 y hpEurope, no mostrando esta última las subdivisiones antes observadas. Las cepas de población chilena se distribuyen a lo largo de todo el árbol, mostrando una tendencia a agruparse con hpAfrica1 (n=7), hpEAsia (n=5) y hpEurope (n=9).

Gen *vacA*. Las cepas de la *fig. 8* muestran una distribución menos clara que para los casos anteriores. hspEAsia y hspMaori se observan en clados definidos, mientras que hpAfrica1, hpAfrica2 y hspAmerind se encuentran dispersos en más de un grupo. hpEurope se encuentra esparcida por todo el árbol y, como para la *fig. 6*, exhibe una inclinación a agruparse con hpEAsia, hpAfrica1 y a conformar un grupo netamente europeo. Las muestras chilenas se agrupan casi en igual número con hpAfrica1 (11) y hpEurope (10).

Las relaciones filogenéticas exhibidas en la *fig. 9*, muestran una mejor definición de los clados que para el caso anterior. hspEAsia, hspMaori y hpAfrica1 se ubican en grupos definidos, mientras que las cepas europeas se distribuyen en hpEAsia y en un grupo conformado por cepas de hpAfrica2 y de población chilena (n=8), estas últimas ubicadas además entre hpAfrica1 (n=8) y hpEAsia (n=4). Únicamente una cepa de población chilena agrupa con Europa.

El test AU ($p=1.0$), realizado en CONSEL, confirma los árboles filogenéticos presentados para ambos genes, rechazando los árboles sometidos a diferentes *constraints*. Los *constraints* equivalen a “forzar” la búsqueda de árboles óptimos a aquel subconjunto del espacio de árboles que es compatible con hipótesis definidas *a priori* acerca de las relaciones de parentesco. De esta manera, el test AU compara la verosimilitud del árbol observado respecto de diferentes hipótesis filogenéticas representadas por árboles “esperados” bajo las diferentes hipótesis, *i.e.* ancestría asiática, amerindia y europea de las cepas de población chilena actual. Estas tres hipótesis sobre el origen de las cepas chilenas definieron tres *constraints* diferentes que fueron usados en la búsqueda, por criterio de parsimonia, del mejor árbol representante de cada una de estas hipótesis. De esta forma, se obtuvieron tres topologías parsimoniosas “esperadas”, derivadas de cada una de las tres hipótesis planteadas.

Supermatriz filogenética. En ésta se observa que los clados hpEAsia y hpAfrica2 se encuentran bien definidos, mientras que Europa se distribuye a lo largo de todo el árbol, subdividiéndose nuevamente en tres grandes grupos, uno más cercano a hpEAsia, otro a los grupos hpAfrica1 y hpAfrica2 y el último formado por cepas europeas, donde se encuentran dispersas todas las de población chilena (=21) (*Fig. 10*).

El test de homogeneidad rechazó la hipótesis de nulidad, la que indica que las diferencias en la señal filogenética entre ambos genes se debe al azar. Sin embargo,

este test debe tomarse con precaución, considerando que la mayor proporción de diferencias nucleotídicas se distribuye dentro de los clados y no entre ellos. Por lo tanto, el análisis de homogeneidad debería considerar, en este caso, un diseño en donde se minimice el efecto de las relaciones intra-clado.

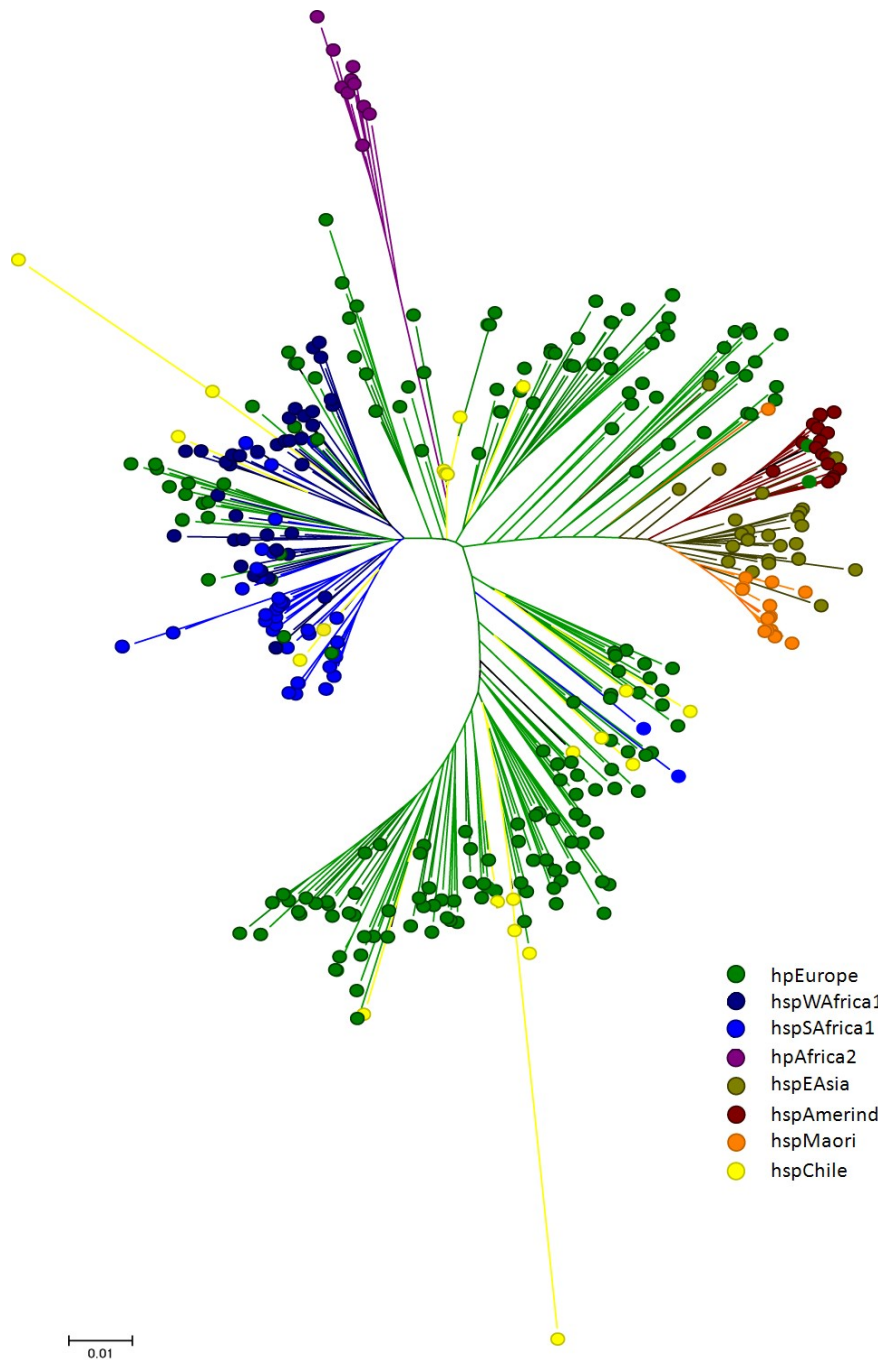


Fig. 6.- Relaciones genéticas deducidas mediante el software MEGA 4, para el gen housekeeping *trpC* ($n=401$), usando el método de Neighbor-Joining (Saitou y Nei 1987), el modelo de parámetros de Kimura 2 (Kimura 1980), un bootstrap de 10.000 y la opción de delección por pares.

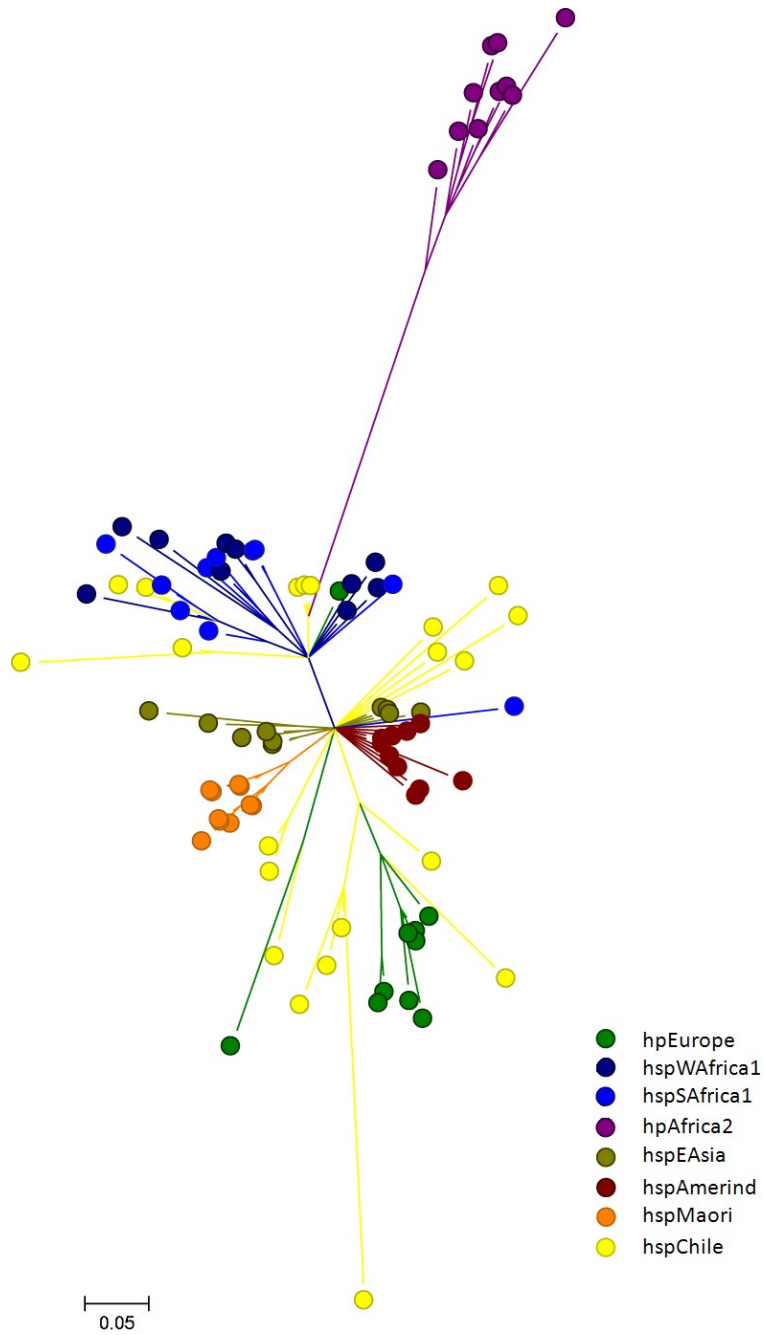


Fig. 7.- Árbol de consenso realizado en Bayes para el gen housekeeping *trpC* (n=91), usando un modelo *K81uf+I+G* de sustitución.

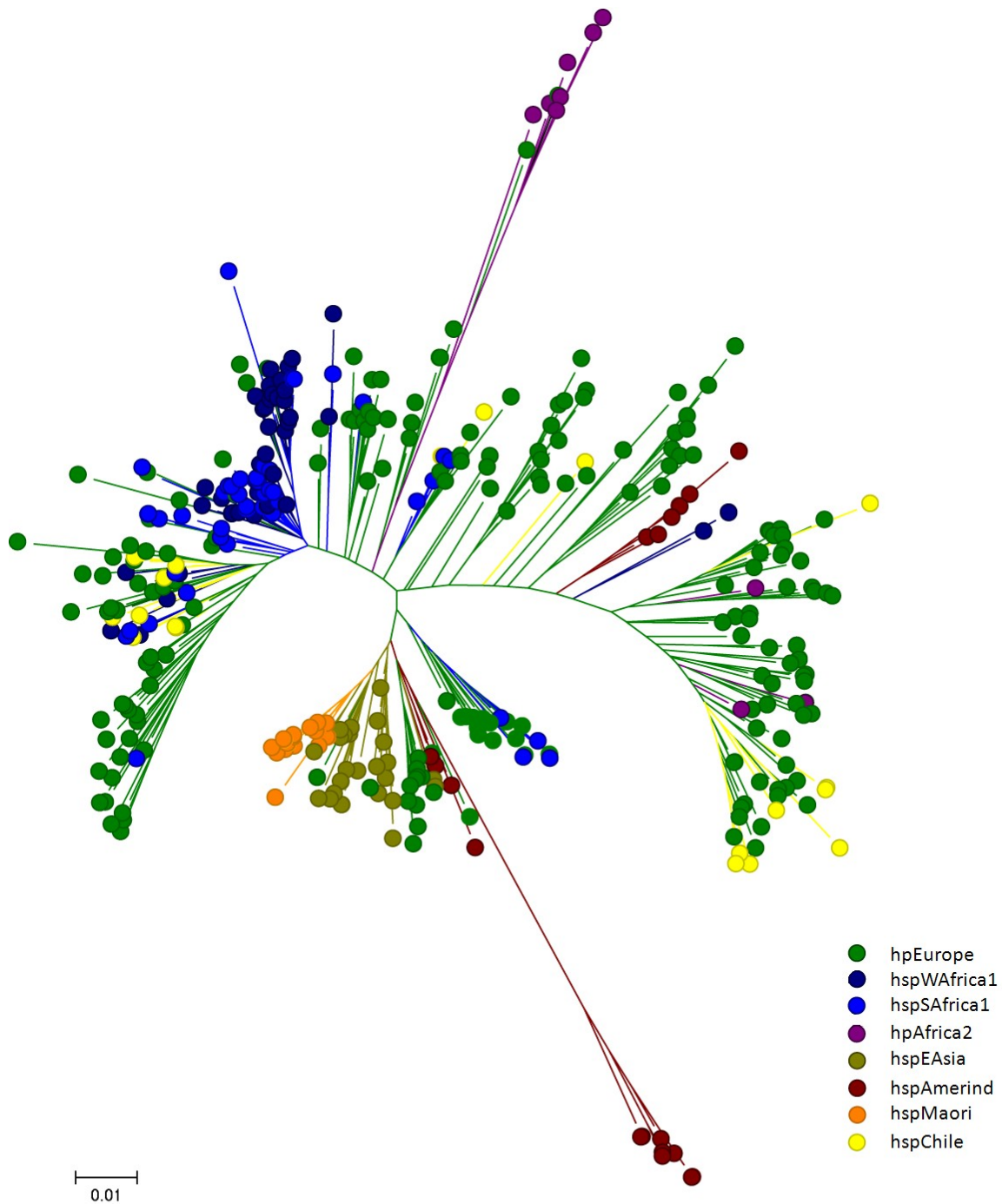


Fig. 8.- Relaciones genéticas deducidas mediante el software MEGA 4, para el gen de patogenicidad *vacA* ($n=401$), usando el método de Neighbor-Joining, el modelo de parámetros de Kimura 2, un bootstrap de 10.000 y la opción de delección por pares.

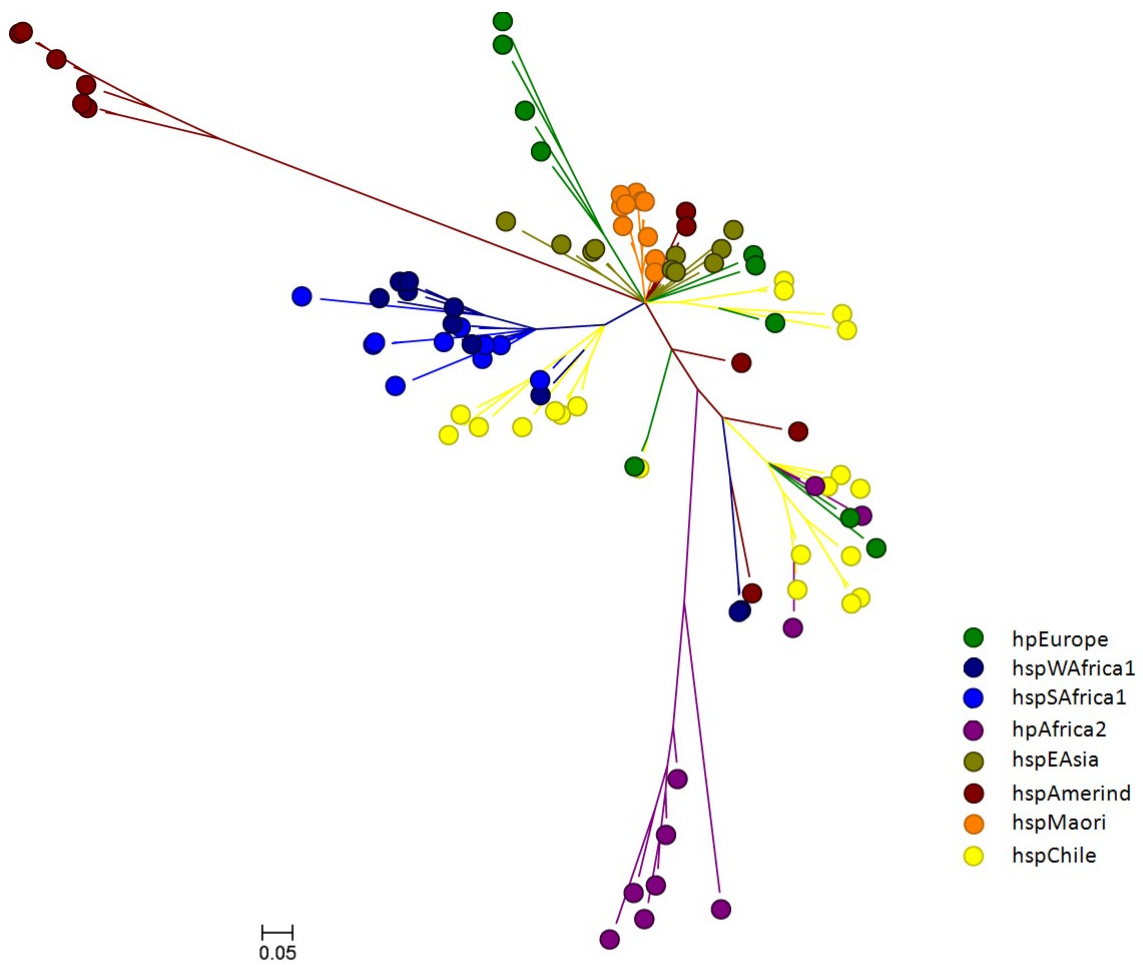


Fig. 9.- *Árbol de consenso realizado en Bayes para el gen de patogenicidad vacA (n=91), usando un modelo GTR+I+G de sustitución.*

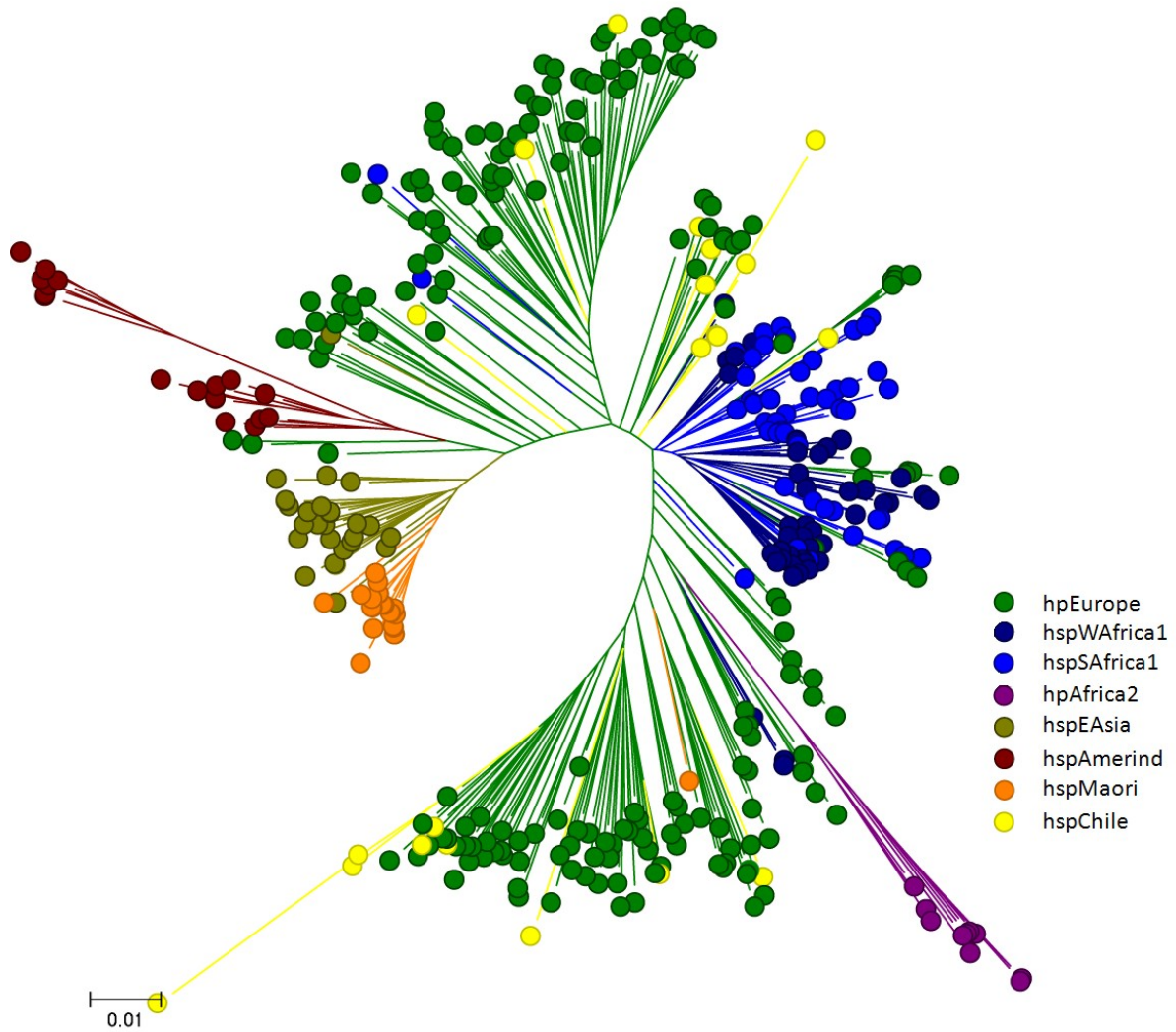


Fig. 10.- Supermatriz de las secuencias de DNA para los genes *vacA* y *trpC* (n=401).

**Análisis comparativo de los árboles filogenéticos en cepas de población chilena
y en cuatro estructuras poblacionales
(Falush *et al.* 2003)**

Gen *trpC*. Los árboles filogenéticos (*fig. 11*), tanto para cepas asiáticas como para amerindias con población chilena, presentan una morfología similar, en donde éstos se bifurcan en dos grupos principales y las muestras chilenas no presentan una agrupación concentrada. El árbol que exhibe la situación del gen *trpC* para cepas africanas y de población chilena, presenta una clara formación de grupos para *hpAfrica1* y *hpAfrica2*, no así para las muestras chilenas que se encuentran incluidas en *hpAfrica1*. La morfología del árbol filogenético que refleja la distribución de cepas europeas y de población chilena difiere considerablemente de los demás, no mostrando agrupaciones de cepas, sino distribuyéndose de forma continua a lo largo de todo el árbol.

Gen *vacA*. La morfología de los árboles (*fig. 12*) para el gen *vacA* no es muy distinta de la de los árboles para el gen *trpC*, sin embargo el modo de agrupación y distancia entre las cepas difiere en cada caso. Los árboles filogenéticos, para cepas asiáticas y amerindias con población chilena, presentan una división principal en dos grandes clados, y luego una secundaria dentro de cada clado. Para el caso de cepas africanas con población chilena, se observa la formación de tres grupos conformados por *hpAfrica1* en la parte superior del árbol, *hpAfrica2* y muestras chilenas en la parte inferior (n=10). Estas últimas se encuentran también distribuidas en *hpAfrica1* (n=11). La morfología del árbol de cepas europeas y muestras chilenas difiere considerablemente de la de los demás, no mostrando agrupaciones de cepas, sino una distribución continua a lo largo de todo el árbol.

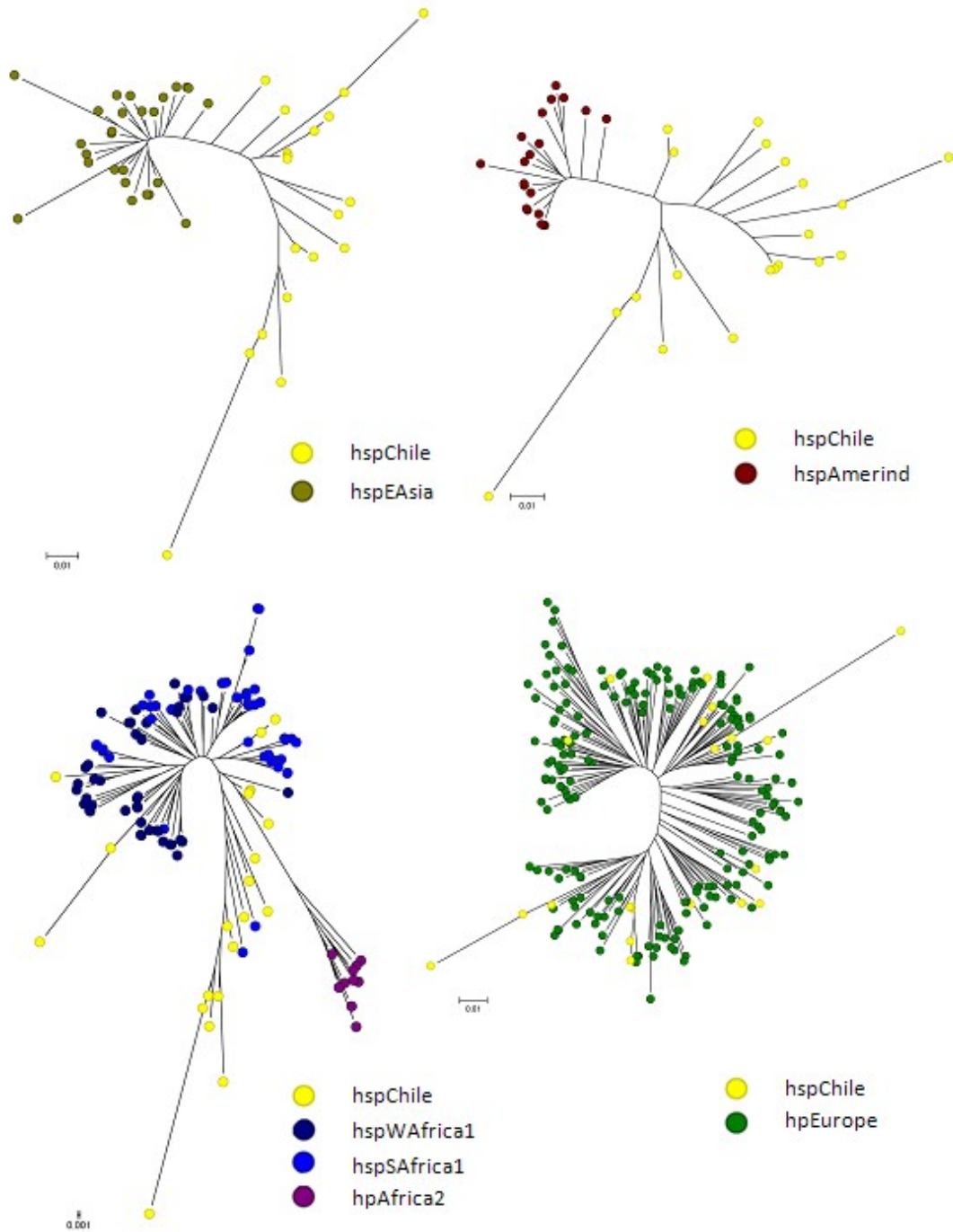


Fig. 11.- Comparación de las relaciones genéticas del gen housekeeping *trpC* para cepas de distintos orígenes y de población chilena (n=21). Árboles construidos mediante el software MEGA 4, usando el método de Neighbor-Joining, el modelo de parámetros de Kimura 2, un bootstrap de 10.000 y la opción de delección por pares.

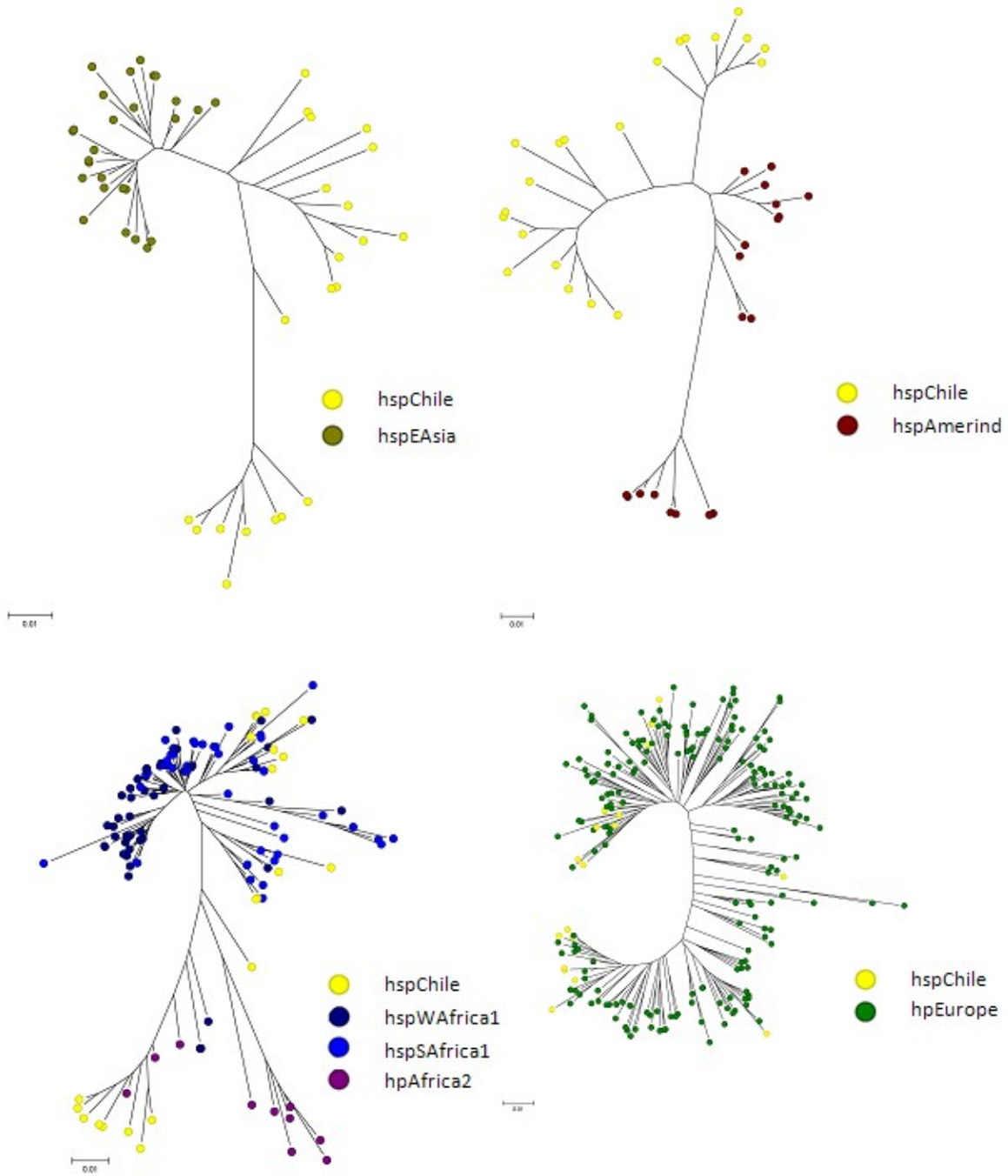


Fig. 12.- Comparación de las relaciones genéticas del gen de patogenidad *vacA* para cepas de distintos orígenes y de población chilena (n=21). Árboles construidos mediante el software MEGA 4, usando el método de Neighbor-Joining, el modelo de parámetros de Kimura 2, un bootstrap de 10.000 y la opción de delección por pares.

Análisis comparativo de árboles filogenéticos en cepas de población chilena y de otros estudios.

En la *fig. 13* es posible comparar la morfología del árbol para los genes *housekeeping* concatenados de 87 cepas de *H. pylori* (Wirth *et al.* 2005) con la del árbol para uno solo de estos genes, el gen *trpC*. Ambos árboles exhiben una distribución continua de las cepas, indicando que el gen *trpC* es representativo de los demás genes *housekeeping*. En la *fig. 14*, en cambio, se compara la morfología del árbol para el gen de patogenicidad *cagA* (*Ibíd.*) con la del árbol para el gen de patogenicidad *vacA*. En el primero se observa una separación entre los grupos africanos e indoeuropeos del asiático. En el segundo, si bien todas las cepas son de población chilena, también puede apreciarse una diferenciación en dos grandes clados. El análisis de *sliding windows* sobre el set de datos reveló que las muestras de población chilena que se agrupan en la parte superior del árbol comparten semejanzas con las cepas de origen amerindio en las primeras 80 pb, y no en las restantes regiones del gen, situación que no se observa en el *cluster* de la parte inferior.

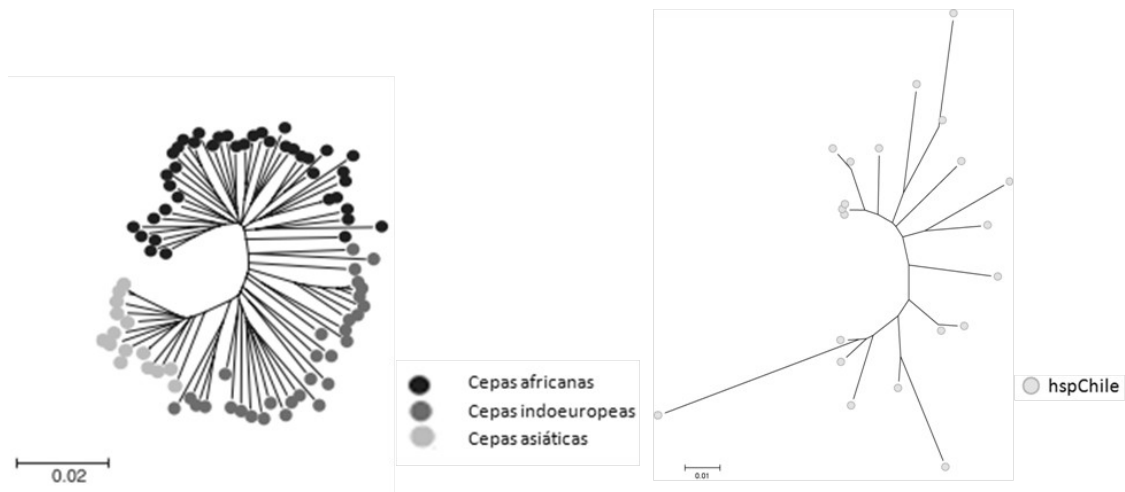


Fig. 13.- Comparación del árbol para los genes *housekeeping* concatenados ($n=87$) con el árbol para el gen *trpC* ($n=21$). El primer árbol fue construido en base al método de Neighbor-Joining con un modelo Tamura-Nei + GTR + I, mientras que el segundo fue construido en base al método de Neighbor-Joining con un modelo de parámetros Kimura 2.

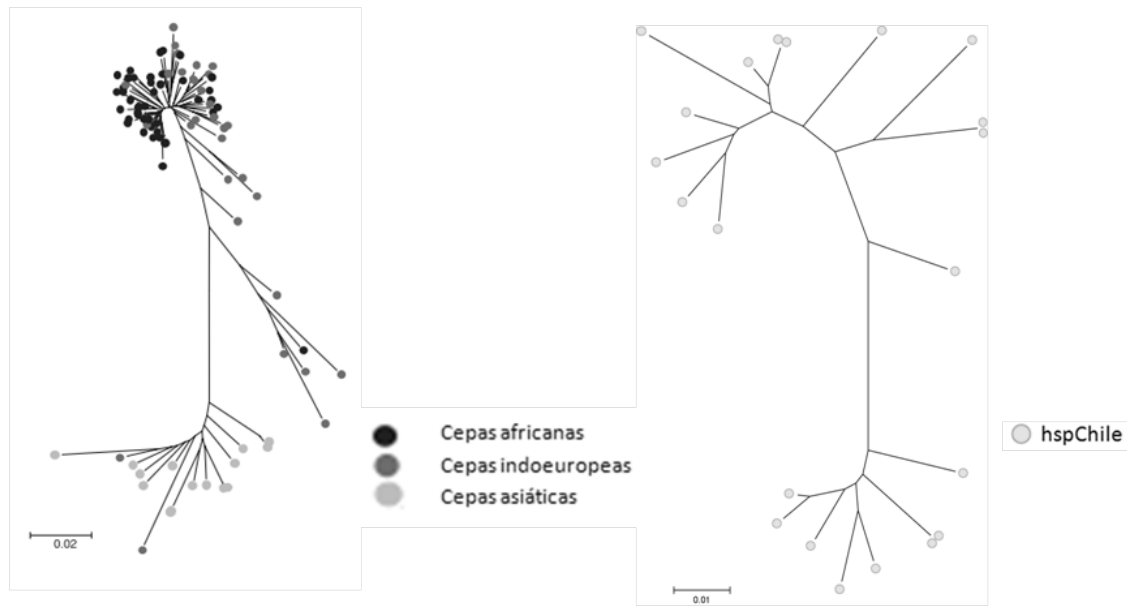


Fig. 14.- Comparación del árbol para el gen de patogenicidad *cagA* (n=87) con el árbol para el gen *vacA* (n=21). El primer árbol fue construido en base al método de Neighbor-Joining con un modelo Tamura-Nei + GTR + I, mientras que el segundo fue construido en base al método de Neighbor-Joining con un modelo de parámetros Kimura 2.

Análisis de Recombinación

Los análisis de recombinación fueron realizados sobre el set de 91 muestras para el gen de patogenicidad *vacA*, debido a la gran variabilidad que éste presenta. El análisis de *sliding windows* resultó en 22 árboles filogenéticos distintos (*fig. 15*), mostrando muy poca señal filogenética el primero, y mucha más el último. En RAT (*fig. 16*), la cepa genéticamente más cercana a la de referencia (Hp100, población chilena) consiste en una de la misma población, compartiendo la misma secuencia durante casi 100 pb. En círculos rojos se destacan posibles eventos de recombinación, observándose la presencia de cuatro en este alineamiento, todos con cepas de hpAfrica1.

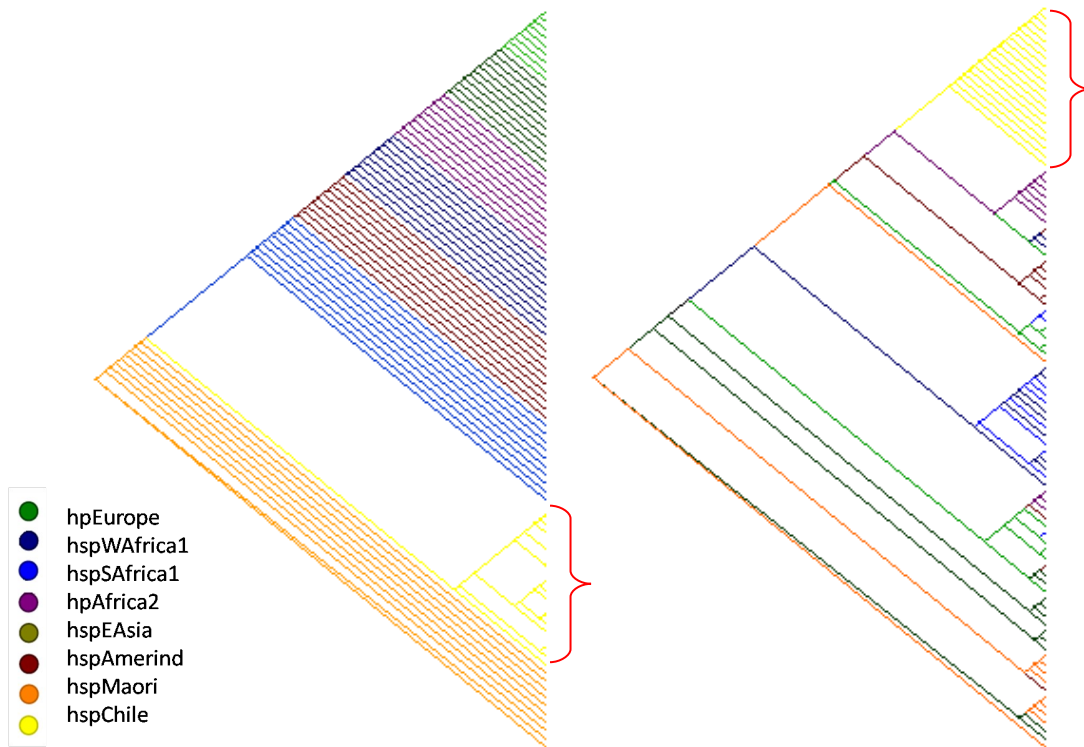


Fig. 15.- Sliding window sobre el set de 91 cepas para el gen vacA.
 (Izq.) Árbol construido para la primera ventana; (der.) árbol construido para la última ventana. Dentro de cada corchete se destacan las muestras de población chilena.

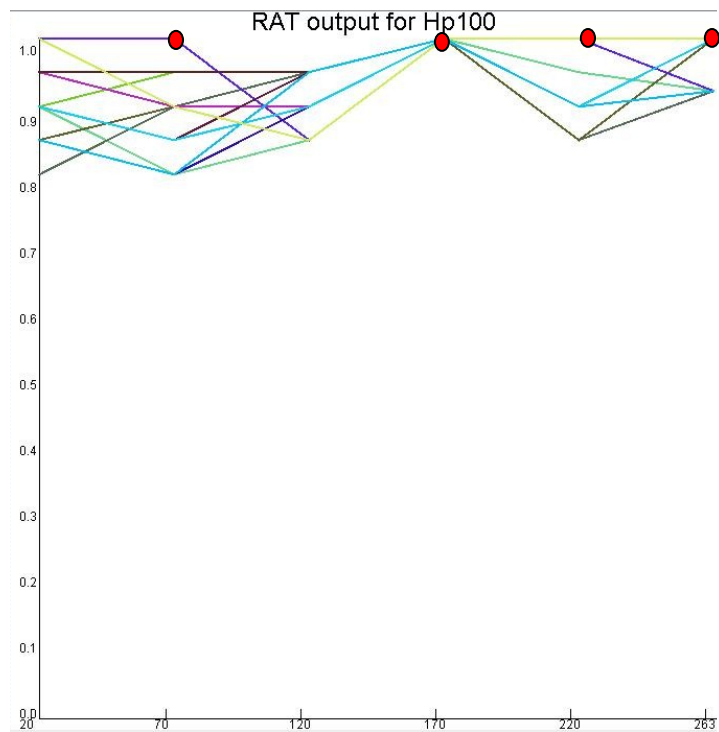


Fig. 16.- RAT sobre un alineamiento representativo de cada estructura poblacional (líneas coloreadas) para el gen *vacA*. El eje X representa la ubicación en la secuencia y el eje Y la distancia genética de cada secuencia del alineamiento en relación a la cepa de referencia de población chilena (Hp100). Los círculos rojos destacan posibles hot spots mutacionales.

Test de selección

Tanto el gen *vacA* como el gen *trpC* rechazan la hipótesis nula de estricta neutralidad, avalando una selección purificadora (negativa) con un número significativamente mayor de sustituciones sinónimas que no-sinónimas (*tabla 5*).

Tabla 5.- Test de selección (Test Z)

	vacA		trpC	
	<i>Prob</i>	<i>Stat</i>	<i>Prob</i>	<i>Stat</i>
hspMaori	0.000	-4.459	0.001	-3.532
hpEurope	0.000	-6.063	0.000	-8.481
hspEAsia	0.000	-5.957	0.000	-6.013
hpAfrica2	0.000	-6.029	0.000	-5.922
hspWAfrica	0.000	-4.834	0.000	-5.732
hspSAfrica	0.000	-5.893	0.000	-5.323
hspAmerind	0.000	-5.935	0.000	-5.146
Chile	0.000	-5.002	0.000	-5.630
Total	0.000	-6.601	0.000	-7.489

El test estadístico (dN-dS) se muestra en la columna Stat. Los valores p menores a 0.05 son considerados significativos. dS y dN son los números de sustituciones sinónimas y no sinónimas por sitio, respectivamente. La varianza de las diferencias fue computada usando el método de bootstrap (100 réplicas). Los análisis fueron realizados usando el método de Nei-Gojobori en MEGA4.

Estimación de diferenciación nucleotídica

El valor L corresponde al coeficiente entre el número de polimorfismos en una secuencia y el total de nucleótidos presentes en ella. Éste alcanza los valores más altos para la cepa europea, tanto en el gen *vacA* como en el gen *trpC*, mientras que los valores más bajos se observan en la cepa maorí. El valor π consiste en la diversidad promedio para cada población y su comparación con θ permite evaluar si la diversidad presente en la población es homogénea a lo largo de la muestra o está concentrada en unos pocos individuos.

En la *tabla 6* se observan los valores para los genes *vacA* y *trpC* en todas las poblaciones, en comparación con un set de secuencias de DNA mitocondrial de un tamaño similar al de los genes ya mencionados. Como únicamente fueron consideradas aquellas secuencias completas de DNA mitocondrial, se excluyó del análisis al subgrupo polinésico. Los subgrupos africanos fueron considerados como uno solo, al igual que el subgrupo amerindio y otros grupos americanos. De la comparación de θ y π , para el gen *vacA*, es posible afirmar que la alta variabilidad observada en *hpEurope* es homogénea a lo largo de la muestra. En cambio, la variabilidad presente en *hspSAfrica* y *hspWAfrica*, que exhiben los niveles más altos después de *hpEurope*, sería producto de unos cuantos individuos. *hpEurope* también muestra los valores más altos de diversidad homogéneamente distribuida para el gen *trpC*, siguiéndola *hspWAfrica* y Chile, donde la variabilidad se distribuye de forma homogénea sólo en la muestra chilena.

Tabla 6.- Comparación entre la diversidad nucleotídica para los genes *vacA*, *trpC* y fragmento de ~500 pb de DNA mitocondrial.

	<i>vacA</i>		<i>trpC</i>		<i>mtDNA</i>	
	θ^1	π^2	θ	π	θ	π
<i>hspMaorí</i>	0.20	0.017	0.16	0.015	---	---
<i>hpEuropa</i>	0.72	0.065	0.93	0.073	0.010	0.001
<i>hspEAsia</i>	0.24	0.026	0.39	0.031	0.024	0.002
<i>hpAfrica2</i>	0.37	0.063	0.26	0.040	0.016	0.001
<i>hspWAfrica</i>	0.40	0.034	0.41	0.042		
<i>hspSAfrica</i>	0.42	0.035	0.39	0.041		
<i>hspAmerind</i>	0.39	0.069	0.25	0.028	0.022	0.001
<i>Chile</i>	0.34	0.067	0.47	0.062		

¹ Proporción de sitios polimórficos (estandarizada por el logaritmo natural)

² Diversidad promedio para la población completa.

Reconstrucción de relaciones entre cepas

La distribución de las cepas de *H. pylori* es muy similar a la observada en la supermatriz, donde *hpEurope* se distribuye a lo largo de todo el eje, conformando grupos consigo misma, con *hpEAsia* y *hpAfrica1*, y las cepas de población chilena siguen una distribución muy similar a *hpEurope* (*fig. 17*).

Bajo un enfoque de parsimonia, la relación entre individuos se observa mucho más clara para aquellos que presentan baja diversidad nucleotídica y escasa recombinación (*fig. 17*). Sin embargo, individuos con altas tasas de recombinación pueden presentar una mayor capacidad de proveer información acerca de la historia de las poblaciones humanas, al exhibir secuencias altamente polimórficas.

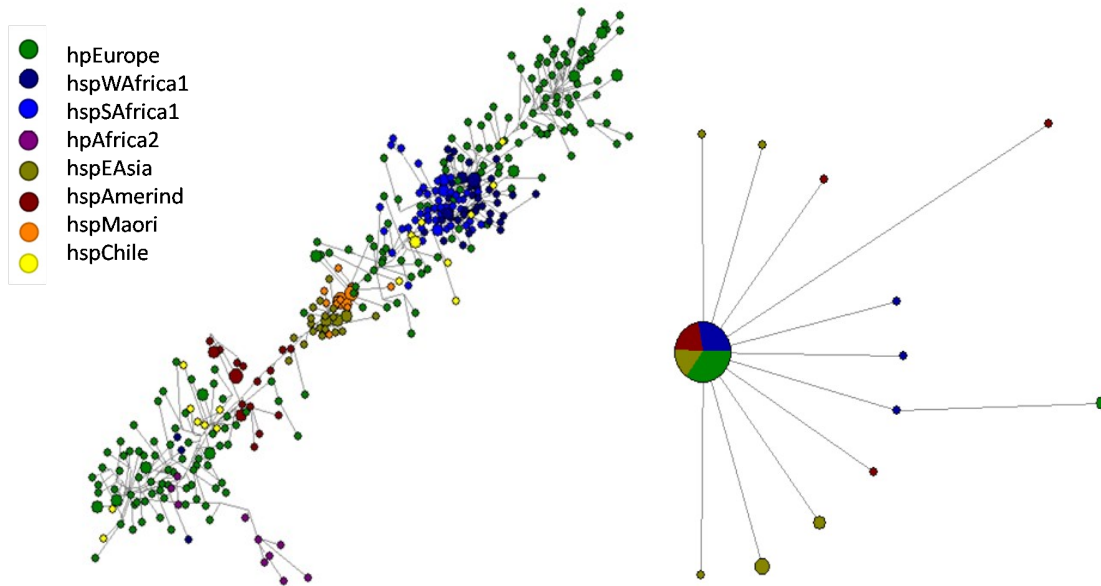


Fig. 17.- (Izq.) Red de relaciones entre las cepas a partir de las secuencias de genes concatenados. (Der.) Red de relaciones entre las secuencias de DNA mitocondrial (n=59).

DISCUSIÓN

Los humanos somos responsables de haber modelado parte de la evolución de los microbios, reflejando en la distribución de sus genomas algunas de las diferencias genéticas, lingüísticas y culturales propias de la población humana. Es por esto que el estudio de microbios que han evolucionado en estrecho contacto con los humanos, permite el estudio indirecto de sus hospederos, otorgando información desde una perspectiva única acerca de procesos evolutivos de las poblaciones.

El aislamiento genético, la deriva y la selección podrían explicar los procesos de radiación de humanos en diferentes grupos y la divergencia de *H. pylori* en tipos geográficos, donde tanto hospedero como bacteria muestran patrones de aislamiento por distancia (Linz *et al.* 2007). Sin embargo, las interacciones entre grupos humanos están continuamente ofreciendo nuevos influjos genéticos, tanto a nivel humano como microbiano, en los que parásitos que alguna vez divergieron en la historia humana se encuentran en proceso de convergencia, volviendo difícil la diferenciación de las señales filogeográficas.

En las relaciones hospedero-parásito, un flujo génico restringido en el hospedero lleva a adaptaciones locales en el microbio (Gandon 1998; Kirkpatrick *et al.* 1997). Un ejemplo de ello es la afinidad de las cepas de *H. pylori* a unirse a antígenos de grupos sanguíneos expresados en la mucosa gástrica humana, donde las cepas europeas se unen a los tres antígenos de grupos sanguíneos (A, B y O), mientras que las cepas amerindias únicamente tienen afinidad por los antígenos del grupo sanguíneo O (Boren *et al.* 1993; Domínguez-Bello *et al.* 2008). La mezcla en humanos modernos resulta, por lo tanto, no sólo en un aumento en la diversidad humana, sino también en un aumento de la diversidad microbiana del microbioma humano. A largo plazo, es posible predecir una superación de las distancias genéticas máximas por la homogenización de la estructura genómica en las distintas cepas (recombinantes) y

humanos (mestizos), resultando en una disminución de las distancias genéticas poblacionales. Observar cómo ocurren estos procesos se vuelve una tarea necesaria, dada la importancia que algunas bacterias humanas tienen para la salud de las poblaciones, ya que alteraciones en las dinámicas de co-evolución podrían traer como consecuencia cambios en el patrón de enfermedades.

A continuación se discuten las hipótesis, planteadas al inicio del presente trabajo, a la luz de los datos obtenidos:

1. *H. pylori* presenta transmisión principalmente matrilineal, desde inicios del poblamiento americano hasta el día de hoy, en población chilena actual. En el presente estudio no se observa un componente asiático en la estructura poblacional de *H. pylori*. Es probable que, debido a la naturaleza de la muestra, no pueda observarse una asociación entre cepas de población chilena con las de origen asiático. No se descarta, sin embargo, la posibilidad de que estudios enfocados en otros genes o en una muestra de individuos sanos permitan aceptar esta hipótesis.

2. Existe una predominancia de la cepa europea de *H. pylori* en población chilena actual. Las cepas de población chilena se agrupan, a lo largo de todos los análisis, con hpAfrica1 y hpEurope en su mayoría, confirmando el fuerte componente europeo de *H. pylori* en población chilena actual. Para explicar este componente tan definido se postulan las siguientes posibilidades:

2.1. Sesgo muestral. Un factor a considerar en este estudio es el sesgo presente en la muestra, el cual incluye únicamente a aquellos individuos que, debido a una enfermedad gástrica, requirieron de procedimientos médicos para ser debidamente diagnosticados. El consumo de antibióticos y otros medicamentos, como los inhibidores de la bomba de protones que reducen el ácido gástrico, podrían estar afectando a las poblaciones bacterianas al

aumentar la presión selectiva a favor de variantes particulares, modelando el genoma de microbios en poblaciones humanas. Esto podría estar explicando la prevalencia de las cepas europeas, ya que presentan una mayor patogenicidad que las demás.

2.2. Competencia entre cepas y dominancia de cepas europeas. Es posible que tanto la diversidad intra-genómica (mosaicismo) como la inter-genómica (diversidad de cepas) hayan jugado un rol importante en la expansión de hpEurope en individuos mestizos. Pareciera que las cepas de origen amerindio carecieran de la diversidad necesaria para sobrevivir en hospederos con una alta diversidad genética, como la que presentarían los mestizos, ocurriendo un desplazamiento de las cepas amerindias por las europeas. Este desplazamiento podría haber ocurrido por competencia durante la colonización, donde cepas generalistas tendrían una mayor oportunidad de colonizar diversos ambientes que las especialistas (como en el caso de cepas que se unen al grupo sanguíneo O amerindio que tendrían menos éxito colonizando individuos carentes de este grupo) o por transformación, donde el DNA de las cepas de origen europeo sería seleccionado y por lo tanto las cepas recombinantes se habrían vuelto “más europeas”. Asumiendo que las cepas de población chilena presentan un grado de recombinación, el gran componente europeo en éstas y el escaso componente amerindio, avalaría esta hipótesis.

3. Las cepas de población chilena recombinan. En el presente estudio la detección de recombinación fue posible, no obstante los eventos de recombinación no fueron caracterizados (identificando secuencias parentales y recombinantes, o puntos de quiebre) ni fue medida la cantidad de recombinación, enfoques que quedarán pendientes para estudios futuros.

--VII--

CONCLUSIONES

Los análisis filogenéticos realizados sobre el set de cepas de distintos orígenes geográficos, incluidas las muestras chilenas, muestran que:

1. hpEAsia y hpAfrica2 mantienen una agrupación más definida que las demás cepas, lo cual puede estar reflejando una mayor diferenciación nucleotídica del resto de los grupos poblacionales. Esto se vuelve evidente cuando se aprecia que las cepas de hspEAsia sólo provienen del este asiático (a excepción de una) y las de hpAfrica2 sólo lo hacen de Sudáfrica. Una situación más compleja se observa en hpEurope que, al encontrarse distribuida a lo largo del árbol, no representa un clado definido, lo que puede estar relacionado con migraciones y eventos de colonización en los distintos continentes a lo largo de la historia humana. Además, su presencia en todas las estructuras poblacionales puede estar evidenciando una alta competencia y mejor adaptación de este tipo de cepa más generalista, pudiendo conllevar incluso a un eventual reemplazo de otras cepas más específicas.
2. Las cepas de población chilena comparten una homología considerable con aquellas pertenecientes a hpEurope y a hpAfrica1, indicando patrones genotípicos en común entre los grupos mencionados. Esto podría explicarse por un lado, por la conquista y posterior colonización europea en el continente americano hace más de 500 años, y por otro, al tráfico de esclavos africanos entre el siglo XVI y XVII, provenientes en su mayoría del oeste y centro de África.
3. La continuidad que presentan cepas europeas y de población chilena, en los árboles filogenéticos, podría estar reflejando una menor diferenciación nucleotídica entre estas cepas, dada por el intercambio génico continuo entre ambas poblaciones.

4. Las diferencias nucleotídicas entre las cepas de distintos orígenes no son tan grandes para genes conservados, lo que permite observar árboles que no forman grupos muy diferenciados, sino una continuidad en la distribución de las cepas que se agrupan según cercanía genética. Por otro lado, la formación de clados muy diferenciados entre cepas distintas se observa en los árboles para genes de patogenicidad, indicando mayores diferencias nucleotídicas y, por tanto, mayor variabilidad. Esta última refleja una mayor tasa de recombinación para este gen que para el conservado y quizá un *hot spot* genético para la recombinación dentro del genoma bacteriano.
5. Es posible que la formación de subgrupos, observada en la muestra chilena (*fig. 14*), esté relacionada con una variabilidad en los genotipos del gen *vacA*.
6. La presencia de *H. pylori* en población extinta permite afirmar su existencia en América precolombina y, por tanto, una ancestría asiática. Sin embargo, la fuerte asociación del gen de patogenicidad *vacA* con *hpEurope*, permite inferir su adquisición en momentos posteriores al contacto europeo.

Direcciones futuras

1. Estudiar la variabilidad en el extremo derecho de la isla de patogenicidad *cag*, que muestra una distribución geográfica no aleatoria, y realizar análisis de los genotipos del gen *vacA*. Este gen se encuentra en todas las cepas de *H. pylori*, existiendo diferentes genotipos de acuerdo a la variabilidad presente en la región señal y media. La discriminación entre los genotipos de este gen puede ser usada para clasificar las cepas asociadas a patologías determinadas (*van Doorn et al. 2009*), siendo de particular interés epidemiológico y farmacogenético en ausencia de trabajos de este tipo para población chilena.

2. Llevar a cabo un análisis de recombinación más profundo. *H. pylori* recombina en fragmentos de 417 pb en promedio (Suerbaum *et al.* 2007), significando esto que deben ser secuenciados fragmentos más extensos que permitan una interpretación a la luz de eventos de recombinación.
3. Hacer uso de otras fuentes de DNA bacteriano, como deposiciones humanas, con el fin de poder acceder a una muestra control de individuos sanos. De esta forma, podría hacerse un muestreo geográfico de la bacteria en población chilena actual y en grupos étnicos del norte, centro y sur, incluyendo Rapa Nui, tomando en consideración tanto la diversidad genética del parásito como la del hospedero.
4. En este estudio se buscaron estómagos de las momias depositadas en las dependencias de la Facultad de Ciencias Sociales (colecciones Az-71 y Ta-40), no encontrando ninguno para poder realizar la detección bacteriana (*anexo: imágenes de algunas de las momias que fueron encontradas y revisadas en el depósito*). Para futuros proyectos, se sugiere buscar la presencia de *H. pylori* en momias precolombinas del norte chileno, con el fin de comparar el origen de su infección en población precolombina y actual. Posteriormente, sería posible comparar las muestras precolombinas con muestras provenientes de grupos Siberianos para evaluar un posible origen centroasiático de estas cepas.

--VIII--

AGRADECIMIENTOS

A los Drs. Sergio Flores y Mauricio Moraga, por guiar los trabajos de laboratorio, análisis bioinformáticos y participar activamente en el diseño y desarrollo de este proyecto; al Dr. Héctor Toledo por facilitar el acceso a muestras de DNA bacteriano; al Dr. Pelayo Correa por facilitar información, textos e imágenes relacionadas a sus trabajos de investigación en Chile; y finalmente, al comité evaluador por revisar el escrito de este proyecto.

--IX--

GLOSARIO

Aislamiento genético: ausencia de intercambio genético entre poblaciones o especies como producto de la separación genética o mecanismos que previenen la reproducción.

Aislamiento por distancia (IBD): tendencia de los individuos de encontrar parejas provenientes de poblaciones cercanas. Como resultado de esta tendencia, las poblaciones que viven cerca son genéticamente más similares que las poblaciones que viven separadas.

Análisis filogenético: estudio de las relaciones evolutivas entre varios grupos de organismos, descubiertas mediante la secuenciación de datos moleculares y matrices de datos morfológicas.

Antígeno: molécula capaz de producir una respuesta inmune adaptativa.

Bacterias gram-negativas: bacterias que no retienen la tinta en el protocolo de tinción Gram y cuyas capacidades patogénicas están asociadas a ciertos componentes de sus paredes celulares.

Bootstrap: es una técnica para evaluar la confianza de cada clado de un árbol observado, basado en la cantidad de árboles que muestran los mismos clados.

Cepa: variante genética o subtipo de un microorganismo.

Clado: cada una de las ramas del árbol filogenético propuesto para agrupar a los seres vivos. Se interpreta como un conjunto de especies emparentadas.

CLUSTALW: programa de alineamiento de múltiples secuencias de DNA o proteínas. Calcula el mejor ajuste para las secuencias seleccionadas y las coloca en fila de tal manera que las identidades, similitudes y diferencias puedan ser vistas. Las relaciones evolutivas pueden ser observadas mediante cladogramas o filogramas (Thompson *et al.* 1994).

Cluster: grupo o conglomerado.

Coevolución: cambio en la composición genética de una especie (o grupo) en respuesta al cambio genético en otra.

Comensalismo: tipo de relación entre dos organismos donde uno se beneficia y el otro no es afectado.

Conjugación: transferencia de información genética desde una célula bacteriana a otra que requiere contacto físico entre ambas bacterias.

Convergencia: proceso por el cual organismos no relacionados adquieren, de forma independiente, características similares.

Cromosomas homólogos: cromosomas que hacen sinapsis durante la meiosis.

Cromosoma Y: uno de los dos cromosomas determinantes de sexo en la mayoría de los mamíferos, incluyendo humanos. Está constituido por 60 millones de pares de bases y es transmitido de padre a hijo.

Crossing over: intercambio de material genético entre cromosomas homólogos.

Cuello de botella: disminución del tamaño de la población en un período de tiempo determinado que reduce la diversidad genética de los individuos.

DAMBE (Data Analysis in Molecular Biology and Evolution): programa que presenta funciones analíticas para genómica y evolución molecular (Xia y Xie 2001).

Deriva génica: mecanismo evolutivo en donde una población experimenta un cambio en la frecuencia de un alelo en particular provocado por azar.

Divergencia genética: proceso en el cual dos o más poblaciones de una especie ancestral acumulan cambios genéticos independientes a través del tiempo.

D-loop mitocondrial: estructura donde dos hebras de una molécula de DNA, separadas por un tramo, se mantienen separadas por una tercera hebra que es complementaria a una de las principales y se acopla con aquella, desplazando a la otra hebra principal en la región. El D-loop mitocondrial ocurre en la principal área no codificante de la molécula de DNA mitocondrial, en un segmento llamado región de control.

Endosimbiosis: asociación estrecha entre especies, en la que los individuos residen dentro de una misma célula. El término endosomático describe con mayor claridad a cualquier simbiote que residiera en el interior del cuerpo de otro ser vivo, como es el caso de muchas de las bacterias que forman parte de la microbiota intestinal.

Estructura poblacional: forma de subdivisión de las poblaciones.

Filogeografía: estudio de los procesos históricos que pueden ser responsables de las distribuciones geográficas contemporáneas de los individuos. Describe las señales genéticas estructuradas geográficamente dentro de una especie.

Fst: medida de diferenciación de una población basada en datos genéticos polimórficos, tales como SNPs o microsatélites. El *Fst* es la correlación de alelos escogidos aleatoriamente dentro de la misma subpoblación en relación a los encontrados en la población completa.

Genes autosómicos: genes ubicados en cromosomas distintos a los cromosomas sexuales que se heredan y expresan en forma idéntica en machos y hembras.

Haplogrupos de DNA mitocondrial humano: grupos determinados por las variaciones encontradas en el DNA mitocondrial humano. Éstos trazan la ascendencia matrilineal hasta los orígenes de la especie humana en África y desde allí, a su dispersión por todo el planeta.

Haplotipos: combinación de alelos ligados a múltiples *loci* que son transmitidos en conjunto y que determinan un cierto fenotipo.

Hot spot genético: región del DNA que es mucho más susceptible a mutaciones que otras.

Inversiones: ocurren cuando un cromosoma se quiebra en dos lugares y el fragmento de DNA resultante se vuelve a insertar de forma invertida, en el cromosoma.

Kimura-2: modelo que toma en cuenta las tasas de sustitución transicionales y transversionales, asumiendo que las cuatro frecuencias nucleotídicas son las mismas y que las tasas de sustitución no varían entre sitios (Kimura 1980).

Linfoma MALT: forma de linfoma que envuelve al tejido linfoide asociado a la mucosa, frecuentemente en el estómago, pero todas las mucosas pueden verse afectadas (Fischbach *et al.* 2009).

Marcadores genéticos: son genes o secuencias de DNA con una ubicación conocida en un cromosoma que pueden ser usados para identificar células, individuos o especies. Algunos tipos de marcadores genéticos son: RFLP, RAPD, VNTR, SSR, SNP, STR, entre otros.

Máxima parsimonia: método estadístico no paramétrico usado para estimar filogenias. Bajo parsimonia, el árbol filogenético preferido es aquel que requiera el mínimo de cambios evolutivos para explicar los datos observados.

Máxima verosimilitud: método estadístico usado para encajar un modelo estadístico a los datos y para proveer estimaciones a los parámetros del modelo.

MEGA4 (Molecular Evolutionary Genetics Analysis): software diseñado para el análisis comparativo de secuencias génicas homólogas, tanto de familias multigénicas como de especies diferentes, con un énfasis especial en la inferencia de relaciones evolutivas y evolución de modelos de DNA y proteínas (Kumar *et al.* 2008).

Meiosis: proceso de división en el cual el número de cromosomas por célula es reducido a la mitad.

Microbioma: total de microbios, sus genomas y sus interacciones en un ambiente definido.

Microsatélites: secuencias repetidas de 1-6 pb usadas como marcadores moleculares para estudios de parentesco y poblaciones, entre otros.

Método de mínima evolución (ME): método en el cual el árbol es adaptado a los datos y el largo de ramas es determinado usando el método de mínimos cuadrados no ponderados.

MLST (Multilocus Sequence Typing): técnica de clasificación de múltiples *loci*, en la que cepas de especies bacterianas son caracterizadas usando las secuencias de DNA de múltiples genes *housekeeping*.

MODELTEST: programa que utiliza los valores de verosimilitud logarítmica para establecer el modelo de evolución nucleotídica que mejor encaje con los datos (Posada *et al.* 1998).

Mosaicismo: alteración genética en la que coexisten dos o más poblaciones de células con distinto genotipo en un mismo individuo.

MrBayes: es un programa para la inferencia bayesiana de filogenias que emplea una variante de Montecarlo basado en cadenas de Markov (Huelsenbeck 2000).

Neighbor-Joining: método de agrupamiento usado para la construcción de árboles filogenéticos basado en el criterio de mínima evolución (Saitou y Nei 1987).

Organismos micro-aerófilicos: microorganismos de tipo específico que requieren oxígeno para sobrevivir, pero en ambientes con niveles más bajos que los presentes en la atmósfera (concentración ~20%).

Panmixis: población en la que todos los individuos son parejas potenciales. Esto asume que no hay restricciones de emparejamiento, ni genéticas ni conductuales, y por lo tanto, la recombinación es posible.

Parasitismo: tipo de relación simbiótica entre organismos de diferentes especies donde uno, el parásito, se beneficia a expensas del hospedero.

RAT (Recombination Analysis Tool): herramienta para investigar eventos de recombinación en un número de secuencias alineadas (de proteínas o DNA) de cualquier tamaño (desde secuencias virales cortas a genomas completos) (Etherington *et al.* 2004).

Selección natural: proceso por el cual ciertos rasgos hereditarios, que favorecen la sobrevivencia y reproducción exitosa, se vuelven más comunes en una población a lo largo de generaciones sucesivas.

Sliding window: procedimiento para detectar potenciales recombinaciones, basado en la inferencia filogenética bayesiana.

SNPs (polimorfismo de nucleótido simple): variación en la secuencia de DNA, que afecta a un solo nucleótido de una secuencia del genoma, que debe darse al menos en un 1% de la población para ser considerada como tal. Debido a que no cambian mucho de una generación a otra, es posible seguir su evolución en estudios poblacionales.

Test AU (Approximate Unbiased Test): valor p de un test aproximado y no sesgado.

Transducción: proceso por el cual el DNA es transferido de una bacteria a otra mediante un vector viral.

Transformación: proceso por el cual algunas bacterias adquieren nuevos genes tomándolos de moléculas de DNA del ambiente.

Translocación: cambio en la localización, donde parte de un cromosoma se transfiere a otro.

Virus RNA: virus que presenta RNA como su material genético. Algunas enfermedades humanas causadas por virus de RNA incluyen el Síndrome Respiratorio Severo Agudo (SARS), influenza y hepatitis C.

--X--

ANEXO



Fig. 18.- Momias de niños enfardados (sin numeración). Se encuentran envueltos en un textil muy dañado y amarrado con una trenza de cáñamo.



Fig. 19.- Momia adulta (AZ71-T21). Consiste únicamente en un torso, donde extremidades superiores e inferiores y cabeza fueron amputadas postmortem, esta última a nivel de C2-C3. Parte del coxal fue aserrado, así como lo fue la parrilla costal. El contenido torácico y abdominal fue removido.



Fig. 20.- Momia de niño (CT95S). Presenta ambos brazos, uno sobre el pecho y otro doblado hacia atrás, y ambas piernas muy dobladas. Cabeza y torso ausentes.



Fig. 21.- Momia de niño (CT5S1). Se encuentra sentado con la pierna derecha sobre la izquierda. Parte del abdomen está comprimido sobre la pelvis y el brazo izquierdo está doblado hacia atrás. No presenta brazo derecho ni cráneo.

REFERENCIAS

Achtman, M., Azuma, T., Berg, D., Ito, Y., Morelli, G., Pan, Z-J., Suerbaum, S., Thompson, S., van der Ende, A., van Doorn, L-J. 1999. Recombination and clonal groupings within *Helicobacter pylori* from different geographical regions. *Molecular Microbiology* 32: 459-470.

Agostini, H. T., Yanagihara, R., Davis, V., Ryschkewitsch, C. F., Stoner, G. L. 1997. Asian genotypes of JC virus in Native Americans and in a Pacific Island population: markers of viral evolution and human migration. *Proceedings of the National Academy of Science (USA)* 94: 14542-14546.

Akopyants, N. S., Clifton, S. W., Kersulyte, D., Crabtree, J. E., Youree, B. E., Reece, C. A., Bukanov, N. O., Drazek, E. S., Roe, B. A., Berg, D. E. 1998. Analyses of the *cag* pathogenicity island of *Helicobacter pylori*. *Molecular Microbiology* 28: 37-53.

Alarcón, T., Domingo, D. Martínez, M. J., López-Brea, M. 1999. *cagA* gene and *vacA* alleles in Spanish *Helicobacter pylori* isolates from patients of different ages. *FEMS Immunology and Medical Microbiology* 24: 215-219.

Alm, R. A., Ling, L. S., Moir, D. T., King, B. L., Brown, E. D., Doig, P. C., Smith, D. R., Noonan, B., Guild, B. C., deJonge, B. L., Carmel, G., Tummino, P. J., Caruso, A., Uria-Nickelsen, M., Mills, D. M., Ives, C., Gibson, R., Merberg, D., Mills, S. D., Jiang, Q., Taylor, D. E., Vovis, G. F., Trust, T. J. 1999. Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*. *Nature* 397: 176-180.

Atherton, J. C., Cao, P., Peek, R. M. J., Tummuru, M. K., Blaser, M. J., Cover, T. L. 1995. Mosaicism in vacuolating cytotoxin alleles of *Helicobacter pylori*. Association of specific *vacA* types with cytotoxin production and peptic ulceration. *Journal of Biological Chemistry* 270: 17771-17777.

Bartolini, M. C., Salzano, F. M., Thomas, M. G., Stuart, S. Nasanen, S. P., Bau, C. H., Hutz, M. H., Layrisse, Z., Petzl-Erler, M. L., Tsuneto, L. T., Hill, K., Hurtado, A. M., Castro-de-Guerra, D., Torres, M. M., Groot, H., Michalski, R., Nymadawa, P., Bedoya, G., Bradman, N., Labuda, D., Ruiz-Linares, A. 2003. Y-chromosome evidence for differing ancient demographic histories in the Americas. *American Journal of Human Genetics* 73: 524-539.

Boren, T., Falk, P., Roth, K. A., Larson, G., Normark, S. 1993. Attachment of *Helicobacter pylori* to human gastric epithelium mediated by blood group antigens. *Science* 262: 1892-1895.

Bunn, J. E. G., Mackay, W. G., Thomas, J. E., Reid, D. C., Weaver, L. T. 2002. Detection of *Helicobacter pylori* DNA in drinking water biofilms: implications for transmission in early life. *Letters in Applied Microbiology* 34: 450-454.

Cann, R. L. 2001. Genetic clues to dispersal in human populations: retracing the past from the present. *Science* 291: 1742-1748.

Castillo-Rojas, G., Cerbón, M., López-Vidal, Y. 2008. Presence of *Helicobacter pylori* in a Mexican pre-Columbian mummy. *BMC Microbiology* 8: 119.

Cavalli-Sforza, L. L. y Feldman, M. W. 2003. The application of molecular genetic approaches to the study of human evolution. *Nature Genetics* 33 Supplement: 266-275.

Chan, S-Y., Ho, L., Ong, C-K., Chow, V., Drescher, B., Dürst, M., Ter Meulen, J., Villa, L., Luande, J., Mgya, H. N., Bernard, H-U. 1992. Molecular variants of human papillomavirus type 16 from four continents suggest ancient pandemic spread of the virus and its coevolution with humankind. *Journal of Virology* 66: 2057-2066.

Correa, P., Willis, D., Allison, M. Gerszten, E. 1998. *Helicobacter pylori* in pre-Columbian mummies [AGA Abstract #G3919]. *Gastroenterology* 114: A956.

Covacci, A., Falkow, S., Berg, D. E., Rappuoli, R. 1997. Did the inheritance of a pathogenicity island modify the virulence of *Helicobacter pylori*? *Trends in Microbiology* 5: 205-208.

Covacci, A., Telford, J. L., Del Giudice, G., Parsonnet, J., Rappuoli, R. 1999. *Helicobacter pylori* virulence and genetic geography. *Science* 284: 1328-1333.

Cover, T. L., Blaser, M. J. 1996. *Helicobacter pylori* infection, a paradigm for chronic mucosal inflammation: pathogenesis and implications for eradication and prevention. *Advanced Internal Medicine* 41: 85-117.

de Thé, G. 2007. Microbial genomes to write our history. *Journal of Infectious Diseases* 196: 499-501.

Delport, W., Cunningham, M., Olivier, B., Preisig, O., van der Merwe, S. W. 2006. A population genetics pedigree perspective on the transmission of *Helicobacter pylori*. *Genetics* 174: 2107-2118.

Delport, W., van der Merwe, S. W. 2007. The transmission of *Helicobacter pylori*: the effects of analysis method and study population on inference. *Best Practice & Research Clinical Gastroenterology* 21: 215-236.

de Queiroz, A., Gatesy, J. 2007. The supermatrix approach to systematics. *Trends in Ecology & Evolution* 22: 34-41.

Devi, S. M., Ahmed, I., Khan, A. A., Asad Rahman, S., Alvi, A., Sechi, L. A., Ahmed, N. 2006. Genomes of *Helicobacter pylori* from native Peruvians suggest admixture of ancestral and modern lineages and reveal a western type cag-pathogenicity island. *BMC Genomics* 7: 191.

Devi, S., Ahmed, I., Francalacci, P., Hussain, M., Akhter, Y., Alvi, A., Sechi, L., Mégraud, F., Ahmed, N. 2007. Ancestral European roots of *Helicobacter pylori* in India. *BMC Genomics* 8:184.

Dijkshoorn, L., Ursing, B. M., Ursing, J. B. 2000. Strain, clone and species: comments on three basic concepts of bacteriology. *Journal of Medical Microbiology* 49: 397-401.

Disotell, T. R. 2003. Discovering human history from stomach bacteria. *Genome Biology* 4: 213.

Domínguez-Bello, M. G., Pérez, M. E., Bortolini, M. C., Salzano, F. M., Pericchi, L. R., Zambrano-Guzmán, O., Linz, B. 2008. Amerindian *Helicobacter pylori* strains go extinct, as european strains expand their host range. *PLoS ONE*: e3307.

Eppinger, M., Baar, C., Linz, B., Raddatz, G., Lanz, C., Keller, H., Morelli, G., Gressmann, H., Achtman, M., Schuster, S. C. 2006. Who ate whom? Adaptive *Helicobacter* genomic changes that accompanied a host jump from early humans to large felines. *PLoS Genetics* 2: e120.

Etherington, G. J., Dicks, J., Roberts, I. N. 2004. Recombination Analysis Tool (RAT): a program for the high-throughput detection of recombination. *Bioinformatics* 21: 278-281.

Falush, D., Kraft, C., Taylor, N. S., Correa, P., Fox, J. G., Achtman, M., Suerbaum, S. 2001. Recombination and mutation during long-term gastric colonization by *Helicobacter pylori*: estimates of

clock rates, recombination size, and minimal age. *Proceedings of the National Academy of Science (USA)* 98: 15056-15061.

Falush, D., Wirth, T., Linz, B., Pritchard, J.K., Stephens, M., Kidd, M., Blaser, M.J., Graham, D.Y., Vacher, S., Perez-Perez, G.I., Yamaoka, Y., Megraud, F., Otto, K., Reichard, U., Katzowitsch, E., Wang, X., Achtman, M., and Suerbaum, S. 2003. Traces of human migrations in *Helicobacter pylori* populations. *Science* 299: 1582–1585.

Farris, J. S., Källersjö, M., Kluge, A. G., Bult, C. 1995. Testing significance of incongruence. *Cladistics* 10: 315-319.

Feldman, R. A. 2001. Epidemiological observations and open questions about disease and infection caused by *Helicobacter pylori*. En: *Helicobacter pylori: Molecular and Cellular Biology*. Achtman, M., Suerbaum, S. editores. Wymondham Horizon, pp. 29-51.

Fischbach, W., Malfertheiner, P., Hoffmann, J. C., Bolten, W., Kist, M., Koletzko, S. 2009. *Helicobacter pylori* and gastroduodenal ulcer disease. *Deutsches Ärzteblatt International* 106: 801-808.

Fisher, M. C., Koenig, G. L., White, T. J., San-Blas, G., Negroni, R., Gutierrez Alvarez, I., Wanke, B., and J. W. Taylor. 2001. Biogeographic range expansion into South America by *Coccidioides immitis* mirrors New World patterns of human migration. *Proceedings of the National Academy of Science (USA)* 98: 4558-4562.

Foster, P. 2001. The idiot's guide to the zen of likelihood in a nutshell in seven days for dummies, unleashed. Distributed by the author. Department of Zoology, Natural History Museum.

Gandon, S. 1998. Local adaptation and host-parasite interactions. *Trends in Ecology & Evolution* 13: 214-216.

Gatesy, J., Matthee, C., DeSalle, R., Hayashi, C. 2002. Resolution of a Supertree/Supermatrix Paradox. *Systematic Biology* 51: 652-664.

Go, M. F., Kapur, V., Graham, D. Y., Musser, J. M. 1996. Population genetic analysis of *Helicobacter pylori* by multilocus enzyme electrophoresis: extensive allelic diversity and recombinational population structure. *Journal of Bacteriology* 178: 3934-3938.

Ghose, C., Perez-Perez, G. I., Dominguez-Bello, M-G., Pride, D. T., Bravi, C. M., Blaser, M. J. 2002. East Asian genotypes of *Helicobacter pylori* strains in Amerindians provide evidence for its ancient human carriage. *Proceedings of the National Academy of Science (USA)* 99: 15107-15111.

Goosen, C., Theron, J., Ntasala, M., Maree, F. F., Olckers, A., Botha, S. J., Lastovica, A. J., van der Merwe, S. W. 2002. Evaluation of a novel heminested PCR assay based on the phosphoglucosamine mutase gene for detection of *Helicobacter pylori* in saliva and dental plaque. *Journal of Clinical Microbiology* 40: 205-209.

Holmes, E. C. 2004. The phylogeography of human viruses. *Molecular Ecology* 13: 745-756.

Hopkins, R. J., Vial, P. A. Ferreccio, C., Ovalle, J., Prado, P., Sotomayor, V., Russel, R. G., Wasserman, S. S., Morris, J. G. J. 1993. Seroprevalence of *Helicobacter pylori* in Chile: vegetables may serve as one route transmission. *Journal of Infectious Diseases* 168: 222-226.

Huelsenbeck, J. 2000. MrBayes: Bayesian inference of phylogeny. Distributed by the author. Department of Biology, University of Rochester.

Israel, D. A., Salama, N., Krishna, U., Rieger, U. M., Atherton, J. C., Falkow, S., Peek, R. M. J. 2001. *Helicobacter pylori* genetic diversity within the gastric niche of a single human host. *Proceedings of the National Academy of Science (USA)* 98: 14625-14630.

Ito, A., Fujioka, T., Kodama, K., Nishizono, A., Nasu, M. 1997. Virulence-associated genes as markers of strain diversity in *Helicobacter pylori* infection. *Journal of Gastroenterology and Hepatology* 12: 666-669.

Ingman, M., Kaessmann, H., Pääbo, S., Gyllensten, U. 2000. Mitochondrial genome variation and the origin of modern humans. *Nature* 408: 708-713.

Ingman, M., Gyllensten, U. 2001. Analysis of the complete human mtDNA genome: methodology and inferences for human evolution. *Journal of Heredity* 92: 454-461.

Jorde, L. B., Rogers, A. R., Bamshad, M., Scott Watkins, W., Krakowiak, P., Sung, S., Kere, J., Harpending, H. C. 1997. Microsatellite diversity and the demographic history of modern humans. *Proceedings of the National Academy of Science (USA)* 94: 3100-3103.

Karafet, T.M., Mendez, F.L., Meilerman, M.B., Underhill, P.A., Zegura, S.L., and Hammer, M.F. (2008). New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. *Genome Research* 18: 830-838.

Ke, Y., Su, B., Song, X., Lu, D., Chen, L., Li, H., Qi, C., Marzuki, S., Deka, R., Underhill, P., Xiao, C., Shriver, M., Lell, J., Wallace, D., Wells, R. S., Seielstad, M., Oefner, P., Zhu, D., Jin, J., Huang, W., Chakraborty, R., Chen, Z., Jin, L. 2001. African origin of modern humans in East Asia: a tale of 12,000 Y chromosomes. *Science* 292: 1151-1153.

Kersulyte, D., Chalkauskas, H., Berg, D. E. 1999. Emergence of recombinant strains of *Helicobacter pylori* during human infection. *Molecular Microbiology* 31: 31-43.

Kersulyte, D., Mukhopadhyay, A., Velapatiño, B., Su, W., Pan, Z., García, C., Hernández, V., Valdez, Y., Mistry, R., Gilman, R., Yuan, Y., Gao, H., Alarcón, T., López-Brea, M., Balakrish Nair, G., Chowdhury, A., Datta, S., Shirai, M., Nakazawa, R., Ally, R., Segal, I., Wong, B., Kam, S., Olfat, F., Borén, T., Engstrand, L., Torres, O., Schneider, R., Thomas, J., Czinn, S., Berg, D. 2000. Differences in genotypes of *Helicobacter pylori* from different human populations. *Journal of Bacteriology* 182: 3210-3218.

Kimura, M. 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution* 16: 111-120.

Kirkpatrick, M., Barton, N. H. 1997. Evolution of a species range. *American Naturalist* 150: 1-23.

Kivi, M., Tindberg, Y., Sörberg, M., Casswall, T. H., Befrits, R., Hellström, P. M., Bengtsson, C., Engstrand, L., Granström, M. 2003. Concordance of *Helicobacter pylori* strains within families. *Journal of Clinical Microbiology* 41: 5604-5608.

Konno, M., Fujii, N., Yokota, S., Sato, K., Takahashi, M., Sato, K., Mino, E., Sugiyama, T. 2005. Five-year follow-up study of mother-to-child transmission of *Helicobacter pylori* infection detected by a Random Amplified Polymorphic DNA fingerprinting method. *Journal of Clinical Microbiology* 43: 2246-2250.

Kraft, C., Stack, A., Josenhans, C., Niehus, E., Dietrich, G., Correa, P., Fox, J. G., Falush, D., Suerbaum, S. 2006. Genomic changes during chronic *Helicobacter pylori* infection. *Journal of Bacteriology* 188: 249-254.

Kumar, S., Nei, M., Dudley, J., Tamura, K. 2008. MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. *Briefings in Bioinformatics* 9: 299-306.

Li, O. L., Genta, R. M., Go, M. F., Gutierrez, O., Kim, J. G., Graham, D. Y. 2002. *Helicobacter pylori* strain and the pattern of gastritis among first-degree relatives of patients with gastric carcinoma. *Helicobacter* 7: 349-355.

Li, J.Z., Absher, D.M., Tang, H., Southwick, A.M., Casto, A.M., Ramachandran, S., Cann, H.M., Barsh, G.S., Feldman, M., Cavalli-Sforza, L.L., and Myers, R.M. 2008. Worldwide human relationships inferred from genome-wide patterns of variation. *Science* 319: 1100-1104.

Linz, B., Balloux, F., Moodley, Y., Manica, A., Liu, H., Roumagnac, P., Falush, D., Stamer, C., Prugnolle, F., van der Merwe, S. W., Yamaoka, Y., Graham, D. Y., Perez-Trallero, E., Wadstrom, T., Suerbaum, S., Achtman, M. 2007. An African origin for the intimate association between humans and *Helicobacter pylori*. *Nature* 445: 915-918.

Liu, H., Prugnolle, F., Manica, A., Piazza, A. 2006. A geographically explicit genetic model of worldwide human-settlement history. *American Journal of Human Genetics* 79: 230-237.

Magalhaes Queiroz, D. M. y Luzzza, F. 2006. Epidemiology of *Helicobacter pylori* infection. *Helicobacter* 11 Supplement 1: 1-5.

Matisoo-Smith, E., Roberts, R. M., Irwin, G. J., Allen, J. S., Penny, D. y Lambert, D. M. 1998. Patterns of prehistoric human mobility in Polynesia indicated by mtDNA from the Pacific rat. *Proceedings of the National Academy of Science (USA)* 95: 15145-15150.

Mellars, P. 2006. Why did modern human populations disperse from Africa ca. 60000 years ago? A new model. *Proceedings of the National Academy of Science (USA)* 103: 9381-9386.

Miura, T., Fukunaga, T., Igarashi, T., Yamashita, M., Ido, E., Funahashi, S., Ishida, T., Washio, K., Ueda, S., Hashimoto, K. 1994. Phylogenetic subtypes of human T-lymphotropic virus type I and their relations to the anthropological background. *Proceedings of the National Academy of Science (USA)* 91: 1124-1127.

Moodley, Y., Linz, B., Yamaoka, Y., Windsor, H., Breurec, S., Wu, J-Y., Maady, A., Bernhöft, S., Thiberge, J-M., Phuanukoonnon, S., Jobb, G., Siba, P., Graham, D., Marshall, B., Achtman, M. 2009. The peopling of the Pacific from a bacterial perspective. *Science* 323: 527-530.

Moraga, M, Llop, E., Carvallo, P., Rothhammer, F. 2004. Polimorfismos de ADN mitocondrial. En: *Poblaciones Chilenas. Cuatro Décadas de Investigaciones Bioantropológicas*. Rothhammer, F., Llop, E. editores. Editorial Universitaria, pp. 189-200.

O'Rourke, J., Bode, G. 2001. Morphology and Ultrastructure. En: *Helicobacter pylori: Physiology and Genetics*. Mobley, H., Mendz, G., Hazell, S. editores. ASM Press, pp. 53-67.

Pakendorf, B., and Stoneking, M. 2005. Mitochondrial DNA and human evolution. *Annual Review of Genomics and Human Genetics* 6: 165-183.

Posada, D., Crandall, K. A. 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics Applications Note* 14: 817-818.

Posada, D., Crandall, K. 2002. The effect of recombination on the accuracy of phylogeny estimation. *Journal of Molecular Evolution* 54: 396-402.

Posada, D., Crandall, K. A., Holmes, E. C. 2002. Recombination in evolutionary genomics. *Annual Review of Genetics* 36: 75-97.

Raymond, O. J., Thiberg, J. M., Chevalier, C., Kalach, N., Bergeret, M., Labigne, A., Dauga, C. 2004. Genetic and transmission analysis of *Helicobacter pylori* strains within a family. *Emerging Infectious Diseases* 10: 1816-1821.

Reich, D. E., Cargill, M., Bolk, S., Ireland, J., Sabeti, P. C., Richter, D. J., Lavery, T., Kouyoumjian, R., Farhadian, S. F., Ward, R., Lander, E. S. 2001. Linkage disequilibrium in the human genome. *Nature* 411: 199-204.

Reidla, M., Kivisild, T., Metspalu, E., Kaldma, K., Tambets, K., Tolk, H. V., Parik, J., Loogväli, E. L., Derenko, M., Malyarchuk, B., Bermisheva, M., Zhadanov, S., Pennarun, E., Gubina, M., Golubenko, M., Damba, L., Fedorova, S., Gusar, V., Grechanina, E., Mikerezi, I., Moisan, J. P., Chaventré, A., Khusnutdinova, E., Osipova, L., Stepanov, V., Voevoda, M., Achilli, A., Rengo, C., Rickards, O., De Stefano, G. F., Papiha, S., Beckman, L., Janicijevic, B., Rudan, P., Anagnou, N., Michalodimitrakis, E.,

Koziel, S., Usanga, E., Geberhiwot, T., Herrstadt, C., Howell, N., Torroni, A., Villems, R. 2003. Origin and diffusion of mtDNA haplogroup X. *American Journal of Human Genetics* 73: 1178-1190.

Robinson, L-G., Black, F., Lee, F., Sousa, A., Owens, M., Danielsson, D, Nahmias, A., Gold, B. 2002. *Helicobacter pylori* prevalence among indigenous peoples of South America. *Journal of Infectious Diseases* 186: 1131-1137.

Rocco, P., Morales, G., Moraga, M., Miquel, J. F., Nervi, F., Llop, E., Carvallo, P., Rothhammer, F. 2002. Composición genética de la población chilena: distribución de polimorfismos de DNA mitocondrial en grupos aborígenes y en la población mixta de Santiago. *Revista Médica de Chile* 130: 125-131.

Rosenberg, N. A., Pritchard, J. K., Weber, J. L., Cann, H. M., Kidd, K. K., Zhivotovsky, L. A. 2002. Genetic structure of human populations. *Science* 298: 2381-2384.

Roumagnac, P., Weill, F. X., Dolecek, C., Baker, S., Brisse, S., Chinh, N. T., Le, T. A., Acosta, C. J., Farrar, J., Dougan, G., Achtman, M. 2006. Evolutionary history of *Salmonella typhi*. *Science* 314: 1301-1304.

Rozen, S., Skaletsky, H. 2000. Primer3. En: *Bioinformatics Methods and Protocols: Methods in Molecular Biology*. Krawetz, S., Misener, S. editores. Humana Press, pp. 365-386.

Ruiz-Linares, A., Ortiz-Barrientos, D., Figueroa, M., Mesa, N., Múnera, J. G., Bedoya, G., Vélez, I. D., García, L. F., Pérez-Lezaun, A., Bertranpetit, J., Feldman, M. W., Goldstein, D. B. 1999. Microsatellites provide evidence for Y chromosome diversity among the founders of the New World. *Proceedings of the National Academy of Science (USA)* 96: 6312-6317.

Saitou, N., Nei, M. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* 4: 406-425.

Schierup, M. E., Hein, J. 2000. Consequences of recombination on traditional phylogenetic analysis. *Genetics* 156: 879-891.

Schreiber, S., Konradt, M., Groll, C., Scheid, P., Hanauer, G., Werling, H-O., Josenhans, C., Suerbaum, S. 2003. The spatial orientation of *Helicobacter pylori* in the gastric mucus. *Proceedings of the National Academy of Science (USA)* 101: 5024-5029.

Schurr, T., Sherry, S. T. 2004. Mitochondrial DNA and Y chromosome diversity and the peopling of the Americas: evolutionary and demographic evidence. *American Journal of Human Biology* 16: 420-439.

Shimodaira, H., Hasegawa, M. 2001. CONSEL: for assessing the confidence of phylogenetic tree selection. *Bioinformatics* 17: 1246-1247.

Suerbaum, S., Smith, J. M., Bapumia, K., Morelli, G., Smith, N. H., Kunstmann, E., Dyrek, I., Achtman, M. 1998. Free recombination within *Helicobacter pylori*. *Proceedings of the National Academy of Science (USA)* 95: 12619-12624.

Suerbaum, S., Josenhans, C. 2007. *Helicobacter pylori* evolution and phenotypic diversification in a changing host. *Nature Reviews: Microbiology* 5: 441-452.

Tamura, K., Dudley, J., Nei, M., Kumar, S. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) Software Version 4.0. *Molecular Biology and Evolution* 24: 1596-1599.

Thompson, J. D., Higgins, D. G., Gibson, T. J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Research* 22: 4673-4680.

Tindberg, Y., Bengtsson, C., Granath, F., Blennow, M., Nyren, O., Granstrom, M. 2001. *Helicobacter pylori* infection on Swedish school children: lack of evidence of child-to-child transmission outside the family. *Gastroenterology* 121: 310-316.

Tomb, J. F., White, O., Kerlavage, A. R., Clayton, R. A., Sutton, G. G., Fleischmann, R. D., Ketchum, K. A., Klenk, H. P., Gill, S., Dougherty, B. A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E. F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H. G., Glodek, A., McKenney, K., Fitzegerald, L. M., Lee, N., Adams, M. D., Hickey, E. K., Berg, D.E., Gocayne, J. D., Utterback, T. R., Peterson, J. D., Kelley, J. M., Cotton, M. D., Weidman, J. M., Fujii, C., Bowman, C., Watthey, L., Wallin, E., Hayes, W. S., Borodovsky, M., Karp, P. D., Smith, H. O., Fraser, C. M., Venter, J. C. 1997. The complete genome sequence of the gastric pathogen *Helicobacter pylori*. *Nature* 388: 539-549

Underhill, P. A., Shen, P., Lin, A. A., Jin, L., Passarino, G., Yang, W. H., Kauffman, E., Bonn -Tamir, B., Bertranpetit, J., Francalacci, P., Ibrahim, M., Jenkins, T., Kidd, J. R., Mehdi, S. Q., Seielstad, M. T., Spencer Wells, R., Piazza, A., Davis, R. W., Feldman, M. W., Cavalli-Sforza, L. L., Oefner, P. J. 2000. Y chromosome sequence variation and the history of human populations. *Nature Genetics* 26: 358-361.

van Doorn, L-J., Figueiredo, C., Sanna, R., Pena, S., Midolo, P., NG, E. K. W., Atherton, J. C., Blaser, M. J., Quint, W. G. V. 1998. Expanding allelic diversity of *Helicobacter pylori* vacA. *Journal of Clinical Microbiology* 36: 2597-2603.

van Doorn, L-J., Figueiredo, C., Megraud, F., Pena, S., Midolo, P., Queiroz, D. M., Carneiro, F., Vanderborcht, B., Pegado, M. D., Sanna, R., De Boer, W., Schneeberger, P. M., Correa, P., Ng, E. K. W., Atherton, J., Blaser, M. J., Quint, W.G. 1999. Geographic distribution of vacA allelic types of *Helicobacter pylori*. *Gastroenterology* 116: 823-30.

van Doorn, L-J., Schneeberger, P. M., Nouhan, N., Plaisier, A. P., Quint, W. G. V., de Boer, W. A. 2009. Importance of *Helicobacter pylori* cagA and vacA status for the efficacy of antibiotic treatment. *Gut* 2000 46: 321-326.

Wang, S., Lewis, C. M., Jakobson, M., Ramachandran, S., Ray, N., Bedoya, G., Rojas, W., Parra, M. V., Molina, J. A., Gallo, C., Mazzotti, G., Poletti, G., Hill, K., Hurtado, A. M., Labuda, D., Klitz, W., Barrantes, R., Bortolini, M. C., Salzano, F. M., Petzl-Erler, M. L., Tsuneto, L. T., Llop, E., Rothhammer, F., Excoffier, L., Feldman, M. W., Rosenberg, N. A., Ruiz-Linares, A. 2007. Genetic variation and population structure in Native Americans. *PLoS Genetics* 3: e185.

Wiens, J. J. 1998. Combining data sets with different phylogenetic histories. *Systematic Biology* 47: 568-581.

Wirth, T., Wang, X., Linz, B., Novick, R. P., Lum, J. K., Blaser, M., Morelli, G., Falush, D., Achtman, M. 2004. Distinguishing human ethnic groups by means of sequences from *Helicobacter pylori*: lessons from Ladakh. *Proceedings of the National Academy of Science (USA)* 101: 4746-4751.

Wirth, T., Meyer, A., Achtman, M. 2005. Deciphering host migrations and origins by means of their microbes. *Molecular Ecology* 14: 3289-3306.

Xia, X., Z. Xie. 2001. DAMBE: Data analysis in molecular biology and evolution. *Journal of Heredity* 92: 371-373.

Xiabo, L., Wenzhong, L., Weiwen, X., Yao, S., Shudong, X. 2001. Clinical implications and prevalence of cagA, cagE and cagT genes in the pathogenicity island of *Helicobacter pylori* strains isolated from Shanghai patients. *Chinese Journal of Digestive Diseases* 2: 133-136.

Yamaoka, Y., Orito, E., Mizokami, M., Gutierrez, O., Saitou, N., Kodama, T., Osato, M. S., Kim, J. G., Ramirez, F. C., Mahachai, V., Graham, D. Y. 2002. *Helicobacter pylori* in North and South America before Columbus. *FEBS Letters* 517: 180-184.

Zhao, Z., Jin, L., Fu, Y-X., Ramsay, M., Jenkins, T., Leskinen, E., Pamilo, P., Trexler, M., Patthy, L., Jorde, L., Ramos-Onsins, S., Yu, N., Li, W-H. 2000. Worldwide DNA sequence variation in a 10-kilobase noncoding region on human chromosome 22. *Proceedings of the National Academy of Science (USA)* 97: 11354-11358.

Zhivotovsky, L. A., Rosenberg, N. A., Feldman, M., W. 2003. Features of evolution and expansion of modern humans, inferred from genome-wide microsatellite markers. *American Journal of Human Genetics* 72: 1171-1186.