

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 377 751**

51 Int. Cl.:  
**C12N 15/31** (2006.01)  
**C07K 14/195** (2006.01)  
**A61K 38/00** (2006.01)  
**A61K 38/16** (2006.01)  
**A61K 39/00** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **08100586 .0**  
96 Fecha de presentación: **10.12.1998**  
97 Número de publicación de la solicitud: **1908772**  
97 Fecha de publicación de la solicitud: **09.04.2008**

54 Título: **Polipéptidos y polinucleótidos de Porfiromonas gingivalis**

30 Prioridad:  
10.12.1997 AU PP083997 31.12.1997 AU PP118297  
30.01.1998 AU PP154698 10.03.1998 AU PP226498  
09.04.1998 AU PP291198 23.04.1998 AU PP312898  
05.05.1998 AU PP333898 22.05.1998 AU PP365498  
29.07.1998 AU PP491798 30.07.1998 AU PP496398  
04.08.1998 AU PP502898

45 Fecha de publicación de la mención BOPI:  
**30.03.2012**

45 Fecha de la publicación del folleto de la patente:  
**30.03.2012**

73 Titular/es:  
**CSL LIMITED**  
**45 POPLAR ROAD**  
**PARKVILLE, VIC 3052, AU y**  
**THE UNIVERSITY OF MELBOURNE**

72 Inventor/es:  
**Ross, Bruce Carter;**  
**Barr, Ian George;**  
**Patterson, Michelle Anne;**  
**Agius, Catherine Therese;**  
**Rothel, Linda Joy;**  
**Margetts, Mai Brigid;**  
**Hocking, Dianna Margaret y**  
**Webb, Elizabeth Ann**

74 Agente/Representante:  
**Carvajal y Urquijo, Isabel**

ES 2 377 751 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

## DESCRIPCIÓN

Polipéptidos y Polinucleótidos de Porfiromonas gingivalis

Campo de la invención

5 La presente invención se relaciona con secuencias de nucleótidos p. Gingivalis, polipéptidos p. Gingivalis y sondas para la detección de p. Gingivalis. Los polipéptidos y nucleótidos de p. Gingivalis se pueden utilizar en composiciones para uso en elevar una respuesta inmune en un sujeto contra p. Gingivalis y tratar o evitar o reducir la gravedad de la afección conocida como periodontitis.

Antecedentes de la invención

10 Las enfermedades periodontales son enfermedades inflamatorias asociadas con bacterias de los tejidos de soporte del diente y varían desde la forma relativamente moderada de gingivitis, la inflamación reversible, no específica del tejido gingival hasta las formas más agresivas de periodontitis que se caracterizan por la destrucción de las estructuras de soporte del diente. La periodontitis se asocia con una infección subgingival de asociación de bacterias Gram-negativas específicas que conduce a la destrucción del periodonto y es un gran problema de salud pública. Una bacteria que atrae interés considerable es P. gingivalis cuando la recuperación de este microorganismo a partir de lesiones de periodontitis de adulto puede ser de hasta 50 % de la flora subgingival cultivable anaeróticamente, mientras que raramente se recupera de P. gingivalis, y luego en bajas cifras, de los sitios saludables. Se ha asociado un aumento proporcional en el nivel de P. gingivalis en la placa subgingival con un aumento grave de periodontitis y erradicación del microorganismo de la población microbiana subgingival cultivable acompañado de resolución de la enfermedad. Se ha demostrado la progresión de las lesiones por periodontitis en primates diferentes a los humanos con la implantación subgingival de P. gingivalis. Estos hallazgos en animales y humanos sugieren una función principal de P. gingivalis en el desarrollo de la periodontitis en adultos.

P. gingivalis es un bacilo Gram-negativo proteolítico, asacarolítico, anaeróbico, pigmentado en negro que obtiene energía del metabolismo de aminoácidos específicos. El microorganismo tiene un requerimiento absoluto de hierro para crecimiento, preferiblemente en la forma de hemo o su Fe (III).

25 Con el fin de desarrollar una vacuna segura y eficaz para evitar, eliminar o reducir la colonización por P. gingivalis es necesario identificar y producir antígenos que estén implicados en la virulencia que tengan utilidad como inmunógenos posiblemente a través de la generación de anticuerpos específicos. Aunque es posible intentar aislar los antígenos directamente de los cultivos de P. gingivalis es frecuentemente difícil. Por ejemplo como se mencionó anteriormente, P. gingivalis es un anaerobo estricto y puede ser difícil de aislar y cultivar. También se sabe que, para un número de organismos, cuando son cultivados in vitro que muchos genes de virulencia se regulan por disminución y las proteínas codificadas no se expresan más. Si se aplican técnicas de química convencional para purificación no se pueden identificar los candidatos de vacuna de moléculas potencialmente importantes (protectoras). Con el secuenciamiento de ADN, cuando está presente el gen (pero no transcrito) aún cuando el organismo se cultiva in vitro se puede identificar, clonar y producir como una proteína de ADN recombinante. De forma similar, se puede expresar transitoriamente un antígeno protector u objetivo terapéutico por el organismo in vitro o producir en bajos niveles que hacen la identificación de estas moléculas extremadamente difíciles mediante métodos convencionales.

40 Con la identificación serológica de los objetivos terapéuticos uno se limita a aquellas respuestas que son detectables utilizando métodos estándar tales como Western Blot o ELISA. La limitación aquí es el nivel de respuesta que se genera por el animal o humano y determina si esta respuesta es protectora, dañina o irrelevante. No está presente tal limitación con un método de secuenciamiento para la identificación de los objetivos profilácticos o terapéuticos potenciales.

45 También se sabe que el P. gingivalis produce un rango de proteasas ampliamente activas (University of Melbourne, Solicitud de Patente Internacional No WO97/16542, Patente Estadounidense Nos 5,475,097 y 5,523,390), que hace que sea difícil la identificación de las proteínas intactas debido a su degradación por estas proteasas.

Resumen de la invención

50 Los presentes inventores han intentado aislar las secuencias de nucleótidos P. gingivalis que se pueden utilizar para producción recombinante de los polipéptidos P. gingivalis y para desarrollar sondas de nucleótido específicas para P. gingivalis. Las secuencias de ADN mencionadas adelante se han seleccionado de un gran número de secuencias P. gingivalis de acuerdo con su potencial indicador como candidatos de vacuna. Esta etapa intuitiva implica la comparación desde la secuencia de proteína deducida de las secuencias de ADN P. gingivalis hasta las bases de datos de secuencia de proteínas conocidas. Algunas de las características utilizadas para seleccionar los candidatos de vacuna útiles incluyen; la ubicación celular esperada, tal como proteínas de membrana externas o proteínas secretadas, las actividades funcionales particulares de proteínas similares tales como aquellas con una actividad enzimática o proteolítica, las proteínas implicadas en las rutas metabólicas esenciales que cuando se inactivan o bloquean pueden ser perjudiciales o letales para el organismo, las proteínas que se pueden esperar que cumplan

una función en la patogenia del organismo por ejemplo lisis de glóbulos rojos, aglutinación celular o receptores celulares y proteínas que son parálogas a las proteínas con eficacia de vacuna mejorada.

En un primer aspecto la presente invención proporciona un polipéptido *Porfiromonas gingivalis* antigénico aislado, el polipéptido comprende;

- 5 (i) la secuencia de aminoácidos de la SEQ ID NO. 426;
- (ii) la secuencia de aminoácidos de los residuos 24 a 821 o 27 a 821 de la SEQ ID NO. 426; o
- (iii) una secuencia de aminoácidos por lo menos 85 % idéntica a una secuencia de aminoácidos de (i).

En una realización de la presente invención, el polipéptido comprende una secuencia de aminoácidos por lo menos 95 % idéntica a la secuencia de aminoácidos de la (i).

- 10 Como se utiliza aquí el % de identidad para los polipéptidos se calcula utilizando el algoritmo de alineación de Needleman y Munsch (9) utilizando una matriz de clasificación de proteína estándar (Blosum 50).

En una realización preferida de la presente invención el polipéptido comprende una secuencia de aminoácidos seleccionada del grupo que consiste de la SEQ. ID. NO. 426 o SEQ ID NO. 301.

- 15 En un segundo aspecto la presente invención proporciona una molécula de ADN aislada, la molécula de ADN comprende una secuencia de nucleótido que codifica el polipéptido del primer aspecto de la presente invención.

Se prefiere que la molécula de ADN aislada comprenda la secuencia de nucleótido de la SEQ. ID. NO. 162 o SEQ ID NO. 37.

- 20 En un tercer aspecto, la presente invención proporciona un vector de expresión recombinante que comprende la molécula de ADN del segundo aspecto de la presente invención ligada operablemente a un elemento regulador de transcripción.

La presente invención también proporciona una célula aislada que comprende este vector de expresión recombinante.

En un cuarto aspecto, la presente invención proporciona un método para producir un polipéptido *P. gingivalis* que comprende cultivar la célula bajo condiciones que permiten la expresión del polipéptido.

- 25 En aún un aspecto adicional, la presente invención proporciona una composición farmacéutica que comprende (i) una cantidad efectiva de (a) por lo menos un polipéptido del primer aspecto de la presente invención, (b) por lo menos una molécula de ADN del segundo aspecto de la presente invención, y/o (c) un polipéptido de por lo menos 40 aminoácidos que tiene una secuencia contigua de por lo menos 40 aminoácidos idénticos a una secuencia de aminoácidos contigua de la SEQ ID NO. 426, en donde el polipéptido es capaz de elevar una respuesta inmune contra *P. gingivalis*, y (ii) un portador farmacéuticamente aceptable. Se prefiere que el portador farmacéuticamente aceptable sea un adyuvante. En otros aspectos, la presente invención proporciona tal composición para uso en elevar una respuesta inmune dirigida a *P. gingivalis* en un sujeto, o para tratar infección por *P. gingivalis*. El tratamiento puede ser profiláctico o terapéutico.
- 30

- 35 En aún otro aspecto, la presente invención proporciona un anticuerpo elevado contra un polipéptido del primer aspecto de la invención. El anticuerpo puede ser policlonal o monoclonal. La presente invención también proporciona composiciones que incluyen estos anticuerpos. Se prefiere que estas composiciones se adapten para uso oral y puedan ser, por ejemplo, pasta dental, enjuague bucal, etc.

La presente invención también proporciona un método para detectar la presencia de una molécula de ADN del segundo aspecto de la invención en una muestra que comprende:

- 40 (a) poner en contacto una muestra con una sonda de nucleótido que comprende por lo menos 18 nucleótidos y que tiene una secuencia contigua de por lo menos 18 nucleótidos idénticos a una secuencia de nucleótido contigua seleccionada del grupo que consiste de la SEQ. ID. NO. 37 bajo condiciones en las que se puede formar un híbrido entre la sonda y la molécula de ADN del segundo aspecto de la invención en la muestra; y

- 45 (b) detectar el híbrido formado en la etapa (a), en donde la detección de un híbrido indica la presencia de la molécula de ADN del segundo aspecto de la invención en la muestra.

Descripción detallada

Definiciones

- 50 Se utilizan intercambiabilmente aquí una preparación sustancialmente pura o polipéptido aislado o purificado de un polipéptido y, como se utiliza aquí, significa un polipéptido que se ha separado de otras proteínas, lípidos, y ácidos nucleicos con los que ocurre en forma natural. Preferiblemente, el polipéptido también se separa de las sustancias,

por ejemplo, anticuerpos o matriz de gel, por ejemplo, poliacrilamida, que se utilizan para purificarlos. Preferiblemente, el polipéptido constituye por lo menos 10, 20, 50 70, 80 o 95 % en peso seco de la preparación purificada. Preferiblemente, la preparación contiene: suficiente polipéptido para permitir el secuenciamiento de la proteína; por lo menos 1, 10, o 100 mg del polipéptido.

- 5 Una preparación purificada de las células se refiere a, en el caso de la planta o células de animal, una preparación in vitro de células y no un animal o planta intacta completa. En el caso de células cultivadas o células microbianas, consiste de una preparación de por lo menos 10 % y más preferiblemente 50 % de las células objeto.

10 Un ácido nucleico sustancialmente puro o aislado o purificado, por ejemplo, un ADN sustancialmente puro, (son términos utilizados intercambiamente aquí) es un ácido nucleico que es uno o ambos de los siguientes: no inmediatamente contiguo con ambas secuencias codificantes con las que es inmediatamente contigua (es decir., uno en el extremo 5' y uno en el extremo 3') en el genoma que ocurre de forma natural del organismo del que se deriva el ácido nucleico; o que es sustancialmente libre de un ácido nucleico con el que ocurre en el organismo del que se deriva el ácido nucleico. El término incluye, por ejemplo, un ADN recombinante que se incorpora dentro de un vector, por ejemplo, en un plásmido o virus que se replica autónomamente, o dentro del ADN genómico de un procarionte o eucarionte, o que existe como una molécula separada (por ejemplo, un cADN o un fragmento de ADN genómico producido por PCR o tratamiento de endonucleasa de restricción) independiente de otras secuencias de ADN. El ADN sustancialmente puro también incluye un ADN recombinante que es parte de un gen híbrido que codifica la secuencia de ADN P. gingivalis adicional.

20 Un "contig" como se utiliza aquí es un ácido nucleico que representa un tramo contiguo de la secuencia genómica de un organismo.

Un "marco de lectura abierto", también denominado aquí como ORF, es una región del ácido nucleico que codifica un polipéptido. Esta región puede representar una porción de una secuencia codificante o una secuencia total y se puede determinar desde un codón de parada hasta el codón de parada o desde un codón de inicio hasta un codón de parada.

25 Como se utiliza aquí, una "secuencia codificante" es un ácido nucleico que se transcribe dentro del ARN mensajero y/o se traduce en un polipéptido cuando se pone bajo el control de las secuencias reguladoras apropiadas. Los límites de la secuencia codificante se determinan mediante un codón de inicio de traducción en los cinco terminales principales y un codón de parada de traducción en los tres terminales principales. Una secuencia codificante puede incluir pero no se limita al ADN sintético del ARN mensajero, y las secuencias de ácido nucleico recombinantes.

30 Un "complemento" de un ácido nucleico como se utiliza aquí se refiere a una secuencia anticodificante o antiparalela que participa en el par base Watson-Crick con la secuencia original.

Un "producto de gen" es una proteína o ARN estructural que se codifica específicamente por un gen.

35 Como se utiliza aquí, el término "sonda" se refiere a un ácido nucleico, péptido u otra entidad química que se une específicamente a una molécula de interés. Las sondas se asocian frecuentemente con o son capaces de asociarse con un marcador. Un marcador es una unidad estructural química capaz de detección. Los marcadores típicos comprenden tintes, radioisótopos, unidades estructurales luminiscentes y quimioluminiscentes, fluoróforos, enzimas, agentes de precipitación, secuencias de amplificación, y similares. De forma similar, un ácido nucleico, péptido u otra entidad química que se une específicamente a una molécula de interés e inmoviliza tal molécula se denomina aquí como un "ligando de captura". Los ligandos de captura se asocian típicamente con o son capaces de asociarse con un soporte tal como nitro-celulosa, vidrio, membranas de nylon, glóbulos, partículas y similares. La especificidad de hibridación es dependiente de condiciones tales como la composición de par base de los nucleótidos, y la temperatura y concentración de sal de la reacción. Estas condiciones son fácilmente discernibles para un experto en la técnica utilizando experimentación de rutina.

45 Homólogos se refieren a la similitud de secuencia o la identidad de secuencia entre dos polipéptidos o entre dos moléculas de ácido nucleico. Cuando se ocupa una posición en las dos secuencias comparadas por la misma base o subunidad de monómero de aminoácido, por ejemplo, si una posición en cada una de las dos moléculas de ADN se ocupa por adenina, entonces las moléculas son homólogas en esta posición. El porcentaje de homología entre las dos secuencias es una función del número de posiciones homólogas o coincidentes compartidas por las dos secuencias divididas por el número de posiciones comparado x 100.

50 Los términos péptidos, proteínas, y polipéptidos se utilizan intercambiamente aquí.

Un "componente inmunogénico" como se utiliza aquí es una unidad estructural, tal como un polipéptido P. gingivalis, análogo o fragmento del mismo, que es capaz de provocar una respuesta inmune celular y/o humoral en un animal anfitrión.

55 Un "componente antigénico" como se utiliza aquí es una unidad estructural, tal como el polipéptido P. gingivalis, análogo o fragmento del mismo, que es capaz de unirse a un anticuerpo específico con afinidad suficientemente alta para formar un complejo de antígeno-anticuerpo detectable.

Como se utiliza aquí, el término "promotor específico de célula" significa una secuencia de ADN que sirve como un promotor, es decir, regula la expresión de una secuencia de ADN seleccionada ligada operativamente al promotor, y que afecta la expresión de la secuencia de ADN seleccionada en células específicas de un tejido. El término también cubre los así llamados promotores de "fuga", que regulan la expresión de un ADN seleccionado principalmente en un tejido, pero que provocan la expresión también en otros tejidos.

Como se utiliza aquí, el término "secuencia de control" se refiere a un ácido nucleico que tiene una secuencia base que se reconoce por el organismo anfitrión para afectar la expresión de las secuencias codificadas a las que se ligan. La naturaleza de tales secuencias de control difiere dependiendo del organismo anfitrión; en procariontes, tales secuencias de control incluyen de manera general un promotor, sitio de unión ribosómico, terminadores, y en algunos casos operadores; en eucariotes, de manera general tales secuencias de control incluyen promotores, terminadores y en algunos casos, mejoradores. El término secuencia de control se pretende que incluya como mínimo, todos los componentes cuya presencia es necesaria para expresión, y también pueden incluir componentes adicionales cuya presencia es ventajosa, por ejemplo, las secuencias líder.

Como se utiliza aquí, el término "ligado operativamente" se refiere a las secuencias unidas o ligadas a la función en su forma pretendida. Por ejemplo, una secuencia de control se liga operativamente a la secuencia codificante mediante ligación en tal forma que la expresión de la secuencia codificante se logra bajo condiciones compatibles con la secuencia de control y la célula anfitriona.

Una "muestra" como se utiliza aquí se refiere a una muestra biológica, tal como, por ejemplo, tejido o fluido aislado de un individuo (que incluye sin limitación plasma, suero, fluido cerebroespinal, linfa, lágrimas, saliva y secciones de tejido) o de constituyentes de cultivo celular in vitro, así como también muestras del ambiente.

La práctica de la invención empleará, a menos que se indique otra cosa, técnicas convencionales de química, biología molecular, microbiología, ADN recombinante, e inmunología bien conocida por aquellos expertos en la técnica. Tales técnicas se describen y se explican a lo largo de la literatura en fuentes tales como, J. Perbal, A Practical Guide to Molecular Cloning, John Wiley and Sons (1984), J. Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbour Laboratory Press (1989), T.A. Brown (editor), Essential Molecular Biology: A Practical Approach, Volumes 1 y 2, IRL Press (1991), D.M. Glover y B.D. Hames (editors), DNA Cloning: A Practical Approach, Volumes 1-4, IRL Press (1995 y 1996), y F.M. Ausubel et al. (Editors), Current Protocols in Molecular Biology, Greene Pub. Associates y Wiley- Interscience (1988, que incluyen todas las actualizaciones hasta ahora presentes).

#### Portadores Farmacéuticamente Aceptables

Los anticuerpos, polipéptidos y ADN de la presente invención se pueden incluir en composiciones que incluyen un portador o diluyente. Estas composiciones incluyen composiciones farmacéuticas en donde el portador o el diluyente serán farmacéuticamente aceptables. Los portadores o diluyentes farmacéuticamente aceptables incluyen aquellos utilizados en las composiciones adecuadas para administración oral, rectal, nasal, tópica (que incluye bucal y sublingual), vaginal, parenteral (que incluye subcutánea, intramuscular, intravenosa, intradérmica, intratecal y epidural). Estos no son tóxicos para los receptores en las dosificaciones y concentraciones empleadas. Los ejemplos representativos de portadores o diluyentes farmacéuticamente aceptables incluyen, pero no se limitan a; agua, soluciones isotónicas que se regulan preferiblemente en un pH fisiológico (tal como solución salina regulada con fosfato o solución salina regulada por Tris) y también pueden contener uno o más de, manitol, lactosa, trehalosa, dextrosa, glicerol, etanol o polipéptidos (tal como albúmina de suero humana). Las composiciones se pueden presentar convenientemente en forma de dosificación unitaria y se pueden preparar mediante cualquiera de los métodos bien conocidos en la técnica de farmacia.

Como se entenderá bien por aquellos expertos en la técnica se pueden hacer alteraciones a las secuencias de aminoácidos establecidas en los Listados de Secuencias. Estas alteraciones pueden ser eliminaciones, inserciones, o sustituciones de los residuos de aminoácidos. Los polipéptidos alterados pueden ocurrir en forma natural (es decir, purificar o aislar de una fuente natural) o sintética (por ejemplo, al realizar mutagenia dirigida a sitio en el ADN codificante). Se pretende que tales polipéptidos alterados que tiene por lo menos 85%, preferiblemente por lo menos 95 % de identidad con las secuencias establecidas en el Listado de Secuencias que están dentro de alcance de la presente invención. Los anticuerpos que surgen contra estos polipéptidos alterados también se unirán a los polipéptidos que tienen una de las secuencias establecidas en los Listados de Secuencias. El nivel de % de identidad se va a calcular según lo establecido anteriormente.

Las secuencias de proteínas son homólogas si se relacionan mediante divergencia de un ancestro común. Posteriormente, un homólogo de especie de la proteína será la proteína equivalente en la que ocurre en forma natural en otras especies. Dentro de cualquiera de una especie puede existir un homólogo como numerosas variantes alélicas, y éstas se considerarán homólogos de la proteína. Las variantes alélicas y homólogos de especies se pueden obtener mediante las siguientes técnicas estándar conocidas por aquellos expertos en la técnica.

Una variante alélica será una variante que ocurre en forma natural dentro de un organismo individual.

Mutantes, Variantes y Homología - Ácidos Nucleicos

Los polinucleótidos mutantes poseerán una o más mutaciones que son eliminaciones, inserciones, o sustituciones de los residuos de nucleótido. Los mutantes pueden ocurrir de forma natural (es decir, aislar de una fuente natural) o sintética (por ejemplo, al realizar mutagenia dirigida a sitio en el ADN). Es así evidente que los polinucleótidos de la invención pueden ocurrir en forma natural o recombinante (es decir preparar utilizando técnicas de ADN recombinante).

Una variante alélica será una variante que ocurre en forma natural dentro de un organismo individual.

Las secuencias de nucleótidos son homólogas si se relacionan mediante divergencia de un ancestro común. Posteriormente, un homólogo de especie del polinucleótido será el polinucleótido equivalente que ocurre en forma natural en otra especie. Dentro de cualquier una especie puede existir un homólogo como numerosas variantes alélicas, y estas se considerarán homólogos del polinucleótido. Las variantes alélicas y los homólogos de las especies se pueden obtener mediante las siguientes técnicas estándar conocidas por aquellos expertos en la técnica.

Producción de Anticuerpo

Los anticuerpos, policlonal o monoclonal, que son específicos para un polipéptido de la presente invención se pueden producir por una persona experta en la técnica utilizando técnicas estándar tales como, pero no limitado a, aquellos descritos por Harlow et al. *Antibodies: A Laboratory Manual*, Cold Springs Harbor Laboratory Press (1988), y D. Catty (editor), *Antibodies: A Practical Approach*, IRL Press (1988).

Se pueden utilizar diversos procedimientos conocidos en la técnica para la producción de anticuerpos policlonales a epítopos de una proteína. Para la producción de los anticuerpos policlonales, es aceptable un número de animales anfitriones para la generación de anticuerpos mediante inmunización con una o más inyecciones de preparación de un polipéptido, que incluye pero no se limita a conejos, ratones, ratas, etc. Se pueden utilizar diversos adyuvantes para aumentar la respuesta inmunológica en el animal anfitrión, dependiendo de las especies anfitriones, que incluyen pero no se limitan a Freund (completo y incompleto), geles minerales tales como hidróxido de aluminio, sustancias de superficie activa tales como lisolecitina, polioles plurónicos, polianiones, emulsiones de aceite, hemocianinas ojo de lapa de cerradura, dinitrofenol, y adyuvantes humanos potencialmente útiles tales como BCG (bacille Calmette-Guerin) y *Corynebacterium parvum*.

Un anticuerpo monoclonal para un epítipo de una proteína se puede preparar al utilizar cualquier técnica que proporciona la producción de las moléculas de anticuerpo mediante estirpes celulares continuas en el cultivo. Estas incluyen pero no se limitan a la técnica de hibridoma originalmente descrita por Kohler y Milstein (1975, *Nature* 256, 493-497), y la técnica de hibridoma de célula B humana más reciente (Kesber et al. 1983, *Immunology Today* 4:72) y la técnica de hibridoma EBV (Cole et al. 1985, *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc. pp. 77-96). Adicionalmente, se pueden utilizar técnicas desarrolladas para la producción de "anticuerpos quiméricos" al cortar y empalmar los genes de la molécula de anticuerpo de especificidad de antígeno apropiado junto con genes de una molécula de anticuerpo humano de actividad biológica apropiada (Morrison et al. 1984, *Proc. Natl. Acad. Sci.*, 81:6851-6855; Neuberger et al. 1984 *Nature* 312:604-608; Takeda et al. 1985 *Nature* 31: 452-454). Alternativamente, las técnicas descritas para la producción de anticuerpos de cadena sencilla (Patente Estadounidense 4,946,778) se pueden adaptar para producir anticuerpos de cadena sencilla específica 4.

Las versiones humanas o humanizadas recombinantes de los anticuerpos monoclonales son una realización preferida para aplicaciones terapéuticas humanas. Se puede preparar anticuerpos humanizados de acuerdo con procedimientos en la literatura (por ejemplo Jones et al. 1986, *Nature* 321:522-25; Reichman et al. 1988 *Nature* 332:323-27; Verhoeyen et al. 1988, *Science* 239:1534-36). La estrategia "mutagenia de conversión de gen" recientemente descrita para la producción del anticuerpo monoclonal humanizado también se puede emplear en la producción de anticuerpos humanizados (Carter et al. 1992 *Proc. Natl. Acad. Sci. U.S.A.* 89: 4285-89). Alternativamente, las técnicas para generar la colección de fase recombinante de combinaciones aleatorias de las regiones pesada y ligera que se pueden utilizar para preparar anticuerpos recombinantes (por ejemplo Huse et al. 1989 *Science* 246:1275-81).

Los fragmentos de anticuerpo que contienen el idiotipo de la molécula tal como  $Fu F(ab1)$  y  $F(ab2)$  se pueden generar mediante técnicas conocidas. Por ejemplo, tales fragmentos incluyen pero no se limitan a: el fragmento  $F(ab) E2$  que se puede producir mediante digestión de pepsina de la molécula de anticuerpo intacto; los fragmentos  $Fab'$  que se pueden generar al reducir los puentes de disulfuro del fragmento  $F(ab')_2$ , y los dos fragmentos  $Fab$  que se pueden generar al tratar la molécula de anticuerpo compuesta con papaína y un agente de reducción. Alternativamente, se pueden construir colecciones de expresión  $Fab$  (Huse et al. 1989, *Science* 246:1275-1281) para permitir la identificación rápida y fácil del fragmento monoclonal  $Fab$  con la especificidad deseada para una proteína.

## Adyuvantes

"Adyuvante" significa una composición comprendida de una o más sustancias que mejora la inmunogenicidad o eficacia de una composición de vacuna. Ejemplos no limitantes de los adyuvantes adecuados incluyen escualano y escualeno (u otros aceites de origen animal); copolímeros de bloque; detergentes tales como Tween®-80; Quail® A, aceites minerales tales como Drakeol® o Marcol®, aceites vegetales tales como aceite de maní; adyuvantes derivados de *Corynebacterium* tales como *Corynebacterium parvum*; adyuvantes derivados de *Propionibacterium* tales como *Propionibacterium acne*; *Mycobacterium bovis* (*Bacillus Calmetic y Guerin* o BCG); interleuquinas tales como interleuquina 2 y interleuquina-12; monoquinas tales como interleuquina 1; factor de necrosis de tumor; interferones tales como interferón gama; combinaciones tales como hidróxido de saponina-aluminio o hidróxido de aluminio Quil-A; liposomas; adyuvante ISCOM; extracto de pared celular micobacteriana; glicopéptidos sintéticos tales como dipéptidos muramilo u otros derivados; Avridina; Lípido A; sulfato de dextrano; DEAE-Dextrano o DHAE-Dextrano con fosfato de aluminio; carboxipolimetileno tal como Carbopol' EMA®; emulsiones de copolímero acrílico tales como Neocryl A640 (por ejemplo Patente Estadounidense No. 5,047,238); proteínas poxvirus de animal o vaccinia; adyuvantes de partícula sub-vírica tal como toxina del cólera, o mezclas de los mismos.

Como se utiliza aquí, las condiciones exigentes son aquellas que (1) emplean resistencia iónica baja y alta temperatura para lavado, por ejemplo, 0.015 M NaCl/0.0015 M citrato de sodio/0.1 % de NaDodSO<sub>4</sub> a 50° C; (2) emplean durante hibridación un agente desnaturizante tal como formamida, por ejemplo, 50 % (vol/vol) de formamida con 0.1 % de albúmina de suero bovino, 0.1 % de Ficoll, 0.1 % de povidona, regulador de fosfato de sodio 50 mM a pH 6.5 con 750 mM NaCl, citrato de sodio 75 mM a 42° C; o (3) emplean 50 % de formamida, 5 x SSC (0.75 M NaCl, citrato de sodio 0.075 M), fosfato de sodio 50 mM (pH 6.8), 0.1 % de pirofosfato de sodio, 5 x solución de Denhardt, ADN sonificado de esperma de salmón (50 µg/ml), 0.1 % de SDS y 10 % de sulfato de dextrano a 42° C en 0.2 x SSC y 0.1 % de SDS

Como se entenderá la presente invención incluye dentro de su alcance la vacuna de ADN. Se puede encontrar ADN con respecto a información adicional en Donnelly et al, *Journal of Immunological Methods* 176(1994) 145-152.

A través de esta especificación la palabra "comprende", o variaciones tales como "comprenden" o "que comprende", se entenderá que implica la inclusión de un elemento indicado o entero o grupo de elementos o enteros pero no la exclusión de cualquier otro elemento o entero, o grupo de elementos o enteros.

Preparación de la colección *P. gingivalis* para secuenciamiento.

Para determinar la secuencia de ADN de *P. gingivalis* el ADN genómico se aísla de la cepa *P. gingivalis* W50 (ATCC 53978) esencialmente por el método descrito por Mamur J. (*J. Mol. Biol.* 3, 208-218, 1961). Se realiza la clonación de los fragmentos de ADN esencialmente como se describe por Fleischmann et al., (*Science*; 269, 496-512, 1995) (2). En resumen, el ADN genómico de *P. gingivalis* se nebuliza para el fragmento de ADN y se trata con nucleasa Bal31 para crear extremos romo luego de correr dos veces a través de 1 % de geles de agarosa preparativos. Los fragmentos de ADN de 1.6-2.0 kb se extrajeron del gel y el ADN recuperado. Este ADN luego se liga a un vector pUC18 (Smal digerido y desfosforilado; Pharmacia) y se realiza electroforesis a través de 1 % de gel de agarosa preparativo. El fragmento que comprende el vector lineal más un inserto se extrae, se purifica y este proceso se repite para reducir cualquier vector sin contaminación del inserto. El vector recuperado más el ADN de inserto tiene extremo romo con polimerasa de ADN T4, luego se realiza un ligado final para producir ADN circular. Se transforman alícuotas de Células Competentes para Electroporación *Epicurian Coli* (Stratagene) con el ADN ligado y se ponen en placas en placas de difusión antibiótica de agar SOB que contienen X-gal y se incuban a 37° C durante la noche. Las colonias con insertos aparecen blancas y aquellas sin insertos (solo vector) aparecen azules. Las placas se almacenan a 4° C hasta que se recogieron los clones blancos y se expandieron para la extracción del plásmido de ADN para secuenciamiento.

## Secuenciamiento de ADN

Se prepara el plásmido de ADN al recoger colonias bacterianas en 1.5 ml de caldo de cultivo LB, TB o SOB complementado con 50-100 µg/ml de Ampicilina en placas de 96 pozos profundos. El plásmido de ADN se aísla utilizando los paquetes QIAprep Spin o QIAprep 96 Turbo miniprep (QIAGEN GmbH, Alemania). El ADN se eluye en una disposición de rejilla de 96 pozos a -20C.

Se realizan reacciones de secuenciamiento utilizando los equipos de Reacción Listos para Secuenciamiento de Ciclo Terminador de Tinte ABI PRISM y Tinte ABI PRISM BIG con polimerasa FS de ADN AmpliTaq (PE Applied Biosystems, Foster City, CA) utilizando los cebadores de secuenciamiento traseros y delanteros M13 Universal. Se conducen reacciones de secuencia en cualquier ciclador térmico Perkin-Elmer GeneAmp 9700 (PE Applied Biosystems) o Hybaid PCR Express (Hybaid, UK). Se analizan las reacciones de secuenciamiento en secuenciadores de ADN ABI PRISM 377 (PE Applied Biosystems).

Las secuencias obtenidas se establecen adelante. Se nota que la presente invención se relaciona con PG21. La relación entre estas secuencias se establece en la Tabla 1. Se calcula el codón de inicio utilizando una combinación

de la alineación de homología de secuencia (FASTA), predicción de la secuencia de señal (PSORT, SignalP) o predicción ORF (GeneMark).

Tabla 1: Tabla de referencia que indica las relaciones de cada secuencia ID a las proteínas seleccionadas.

Nombre de proteína	Secuencia de ADN de ORF completo	Secuencia de aminoácidos de ORF completo	Secuencia de ADN de proteína	Secuencia de aminoácidos de proteína
PG1	1	265	122	386
PG21	37	301	162	1426
PG3	45	309	170	434
PG30	46	310	171	435
PG96	118	382	261	525
PG97	119	383	262	526

5 **Análisis de secuencia de ADN**

Los archivos de ADN en formato FASTA se convierten a archivos de formato GCG y se importan en una base de datos. Los archivos de ADN se traducen en archivos de aminoácido utilizando el programa Flip obtenido de ANGIS (Australian Genomic Information Service, University of Sydney, Australia). Se realiza una serie de análisis bioinformáticos en las proteínas con el fin de seleccionar candidatos de vacuna potenciales. Los programas utilizados son búsqueda de homología FASTA (1), PSORT (2,3), SignalP (4), TopPred (5), y GeneMark (6). Las proteínas y sus resultados bioinformáticos se almacenan en la base de datos escrita habitual y se recupera de las proteínas con las características deseadas.

La homología FASTA resulta para estas proteínas que se examinan para cualquier alineación con una proteína que sugiere la ubicación de superficie o eficacia de vacuna. Todas las proteínas se buscan para homología contra una base de datos de proteína bacteriana no redundante compilada mediante ANGIS utilizando el algoritmo FASTA. Las configuraciones utilizadas para las búsquedas FASTA son Ktup = 2, penalidad de creación de espacio = -12, penalidad de extensión de espacio = -2, ancho para derivar la alineación en opt = 16 y la matriz de clasificación Blosum 50. Se examinan resultados de búsqueda FASTA individuales para homología significativa mediante probabilidad estadística y alineaciones de aminoácido. Los resultados se establecen en la Tabla 2.

Los archivos de proteína luego se ajustan al primer, segundo, tercero, cuarto y quinto residuos de metionina utilizando un ajuste de proteína, programa (ANGIS). Las proteínas ajustadas luego se someten a análisis PSORT para la detección de las secuencias de señal y la predicción de la ubicación celular. Las proteínas que exhiben una probabilidad PSORT de la membrana externa >0.8 se considera que indica la ubicación de la superficie. Un segundo programa de detección de la secuencia de señal SignalP también se realiza y, en ciertos casos, este programa detecta señales no identificadas con PSORT. Todas las proteínas identificadas por otros métodos también se analizan por PSORT y SignalP. Previamente, el aminoácido de terminal C de las proteínas de membrana externa bacteriana se ha mostrado que es importante para el ensamble de la proteína en la membrana externa (7). Una definición de estructura típica para las proteínas de membrana externa se ha determinado como la presencia de una secuencia de señal al terminal N y una tirosina o fenilalamina en el terminal C. Un número de las proteínas seleccionadas exhibe esta estructura característica. El programa TopPred se utiliza para determinar la presencia y número de dominios periféricos de membrana (MSD) y la presencia de tales secuencias indica una preferencia que se va a unir a las membranas tal como la membrana externa. Los resultados de los análisis PSORT, SignalP y TopPred con los aminoácidos de terminal C de las proteínas seleccionados se establecen en la Tabla 3.

Los 70 aminoácidos del terminal C de un número de proteínas de membrana externas *P. gingivalis* comparten 50-100 % de identidad de secuencia de proteína. Estas proteínas incluyen RGP1, RGP2, KGP, HagA, HagC, HagD, prtH y prtT. Este motivo conservado puede estar implicado en la unión o clasificación de la proteína a la membrana externa. El conjunto de datos de proteína se busca utilizando la homología FASTA como se describió anteriormente y se identifica un número de proteínas novedosas que demuestran motivos similares en su terminal C. Los resultados se enumeran en la Tabla 4.

40

ES 2 377 751 T3

Tabla 2: Homología de proteína FASTA resulta de los ORF completos contra una base de datos de proteína no redundante.

Nombre de la proteína	Descripción de homología	Número de acceso Genbank	longitud del homólogo	Longitud de la proteína P. gingivalis	Resultados de homología FASTA		
					% de identidad	Sobreposición	Valor E
PG1	Proteína de membrana externa 48kD, Actinobacillus pleuropneumoniae	U24492	449aa	451aa	32	454aa	1.40E-42
PG3	Adhesina porina de membrana externa, Pseudomonas fluorescens	U19743	317aa	223aa	35	187aa	1.10E-10
PG21	Gen de antígeno de superficie, Methanosarcina mazei	X84710	783aa	821aa	37	331aa	6.20E-33
PG30	Lipoproteína putativa NlpD, Aquifex aeolicus	AE000754	187aa	337aa	42	142aa	1.80E-12
PG75	Porina de membrana externa clase 3 (porB), Neisseria meningitidis	U07191	332aa	391aa	23	239aa	4.60E-01

Tabla 3: Resultados de los análisis PSORT, SignalP y TopPred de las proteínas. La columna presente de señal indica la presencia de una secuencia de señal detectada con PSORT o SignalP. Los términos en paréntesis indican el tipo de secuencia de señal como se determina por PSORT. La ubicación de la célula & los valores de probabilidad se generan mediante PSORT y representan la probabilidad de la proteína que está en los compartimientos celulares de la membrana externa (OM), la membrana interna (IM), el espacio periplásmico (PC) o el citoplasma (C). El número de dominios de transmembrana (TMD) se determina por TopPred y no incluye las secuencias de señal no divisibles.

Nombre de la proteína	Número de SeqID de proteína	Longitud de la proteína	Señal presente	Metionina en ORF	Sitio de división de señal IP	Sitio de división PSORT	Ubicación celular & probabilidad				Aminoácido de terminal C	Número de los TMD
							OM	IM	PS	C		
PG1	386	451aa	Y	1	14	34	0	0	0	0.22	N	0
PG3	434	223aa	Y (lipoproteína)	1	-	18	0.79	0.76	0	0	K	3
PG21	426	821aa	Y	2	24	27	0.34	0	0.37	0	G	1
PG30	435	337aa	Y	1	21	21	0.24	0	0.4	0	K	0
PG75	493	391aa	Y	1	26	26	0.94	0	0.3	0	H	1
PG96	525	563aa	Y	1	23	23	0.40	0	0.33	0	K	0
PG97	526	437aa	Y	1	23	23	0.32	0	0.65	0	Q	0

## ES 2 377 751 T3

Tabla 4: Porcentaje de identidad y porcentaje de similitud de diversas proteínas con los 70 aminoácidos del terminal C de la proteasa 1 de arginina *P. gingivalis* 1 (RGP1), proteasa 2 de arginina (RGP2), y la proteasa cisteína/hemaglutinina (prtT).

Nombre de la proteína	Porcentaje de identidad			Porcentaje de similitud		
	RGP1	RGP2	prtT	RGP1	RGP2	prtT
PG21	17	29	21	40	57	49
PG96	0	13	20	0	24	43
PG97	10	26	33	14	47	61

### 5 Clonación, expresión y purificación de los genes recombinantes *P. gingivalis*.

#### PG1

Los oligonucleótidos en las regiones 5' y 3' de la proteína deducida se utilizan para amplificar el gen de interés de una preparación del ADN genómico *P. gingivalis* W50 utilizando el Sistema PCR de Precisión TaqPlus (Stratagene) y un ciclador térmico PTC- 100 (MJ Research) o dispositivo similar. La secuencia de cebador 5' de oligonucleótido es GCGCCATATGCTGGCCGAACCGGCC, la secuencia de cebador 3' de oligonucleótido es GCGCCTCGAGTCAATTCATTTCTTATAGAG. Se purifica el fragmento PCR, digerido con las enzimas de restricción Nde I, Xho I (Promega) y se liga en los sitios correspondiente del plásmido pProEx-1 (Gibco-BRL) y se transforma en células *E. coli* ER1793 (un regalo de Elizabeth Raleigh, New England Biolabs). Se selecciona un clon resultante que expresa el inserto correcto y se induce con o sin 0.1mM IPTG (Promega) para la expresión de la proteína recombinante. Se determina la expresión de la proteína recombinante mediante análisis SDS-PAGE y Western Blot utilizando un antisuero de conejo descrito anteriormente o un anticuerpo anti-hexahistidina (Clontech) que detecta la etiqueta hexahistidina que se fusiona a la proteína recombinante *P. gingivalis*. Se purifica PG1 mediante la interrupción de las células *E. coli* mediante sonicación en regulador de unión (Novagen) y solubilización mediante la adición de sarcosil (N-Lauroil sarcosina) a una concentración final de 1%. Después la preparación se diluye a 0.1 % de sarcosil en regulador de unión, una a una columna de ácido Níquel-nitilotriacético (Ni-NTA; Qiagen), después de lavado las proteínas unidas se eluyen con 1M imidazol en regulador de elución (Novagen) de acuerdo con las recomendaciones Qiagen con 0.1 % de sarcosil agregado a todos los reguladores. Después de purificación las muestras se dializan contra 500mM NaCl, 20mM Tris, 0.1 % de sarcosyl a pH 7.4 para retirar el imidazol, concentrado cuando se requiera y se almacena a 4° C hasta que se utilizan. Se evalúan la pureza y antigenicidad mediante SDS-PAGE y Western blot utilizando el antisuero seleccionado (de aquellos descritos anteriormente) y la concentración de proteína se determina por el ensayo BCA (Pierce).

#### PG3

Los métodos utilizados para PG3 son esencialmente iguales para PG1 con las siguientes excepciones. La secuencia de cebador 5' de oligonucleótido es GCGCGTATACATGAAGAAATCAAGTGTAG, la secuencia de cebador 3' de oligonucleótido es GCGCAGATCTCTTCAGCGTACCTTGCTGTG y el ADN se amplifica con polimerasa de ADN Pfu (Stratagene). Se clona el producto PCR directamente en pCR-Blunt y se transforma en *E. coli* Top10F' (In Vitrogen) antes de subclonar en el plásmido de expresión pGex-parada RBS(IV) utilizando los sitios de restricción Bst Z171 y Bgl II y se transforma en *E. coli* BL21DE3 (Pharmacia Biotech). Se hacen las siguientes modificaciones para la purificación de PG3 del método PG1. Las células que expresan la proteína recombinante se interrumpen mediante sonicación en regulador de unión y los cuerpos de inclusión insolubles se concentran mediante centrifugación. Luego se solubilizan los cuerpos de inclusión en 6M urea (Sigma) en el regulador de unión y se eluyen con 6M urea agregado al regulador de elución. En algunos casos se utiliza clorhidrato de guanidina 6M (Sigma) en lugar de urea para estas etapas. Se retira la urea (o clorhidrato de guanidina cuando se sustituye) de la proteína purificada mediante diálisis secuencial contra niveles reducidos de urea (3M luego 1.5M luego 0.5M luego 0M urea toda en 50mM Tris, 500mM NaCl, 8 % de glicerol, pH 7.4). La proteína purificada se almacena congelada a -80° C hasta que se requiera. Se determina la concentración de la proteína mediante ensayo de proteína Coomassie Plus (Pierce).

#### PG30

Los métodos utilizados para PG30 son esencialmente iguales como para PG3 con las siguientes excepciones. Se retira la secuencia de señal de termina N predicha de la proteína recombinante. La secuencia de cebador 5' de oligonucleótido es TACGGAATTCGTGACCCCGTCAGAAATGTGCGC, la secuencia de cebador 3' de oligonucleótido es CTATGCGGCCGCTTTGATCCTCAAGGCTTTGCCCGG y se amplifica el ADN con el equipo PCR Tth XL. El producto PCR se clona en el plásmido de expresión pET24a utilizando los sitios de restricción Eco RI y Not I y se transforma en *E. coli* BL21DE3. Se llevan a cabo estudios de expresión y estudios de inmunoreactividad en lisados completos *E. coli* de PG30. Se cultivan cultivos de 10 ml de *E. coli* recombinante en un OD de 2.0 (A600 nm en caldo de cultivo Terrific y las células se inducen con 0.5 mM IPTG y se toman muestras para análisis a 4 horas post inducción. No se hace purificación para estos estudios.

#### PG75

Los métodos utilizados para PG75 son esencialmente iguales como para PG30 con las siguientes excepciones. La secuencia de señal de terminal N predicha se retira de la proteína recombinante. La secuencia de cebador 5' de oligonucleótido es GGCGGGATCCGCTCAGGAGCAACTGAATGTGGTA, la secuencia de cebador 3' de oligonucleótido es GAGTGC GGCCGCTGTGGAACAAATTGCGCAATCCATC y se amplifica el ADN con el equipo de PCR Tth XL. El producto PCR se clona en el plásmido de expresión pET24a utilizando los sitios de restricción Bam HI y Not I y se transforma en *E. coli* BL21DE3. Se llevan a cabo estudios de inmunoreactividad y estudios de expresión en lisados *E. coli* completos. No se hace purificación para estos estudios.

PG96

Los métodos utilizados para PG96 son esencialmente iguales como para PG30 con las siguientes excepciones. La secuencia de señal de terminal N predicha se retira de la proteína recombinante. La secuencia de cebador 5' de oligonucleótido es TGCTGAGCTCCAAACGCAATGCAAGCAGACCGA, la secuencia de cebador 3' de oligonucleótido es GAGTGC GGCCGCTTTTGGAGAATTTTCATTGTCTCACG y se amplifica el ADN con el equipo de PCR Tth XL. El producto PCR se clona en el plásmido de expresión pET24a utilizando los sitios de restricción Sac I y Not I y se transforma en *E. coli* BL21DE3. Se llevan a cabo estudios de inmunoreactividad y estudios de expresión en lisados *E. coli* completos. No se hace purificación para estos estudios.

PG97

Los métodos utilizados para PG97 son esencialmente iguales como para PG30 con las siguientes excepciones. La secuencia de señal de terminal N predicha se retira de la proteína recombinante. La secuencia de cebador 5' de oligonucleótido es GGCGGGATCCAGTTTGTTCGGCTCCACCACA, la secuencia de cebador 3' de oligonucleótido es GAGTGC GGCCGCTCTGTTTGATGAGCTTAGTGGTATA y se amplifica el ADN con el equipo de PCR Tth XL. El producto PCR se clona en el plásmido de expresión pET24a utilizando los sitios de restricción Bam HI y Not I y se transforma en *E. coli* BL21DE3. Se llevan a cabo estudios de inmunoreactividad y estudios de expresión en lisados *E. coli* completos. No se hace purificación para estos estudios.

Antisuero de animal y suero de paciente humano.

Se elevan diversos antisueros para detectar la expresión y el replegado de las proteínas *P. gingivalis* recombinantes. Se eleva un antisuero de célula completa al inyectar conejos Blandos de Nueva Zelanda con 3 dosis de *P. gingivalis* sonicados (cepa W50) que contiene aproximadamente 2 mg de proteína. La primera dosis se da en adyuvante completo de Freund's (FCA) y la segunda y tercera dosis se dan en adyuvante incompleto de Freund's (IFA) en intervalos de 3 semanas. Se suministran dosis intramusculares (1 ml) en las patas traseras y se desangran los conejos 7 días después de la última dosis, la sangre coagulada y el suero se retiran y se almacenan a -20° C hasta que se requiera. Se produce un segundo antisuero de conejo en una forma similar pero utilizando una fracción sarcosil insoluble (cada dosis tiene 0.69 mg de proteína) derivada de *P. gingivalis* W50 de acuerdo con el método de Doidg y Trust T. et al 1994 como el inmunógeno. Se produce un tercer antisuero de conejo en una forma similar para solo la primera fracción sarcosil soluble (1 mg de proteína por dosis) derivada de células *P. gingivalis* W50 de acuerdo con el método de Doidg P. y Trust T.J. (1994 Infect Immun 62:4526-33) se utiliza como el inmunógeno.

También se utiliza un grupo de "suero de rata protegido" en estos estudios y se obtiene de ratas inmunizadas con células *P. gingivalis* completas muertas con formalina en FIA (cepa ATCC 33277; 2 dosis de  $2 \times 10^9$  células, 3 semanas aparte). Las ratas luego se exponen 2 semanas después de su última dosis con células *P. gingivalis* (cepa 33277) dada oralmente como se describió previamente (Klaussen B. et al. 1991, Oral Microbiol Immunol 6:193-201) y el suero obtenido de estas ratas 6 semanas después de la inoculación de exposición final al momento del sacrificio.

Se obtienen suero humano de pacientes adultos que experimenta tratamiento o evaluación para periodontitis en una clínica de pacientes ambulatorios. Estos pacientes tienen por lo menos 6 dientes con pérdida de adhesión de 6 mm y tienen *P. gingivalis* presente en su placa sub-gingival cuando se detecta utilizando una sonda de ADN específica *P. gingivalis*. Se agrupa el suero de estos pacientes y se compara con un grupo de suero de pacientes periodontalmente saludables.

Inmunización y Protocolos del Modelo de Lesión de Murino

Se utiliza el modelo de absceso de ratón mouse para evaluar la eficacia de ratones inmunizados con proteínas recombinantes *P. gingivalis* en ratones protegidos de la formación de un absceso subcutáneo. Este modelo se ha utilizado por otros como un predictor de vacunas potenciales contra enfermedad periodontal (Bird PS, et al. 1995 J. Periodontol. 66:351-362). Se inmunizan ratones BALB/c de 6-8 semanas de edad al inyectarlos subcutáneamente con 0.1 ml que contiene 10 o 20 µg de proteína recombinante *P. gingivalis*, 20 µg de la proteína de lisado *E. coli*,  $2 \times 10^9$  células muertas con formalina de la cepa *P. gingivalis* 33277 emulsificada en adyuvante de Freund incompleto (IFA; Sigma) en el día 0. En el día 21 se inyectan ratones con la misma dosis y luego se desangran 1 semana después y se evalúan para niveles de anticuerpo. Al día 35 todos los ratones se exponen con aproximadamente  $2 \times 10^9$  células de *P. gingivalis* vivo (ATCC 33277) mediante inyección subcutánea en el abdomen. Luego de exposición los ratones se monitorean diariamente para la pérdida de peso y el tamaño de la lesión se mide durante los siguientes 10 días. Se miden los tamaños de la lesión mediante longitud y ancho y se expresa como mm<sup>2</sup>. Los

grupos se analizan estadísticamente utilizando un ANOVA de una vía Kruskal-Wallis y también se examinan individualmente utilizando la prueba t no pareada o la prueba de suma de clasificación Mann-Whitney utilizando el paquete estadístico Instat.

5 La Figura 1 muestra los resultados de un experimento en el día 4 después de exposición (las lesiones tienen un tamaño máximo al mismo punto). Los ratones de control inmunizados con el lisado E. coli muestran lesiones grandes mientras que los ratones inmunizados con células muertas de la cepa P. gingivalis 33277 se protegen completamente. Esto indica que no se protegen las células completas que proporcionan protección contra la proteína E.coli blanca P. gingivalis de ratones inmunizados. A los ratones que se les da diversas proteínas recombinantes PG muestran niveles significativos de protección para PG2, PG22, PG24 y PG29 ( $p < 0.05$  prueba t no  
10 pareada) mientras que el PG8A no es muy significativamente diferente ( $p = 0.07$ ) comparado con el grupo de control E. coli.

La Figura 2 muestra los resultados de un experimento separado utilizando combinaciones de proteínas recombinantes. A los ratones que se les da PG1 + PG2 muestran un nivel significativo de protección comparado con ratones de control que da el lisado E. coli ( $p < 0.026$  prueba t no pareada).

15 Inmunodetección

Se cultivan candidatos clonados en 15 ml de caldo de cultivo Terrific, inducido con IPTG y se toman muestras 4h post-inducción. Se retira un ml de cultivo, se peletiza y las células se resuspenden en un volumen de PBS determinado al dividir el OD  $A_{600nm}$  del cultivo mediante 8. Se agrega una alícuota de lisado (100  $\mu$ l) a 100  $\mu$ l de regulador de reducción de la muestra 2x (125 mM Tris pH 6.8, 20 % de glicerol, 4 % de SDS, 80mM DTT, 0.03 % de azul bromofenol) y se hierve durante 10 min. Se realiza SDS-PAGE de acuerdo con el método de Laemmli UK. 1970  
20 (Nature 227:680-685) utilizando 4-20 % de 1.0mm geles Tris-Glicina (Novex) de acuerdo con las recomendaciones de los fabricantes. Las proteínas se transfieren en membranas de nitrocelulosa Hybond-C Extra (Amersham) mediante "transblotting" y las membranas luego se bloquean durante 2h a temperatura ambiente (TA) en 5 % de leche descremada en 20mM Tris, 0.5M NaCl, 0.05 % de Tween-20, pH 7.5 (TTBS).

25 Se realiza inmunodetección separadamente con el suero de célula completa anti- P. gingivalis de conejo, el suero protector de rata, un grupo de pacientes periodontales humanos, y en muchos casos un conjugado HRP de anticuerpo anti-T7-Tag (Novagen). Antes de uso, se diluyen el suero de conejo, rata y humano 1/5000, 1/1000 y 1/500 respectivamente en 5 % de leche descremada en TTBS y se absorbe con 100  $\mu$ l (para el suero de conejo) o 250  $\mu$ l (para el suero de rata y humano) de extracto E. coli (20 mg/ml; Promega) durante 6h a TA.

30 Se incuban durante la noche membranas a TA con el antisuero absorbido, o durante 1 hr a YA con 1/5000 de conjugado anti-T7-Tag diluido. Luego de lavados de 3x10 min con TTBS, anticuerpo anti-conejo (Silenus), anti-ratón (Silenus) o anti-humano (KPL) conjugado con HRP, se diluye 1/5000 en 5 % de leche descremada en TTBS, se agrega durante 1h a TA. Las membranas se lavan como anteriormente, antes de la adición del sustrato de peroxidasa de membrana TMB (KPL) para la detección de las proteínas inmunoreactivas. Los resultados de  
35 reactividad de las proteínas P. gingivalis recombinantes se muestra en la Tabla 5.

Tabla 5: Resultados de inmunotransferencia de las proteínas expresadas en E.coli contra suero de conejo, rata y humano. Se calcula el MW deducido de la secuencia de aminoácidos de las proteínas P. gingivalis, algunas de las cuales tienen sus secuencias de señal de terminal N retiradas. Se determina el MW evidente de geles SDS-PAGE. Las etiquetas de terminal N- y C agregan aproximadamente 2.5 KDa al PM deducido de las proteínas recombinantes. Los símbolos son + positivo, - negativo, +/- positivo débil, ND no hecho.  
40

Número de la proteína	PM deducido (KDa)	PM evidente (KDa)	Reactividad del antisuero			
			T7	Conejo	Rata	Humano
PG1	47.5	63	ND	-	-	-
PG3	22.6	18.3	ND	- <sup>a</sup>	-	-
PG30	35.1	46.9	+	-	-	-
PG75	40.7	46.7	+	-	-	-
PG96	59.3	70.3	+	+	+	+
PG97	44.4	57.5	+	-	+	+

a. Reacción positiva detectada con el antisuero de conejo al antígeno sarcosil insoluble P. gingivalis.  
b. La proteína purificada demuestra reacción positiva débil con el antisuero de conejo a P. gingivalis completo.

(2) INFORMACIÓN PARA LA SEQ ID NO: 1

(i) SECUENCIA: CARACTERÍSTICAS:

(A) LONGITUD: 1362 pares base

5 (B) TIPO: ácido nucleico

(C) CADENA: doble

(D) TOPOLOGÍA: circular

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(iii) HIPOTÉTICO: NO

10 (iv) ANTI-CODIFICANTE: NO

(vi) FUENTE ORIGINAL:

(A) ORGANISMO: PORYPHYROMOHA GINGIVALIS

(ix) CARACTERÍSTICA:

(A) NOMBRE/CLAVE: misc\_feature

15 (B) UBICACIÓN 1...1362

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 1

TTCTGTGTCA	TGGCAAAAGT	TATAAAAACA	AAAAAAGGCC	TTGCACTTAA	TCTGAAAGGA	60
AAACCGCTSC	CCGAGATGCT	GGCCGAACCG	GCCCAAAGTC	CTACTTACGC	GSTCGTGCCC	120
GACGATTTTG	AAGGTGTTAT	CCCCAAGGTG	ACGGCTCGTC	CGGGGGATAA	GSTGCGTGCC	180
GGCTCAGCAC	TGATGCACCA	CAAGGCATAT	CCGGAGATGA	AGTTTACAAG	TCCGGTTAGC	240
GGCGAAGTGA	TCGCGGTGAA	TCGCGGTGCC	AAGCGCAAGG	TGTTGAGCAT	CGAGGTGAAA	300
CCGGACGGAC	TGAACGAATA	CGAGTCATTG	CCTGTCGGGG	ATCCGTCTGC	CCTCTCTGCC	360
GAACAGATCA	AGGAGCTTTT	ACTGTCGAGC	GGTATGTGGG	GTTTTATTAA	GCAACGTCTT	420
TACGACATAG	TGGCTACACC	GGATATAGCT	CCACGCGACA	TTTATATTAC	TGCCAACTTT	480
ACTGCACCAT	TGGCTCCGGA	CTTCGATTTC	ATCGTTCGAG	GAGAAGAACG	CGCCCTGCAG	540
ACTGCCATCG	ATGCCCTGGC	CAAACCTCAG	ACAGGAAAGG	TGTATGTGGG	CCTGAAGCCG	600
GGTTCATCTC	TGGGCTTGCA	CAATGCAGAA	ATCGTAGAAG	TACACGGACC	TCATCCGGCA	660
GGTAACCTGG	GCGTGTGAT	CAATCATACG	AAGCCAATCA	ATCGGGGCGA	AAUGGTGTGG	720
ACGCTCAAGG	CTACCGACCT	GATCGTGATC	GGACGTTTCC	TGCTTACGGG	CAAAGCCGAT	780
TTTACCAGAA	TGATTGCCAT	GACCGGCTCA	GACGCTGCAG	CTCAGGATA	CGTCCGTATT	840
ATGCCGGSTT	GCAATGTCTT	TGCTTCCTTC	CCCGGCCGAC	TGACAATAAA	GGAATCTCAC	900
GAGCGTGTGA	TCGATGGCAA	TGTGCTGACC	GSTAAGAAGC	TCTGCGAGAA	GGAGCCTTTC	960
CTGTFCAGCCC	GGTGTGACCA	GATCACGGTG	ATCCCCGAAG	GCGACGATGT	GGACGAACTC	1020
TTGCGGTGGG	CTGCACCCCG	TCTCGATCAG	TACAGCATGA	GCAGAGCTTA	TTTCTCTTGG	1080
TTGCAGGGGA	AAAACAAGA	GTACGTACTC	GATGCCCGGA	TCAAGGGTGG	CGAACGTGCT	1140
ATGATCATGA	GCAACGAGTA	TGACCGCGTT	TTCCCGATGG	ACATCTATCC	GGAGTATTTG	1200
CTCAAGGCTA	TTATAGCATT	CGACATCGAC	AAGATGGAGG	ACTTAGGCAT	ATATGAAGTG	1260
GCTCCGGAGG	ACTTTGCCAC	TTGCGAATTT	GTGGATACAT	CCAAGATCGA	GCTGCAGCGT	1320
ATCGTTCGCG	AGGGCTTGA	TATGCTCTAT	AAGGAAATGA	AT		1362

(2) INFORMACIÓN PARA LA SEQ ID NO: 37

(i) CARACTERÍSTICAS DE LA SECUENCIA:

20 (A) LONGITUD: 2607 pares base

(B) TIPO: ácido nucleico

(C) CADENA: doble

(D) TOPOLOGÍA: circular

(ii) TIPO DE MOLÉCULA: ADN (genómico)

25 (iii) HIPOTÉTICO: NO

(iv) ANTI-CODIFICANTE: NO

(vi) FUENTE ORIGINAL:

(A) ORGANISMO: PORTPHYROMONAS GINGIVALIS

(ix) CARACTERÍSTICA:

(A) NOMBRE/CLAVE: misc\_feature

5 (B) UBICACIÓN 1...2607

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 37

TGGCATAGGA	ATATTTTAT	CTTTGCGAGT	ACATTTAGCC	CGAAAAATAT	GCTCCCACTG	60
CCATACCGTT	ATGCAAAAAC	CGAGCACCTT	TTTCTCGCAA	AAGGATACTG	CAAGAATCCA	120
ATAACAAACA	TAATTTACCT	ATTTATGAAG	AAAAAGAATT	TTTTGCTTCT	TGGCATTTTC	180
GTTGCTTTGC	TGACTTTTAT	CGGCAGCATG	CAGGCACAAAC	AGGCCAAAAGA	TTATTTCAAC	240
TTTGACGAAC	GGGGCGAGGC	CTACTTCTCA	TTCAAAGTGC	CTGATAGGGC	CGTTCTACAA	300
GAGCTGGCTC	TGATCATGTC	CATCGACGAG	TTTGACCCCG	TAACCAATGA	AGCCATTGCC	360
TATGCCAGCG	AAGAGGAGTT	CGAGGCATTC	CTGCGCTATG	GGCTCAAGCC	TACATTCTTG	420
ACTCCTCCAT	CCATGCAGCG	CGCTGTGAG	ATGTTCCGACT	ACCGCTCAGG	AGAAAAATAC	480
GAATGGAAFG	CTTACCCAC	CTATGAAGCC	TATATCAGCA	TGATGGAAGA	GTTCCAAACA	540
AAGTATCCAT	CATTTTGTAC	TACTTCCGTC	ATTGSCAAGT	CCGTAAAAGG	TGTTAAACTG	600
ATGATTTGCA	AGCTGACGTC	CTCTGCCAAT	ACAGGGAAAA	AGCCTCGGCT	GCTCTATACT	660
TCTACGATGC	ACGAGAGACG	AACGACCGGA	TATGTGGTAC	TGCTCCGACT	CAIAGACCAT	720
CTGCTGTGCG	ACTACGAATC	CGATCCGAGG	ATTAAGAACA	TTCTGGATAA	AACGGAAGTA	780
TGGATCTGCC	CTTTGACCAA	TCCGGACGGA	GCATACAGAG	CCGGAAAACA	CACCGTACAA	840
GGAGCTACTC	GCTACAATGC	CAACAATGTC	GATTTGAACC	GTAACCTCAA	GGATGATGTA	900
GCCGGTGATC	ACCCCGATGG	AAAACCTTGG	CAGCCGGAGG	CAACTGCATT	CATGGATTTG	960
GAAGGAAACA	CCTCTTTCGT	GCTCGGTGCC	AATATACATG	GAGGAACAGA	GGTGGTGAAC	1020
TATCCATGGG	ATAATAAAAA	AGAAAGACAT	GCAGACGATG	AGTGGTACAA	ACTGATCAGT	1080
CGCAACTACG	CAGCCGCTTG	TCAGAGTATT	TCCGCCAGCT	ACATGACCTC	CGAAACCAAT	1140
TCGGGAATCA	TCAACGGTTC	AGACTGGTAT	GTAATTCGCG	GAAGTCGTCA	GGACAATGCA	1200
AATTATTTCC	ATCGTCTGCG	AGAAATTACC	CTTGAATCA	GCAACACGAA	GTGGTGCCG	1260
GCCTCTCAAC	TTCCAAAGTA	TTGGAATCTG	AACAAAGAAT	CTCTGCTTGC	TCTGATCGAA	1320
TAGTCCATTAT	ACGGCATCCA	TGGTACAGTG	ACTTCCGCTG	CGAACGGACA	GCCTCTCAAA	1380
TGCCAGATCT	TGATAGAAAA	CCATGACAAG	CGCAACTCCG	ATGTTTACTC	CGATGCTACC	1440
ACAGGCTACT	ACGTACGTC	TATCAAAAGCC	GGCACTTATA	CGGTGAAATA	CAAAGCCGAG	1500
GGTTATCCTG	AGGCAACTCG	TACCATTACG	ATCAAGGACA	AAGAAACCGT	CATCATGGAC	1560
ATTGCATTGG	GCAACTCGGT	TCCTCTGCCT	GTACCCGATT	TCACAGCTTC	TCCTATGACC	1620
ATCTCAGTAG	GCGAAAAGCGT	CCAATTCCAA	GATCAAAACGA	CAAATAACCC	CACGAATTGG	1680
GAGTGGACGT	TCGAAAGCGG	ACAGCCTGCC	ATGAGTACAG	AGCAGAATCC	GCTCGTATCC	1740
TATAGTCATC	CCGGTCAGTA	CGACGTTACG	CTCAAAGTGT	GGAAATGCAAG	TGGTTCCAAC	1800
ACGATTAACA	AAGAAAAATT	CATCACTGTC	AATGCCGTTA	TGCTGTAGC	TGAATTCGTC	1860
GGTACCCCGA	CGGAAATAGA	AGAGGGCCAG	ACGGTATCTT	TCCAAAACCA	ATCCACCAAT	1920
GCCACCAACT	ACGTATGGAT	ATTCGATGGC	GGCACTCCCG	CTACCAAGTA	AGACGAAAAC	1980
CCGACTGTGC	TTTACAGCAA	AGCCGGCCAA	TACGATGTCA	CGCTCAAGGC	GATCASTGCT	2040
TCCGGTGAAA	CGGTGAAGAC	GAAAGAAAAA	TACATCACTG	TCAAGAAAGC	TCCGGTCCCT	2100
GCTCCGGTAG	CCGACTTCGA	AGGAACACCT	CGAAAAGTAA	AGAAAGGCGA	GACAGTTACT	2160
TTCAAAGACT	TGTCTAGCAA	CAATCCGACT	TCATGGCTTT	GGGTGTTTCA	AGGCGGCTCT	2220
CCTGCCACCA	GCACGGAGCA	AAACCCGGTG	GTCACTTACA	ATGAAACAGG	CAAGTACGAT	2280
GTCCAGCTGA	CTGCCACCAA	CGAGGGCCGA	AGCAATGTGA	AGAAAGCAGA	AGACTACATT	2340
GAGGTTATCC	TCGATGACAG	TGTCGAGGAC	ATAGTGGCAC	AGACGGGTAT	CGTCATTGGT	2400
CCGCAAAACG	GAACGAAGCA	GATCCTCATA	GAAGCCAAAG	CTGCTATCAA	AGCGATCGTT	2460
CTCTATGACA	TCAATGGAGC	GGTCGTACTC	AAACTACTC	CGAATCAGCT	CCGCTCGACC	2520
GTAGATCTTT	CCATCCTGCC	CGAAGGAATC	TACACCATCA	ATATCAAAAC	GGAAAAATCC	2580
GCTCGCACGG	AAAAGATCCA	TATCGGG				2607

(2) INFORMACIÓN PARA LA SEQ ID NO: 45

(i) CARACTERÍSTICAS DE SECUENCIA:

10 (A) LONGITUD: 690 pares base

(B) TIPO: ácido nucleico

(C) CADENA: doble

(D) TOPOLOGÍA: circular

(ii) TIPO DE MOLÉCULA: ADN (genómico)

15 (iii) HIPOTÉTICO: NO

(iv) ANTI-CODIFICANTE: NO

(vi) FUENTE ORIGINAL:

(A) ORGANISMO: PORYPHYROMONAS GINGIVALIS

(ix) CARACTERÍSTICA:

(A) NOMBRE/CLAVE: misc\_feature

5 (B) UBICACIÓN 1...690

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 45

```

ACGAATAAAA AAGAAGAGAC AATGAAGAAA TCAAGTGTAG TAGCCTCAST TTTGGCCGTG      60
GCCTCGTGT TCGCCGTTG CGGACTGAAC AATATGGCAA AAGGCGCCT TATCGGCGCC      120
GGAGTAGGAG GTGCCATTGG TGCCGGAGTA GGTAACGTAG CCGGAAATAC GGCTGTCGGT      180
GCCATCGTCG GTACTGCAGT CGGTGGAGCA GCCGGTGCTC TCATCGGAAA GAAGATGGAC      240
AAGCAGAAAA AAGAAGTGGG GGCCGCAGTA CCCGATGCTA CGATTCAGAC AGTAAATGAC      300
GGAGAGGCTA TTCTGTTTAC TTTCGATAGC GGTATCCTCT TTGCGACGAA CTCCAGCACT      360
CTGAGTCCCA ACTCACGCAC TCGCGTGACG AAGTTTGCTG CAAACATGAA CAAAAACCCC      420
GACACGGATA TTCGTATCGT AGGCCATACG GACAATACCG GCTCCGACAA GATCAACGAT      480
CCTCTGTCTG AGAGACGTGC AGCCAGCGTA TATTCTTTC TGAATTCTCA GGGTGTGAGT      540
ATGTCGCGCA TGGCAGCCGA AGGGCGTGGG AGCCATGAAC CGGTTGCAGA CAATAGCACA      600
GTTGCCGGAC GTTCGGCCAA CCGCCGTGTG GAGGTTTATA TCTTGCCGAA TGCCAAGATG      660
ATCGAACAAG CACAGCAAGG TACGCTGAAG
    
```

(2) INFORMACIÓN PARA LA SEQ ID NO: 46

(i) CARACTERÍSTICAS DE SECUENCIA:

10 (A) LONGITUD: 1026 pares base

(B) TIPO: ácido nucleico

(C) CADENA: doble

(D) TOPOLOGÍA: circular

(ii) TIPO DE MOLÉCULA: ADN (genómico)

15 (iii) HIPOTÉTICO: NO

(iv) ANTI-CODIFICANTE: NO

(vi) FUENTE ORIGINAL:

(A) ORGANISMO: PORYPHYROMONAS GINGIVALIS

(ix) CARACTERÍSTICA:

20 (A) NOMBRE/CLAVE: misc\_feature

(B) UBICACIÓN 1...1026

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 46

```

AACAGGAACA GAAATATGTC GAAAAAATCG ATCCTTCTGC TTTGCTGTTT GCTGTGCTTC      60
ATTTCTGCTA CGAAGGCTGT GACCCCGTTC AGAAATGTGC GCATATAGCCA AGTGAACAGC      120
AAAGCAAAGA CCGAACGTAC AAAGCCCTCG GACTCTGTAC GGTACATTAG CAACATGATT      180
GCAGATCGGC TGGAGTTCCG CAACAAGATT TCTTCCGAAA AAGAGGTAAG AAAAGCCGAA      240
TATGAAAATC GGCTGGCSAT GGAAGCACTC AATTACCCTG CCATAGATTT ATATGGTGAA      300
GATTCTTGGA GCGAGTATGT AAACCCTTTC GTGGGTGCAG GAACCGATGT CGAAATTCGG      360
AACTCCTATG ACATTGATTG CTCTTCGTTT GTGATGCCCG TCGAAGATAA GCAGGTCACC      420
TCTCAATTTG GCTACCGTGG GCGTTTCGGA CGGATGCACT ATGGTATTGA TCTTTCAGTG      480
AATCGTGGOS ATACGATACG AGCAGCCTTT GACGGGAAAG TTCGTGTACG CAGCTATGAA      540
GCGCGTGGCT ATGGCTACTA CATAGTCTTG CGCCATCCGA ACGGACTGGA GACTGTGTAC      600
GGACACATGA GTCGCCAATT GGTAGACGAG AATCAGATCG TTCCGAGCAGG ACAACCGATC      660
GGATTAGGAG GCAGCAGCGG TCGAAGCACC GGTCTCTATC TTCACTTCGA GACCCGCTTC      720
ATGGGTATTC CCATCAATCC GAGTACCATT ATAGACTTCG ATAACGGAGT GCCGCTCCGA      780
GACATTTACA CATTCAAACG AGGGAGCAAT TCTCGCTATG CAAAAGCCTC TAAGACTTCT      840
TCTCGCTATG CAAAAAAGG GAAGAAAGGC AGACAAAGCT CTTCTCTAT GACCTATAGA      900
ATCAAAAAG GCGATACTTT GGAACAATA GCCAAAAGGC ACGGCACCTC TGTTAGAAA      960
CTCTGTGCTA CCAATGGCAT TGGCAAGAGT AAAATTTTGA CTCGGGCAA AGCCTTGAGG      1020
ATCAAA
    
```

ES 2 377 751 T3

(2) INFORMACIÓN PARA LA SEQ ID NO: 118

(i) CARACTERÍSTICAS DE SECUENCIA:

(A) LONGITUD: 1689 pares base

(B) TIPO: ácido nucleico

5 (C) CADENA: doble

(D) TOPOLOGÍA: circular

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(iii) HIPOTÉTICO: NO

(iv) ANTI-CODIFICANTE: NO

10 (vi) FUENTE ORIGINAL:

(A) ORGANISMO: PORYPHYROMONAS GINGIVALIS

(ix) CARACTERÍSTICA:

(A) NOMBRE/CLAVE: misc\_feature

(B) UBICACIÓN 1...1689

15 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 118

```
CATCATAAAA CATATCAAAC AATGAAAAAG CTTTACAGG CTAAGCCTT GATTCTGGCA 60
TTGGGACTCT TCCAAGTCC CGCAATCGCC CAAACGCAA TGCAAGCAGA CCGAACAAAC 120
GGTCAATTG CAACAGAAGA GATGCAACGA GCATTCCAGG AAACGAATCC CCCTGCAGGT 180
CCTGTGCGTG CTATCGCTGA GTACGAACGC TCTGCAGCCG TTTTGGTACG CTACCCGTTT 240
GGTATCCCGA TGGAAATTGAT CAAAGAGCTG GCCAAGAAGC ACAAGGTGAT TACCATTGTG 300
GCGAGTGAAA GCCAAAAAAA CACCGTTATA ACCCAGTACA CCCAAAGCGG TGTGAATCTC 360
TCTAATTGCG ATTTTCATCAT TGGGAAAAGT GACTCTTACT GGACACGCGA CTATACCGGT 420
TGTTTCGCAA TGTACGATAC GAACAAAGTA GGTCTCGTGG ACTTTATTTA TAACCGCCCT 480
CGTCCTAACG ATGATGAATT CCCCAAATAC GAAGCACAAAT ATCTGGGCAT CGAGATGTTC 540
GGGATGAAAC TCAAGCAGAC CGGTGCAAC TACATGACGG ACGGATATGG ATCCGCTGTG 600
CAGTCACATA TCGCATATAC GGAGAAGTCC TCTCTGTCTC AAGCTCAAGT AAATCAAAG 660
```

```
ATGAAAGACT ATCTGGCAT CACACATCAT GATCTGGTAC AAGATCCGAA CGGCGAATAT 720
ATCAACCATG TGGACTGTTG GGGCAAGTAT TTGGCACCGA ACAAATCCT CATCAGGAAA 780
GTGCCTGACA ATCACCTCA GCACCAAGCC CTGGAAGATA TGGCAGCCTA CTCGCAGCA 840
CAGACCTGCG CATGGGGAAC GAAGTACGAG GTATATCGCG CTTTGGCCAC CAATGAACAA 900
CCGTACAGSA ACTCTCTGAT TCTGAACAAC AGGGTATTTG TTCCTGTCAA TGGCCCGGCC 960
TCCGTGGACA ACGATGCTCT GAACGTCTAT AAGACGGCAA TGCCCGGTTA CGAAATTTATA 1020
GGTGTCAAAG GGGCTTCAGG AACACCTTGG TTAGGAACAG ATGCCCTGCA TTGTCTACT 1080
CACGAGGTAG CCGATAAGGG CTATCTCTAT ATCAAGCACT ACCCGATACT GGGCGAACAG 1140
GCAGGCCCTG ATTATAAGAT GGAAGCAGAT GTCGTCTCAT GCGCCAATGC TACTATCTCG 1200
CCGGTACAAT GTTACTATCG TATCAATGGT TCCGGTAGCT TTAAGGCTGC TGATATGACG 1260
ATGGAATCAA CAGTCACTA TACTTATAGC TTTACAGSTC TTAACAAGAA TGATAAGGTA 1320
GAATACTATA TCTCTGCCGC TGACAATAGT GGTGCGCAAAG AGACTTATCC CTTTATCGGC 1380
GAACCTGATC CTTTCAAGTT TACGTGTATG AACGAAACCA ATACATGTAC TGTGACCGGA 1440
GCTCCCAAAG CTCTTCGTGC ATGGTTCAAC GCCGGTCTGT CAGAAGTGGC TGTTCGGTA 1500
AGTTTGAATA TTGCGGCAC ATATCGGATA AAGCTTTATA ACACCGCAGG AGAAGAAGTC 1560
GCTGCAATGA CCAAGGAATT AGTAGCAGGG ACGAGTGTCT TCAGTATGGA TGTGTATTCT 1620
CAGGCTCCGG GCACATATGT TCTGGTTGTT GAAGGAAATG GAATCCGTGA GACAATGAAA 1680
ATTCTCAA 1689
```

(2) INFORMACIÓN PARA LA SEQ ID NO: 119

(i) CARACTERÍSTICAS DE SECUENCIA:

(A) LONGITUD: 1311 pares base

20 (B) TIPO: ácido nucleico

(C) CADENA: doble

(D) TOPOLOGÍA: circular

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(iii) HIPOTÉTICO: NO

5 (iv) ANTI-CODIFICANTE: NO

(vi) FUENTE ORIGINAL:

(A) ORGANISMO: PORYPHYROHONAS GINGIVALIS

(ix) CARACTERÍSTICA:

(A) NOMBRE/CLAVE: misc\_feature

10 (B) UBICACIÓN 1...1311

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 119

```

ACCACAAATA  GAAAACCAAA  TACTAATATG  AACTTTTCAT  CTAAGAAAAT  CTTAGCAATC  60
ATTGCATTGC  TGACGATGGG  ACATGCTGTG  CAGGCACAGT  TTGTTCCGGC  TCCCACCACA  120
GGGATTCCGA  TGTCTGTAC  TACAACCAAG  GCCGTAGGCG  AAAAAATCGA  ATTGTTGGTT  180
CATTCCATAG  AGAAGAAAGG  CATCTGGATC  GATCTCAATG  GGGATGCCAC  TTACCAACAA  240
GGAGAGGAAA  TAACCGTATT  CGATGAGGCA  TACCACGAAT  ACACGATCGG  GACGCAAACC  300
CTCACTATCT  ATGGTAATAC  GACCCGATTG  GGCTGTGGAT  CTACCGGTGC  AACGGCTGTC  360
GATGTAACGA  AAAACCCFAA  TCTGACCTAT  CTCGCATGCC  CGAAAAATAA  TCTGAAATCA  420
TTGGACTTGA  CGCAAACCC  AAAGCTGCTG  CGAGTTTGGT  GCGACTCTAA  CGAAATAGAA  480
AGTTTGGACC  TGAGTGGCAA  TCCGGCTTTG  ATCATCCTCG  GCTGTGACAG  GAATAAGCTG  540
ACTGAGCTGA  AGACCGATAA  CAACCCCAAG  TTGGCCTCTC  TTTGGTGTTC  TGATAATAAC  600
CTGACGGAGT  TGGAACTCAG  TGCCAACTCT  CGTCTCAATG  ATCTTTGGTG  CTTCCGGTAA  660
CGGATCAGCA  AACTCGATCT  GAGTGCCAAT  CCTCTATTGG  TAACACTTTG  GTGCAGTGAC  720
AATGAGCTTT  CGACCTTGG  TCTTTCCAAG  AATTCGGACG  TTGCTTACCT  TTGGTGTTC  780
TCGAACAAC  TTACATCCTT  GAATCTGTG  GGGGTGAAGG  GACTGAGTGT  TTTGGTTTGT  840
CATTCCAATC  AGATCGCAGG  TGAAGAAATG  ACGAAAGTGG  TGAATGCTTT  GCCCACACTA  900
TCTCCCGCG  CAGGCGCTCA  GAGCAAGTTC  GTGGTTGTAG  ACCTCAAGGA  CACTGATGAG  960
AAGAATATCT  GTACCGTAAA  GGATGTGGAA  AAAGCTAAAA  GTAAGAACTG  GCGAGTATTT  1020
GACTTCAACG  GTGATTCTGA  CAATATGCTT  CCATACGAAG  GAAGTCCGAC  ATCGAACTTG  1080
GCAGTAGATG  CTCCCCTGT  CAGGATATAT  CCAATCCGG  TAGGAAGATA  TGCGCTCGTC  1140
GAGATCCCG  AGTCTCTTTT  AGGGCAGGAA  GCTGCTTTAT  ACGATATGAA  TGGGGTAAAA  1200
GTCTATAGT  TCGCGGTAGA  GTCTCTTCGT  CAGAACATTG  ACCTGACACA  TCTTCCCGAC  1260
GGCACTTAT  TCTTCCGTCT  CGATAACTAT  ACCACTAAGC  TCATCAAACA  G  1311
    
```

(2) INFORMACIÓN PARA LA SEQ ID NO: 122

(i) CARACTERÍSTICAS DE SECUENCIA:

15 (A) LONGITUD: 1353 pares base

(B) TIPO: ácido nucleico

(C) CADENA: doble

(D) TOPOLOGÍA: circular

(ii) TIPO DE MOLÉCULA: ADN (genómico)

20 (iii) HIPOTÉTICO: NO

(iv) ANTI-CODIFICANTE: NO

(vi) FUENTE ORIGINAL:

(A) ORGANISMO: PORYPHYROMONAS GINGIVALIS

(ix) CARACTERÍSTICA:

25 (A) NOMBRE/CLAVE: misc\_feature

(B) UBICACIÓN 1...1353

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 122

ATGGCAAAAG	TTATAAAAAC	AAAAAAAGGC	CTTGCACTTA	ATCTGAAAGG	AAAACCGCTG	60
CCCGAGATGC	TGGCCGAACC	GGCCCAAAGT	CCTACTTACG	CGGTGCTGCC	CGACGATTTT	120
GAAGGTGTTA	TCCCAAGGT	GACGGCTCGT	CCGGGGGATA	AGGTGCGTGC	CGGCTCAGCA	180
CTGATGCACC	ACAAGGCATA	TCCGGAGATG	AAGTTTACAA	GTCCGGTTAG	CGGCGAAGTG	240
ATCGCGGTGA	ATCGCGGTGC	CAAGCGCAAG	GTGTTGAGCA	TCGAGGTGAA	ACCGGACGGA	300
CTGAACGAAT	ACGAGTCATT	CCCTGTGCGG	GATCCGTCTG	CCCTCTCTGC	CGAACAGATC	360
AAGGAGCTTT	TACTGTGAG	CGGTATGTGG	GGTTTTATTA	AGCAACGTCC	TTACGACATA	420
GTGGCTACAC	CGGATATAGC	TCCACGCGAC	ATTTATATTA	CTGCCAACTT	TACTGCACCA	480
TTGGCTCCGG	ACTTCGATTT	CATCGTTCGA	GGAGAAGAAC	GCGCCCTGCA	GACTGCCATC	540
GATGCCTTGG	CCAAACTCAC	GACAGGAAAG	GTGTATGTGG	GCCTGAAGCC	GGGTTCATCT	600
CTGGGCTTGC	ACAATGCAGA	AATCGTAGAA	GTACACGAC	CTCATCCGGC	AGGTAACGTG	660
GGCGTGTGA	TCAATCATAC	GAAGCCAATC	AATCGGGGCG	AAACGGTGTG	GACGCTCAAG	720
GCTACCGACC	TGATCGTATC	CGGACGTTTC	CTGCTTACGG	GCAAAGCCGA	TTTTACCAGA	780
ATGATTGCCA	TGACCGGCTC	AGACGCTGCA	GCTCACGGAT	ACGTCCGTAT	TATGCCGGGT	840
TGCAATGTCT	TTGCTTCTTT	CCCCGGCCGA	CTGACAATAA	AGGAATCTCA	CGAGCGTGTG	900
ATCGATGGCA	ATGTGCTGAC	CGGTAAGAA	CTCTGCGAGA	AGGAGCCTTT	CCTGTCAGCC	960
CGGTGTGACC	AGATCACGGT	GATCCCCGAA	GGCGACGATG	TGGACGAACT	CTTCGGGTGG	1020
GCTGCACCCC	GTCTCGATCA	GTACAGCATG	AGCAGAGCTT	ATTTCTCTTG	GTTGCAGGGG	1080
AAAACAAG	AGTACGTACT	CGATGCCCGG	ATCAAGGGTG	GCGAACGTGC	TATGATCATG	1140
AGCAACGAGT	ATGACCGCGT	TTTCCCAGTG	GACATCTATC	CGGAGTATTT	GCTCAAGSCT	1200
ATTATAGCAT	TCGACATCGA	CAAGATGGAG	GACTTAGGCA	TATATGAAGT	GGCTCCGGAG	1260
GACTTTGCCA	CTTGCGAATT	TGTGGATACA	TCCAAGATCG	AGCTGCAGCG	TATCGTTCCG	1320
GAGGGCTTGG	ATATGCTCTA	TAAGGAAATG	AAT			1353

(2) INFORMACIÓN PARA LA SEQ ID NO: 162

(i) CARACTERÍSTICAS DE SECUENCIA:

5 (A) LONGITUD: 2463 pares base

(B) TIPO: ácido nucleico

(C) CADENA: doble

(D) TOPOLOGÍA: circular

(ii) TIPO DE MOLÉCULA: ADN (genómico)

10 (iii) HIPOTÉTICO: NO

(iv) ANTI-CODIFICANTE: NO

(vi) FUENTE ORIGINAL:

(A) ORGANISMO: PORYPHYROMONAS GUIGIVALIS

(ix) CARACTERÍSTICA:

15 (A) NOMBRE/CLAVE: misc\_feature

(B) UBICACIÓN 1...2463

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 162

ATGAAGAAAA	AGAATTTTTT	GCTTCTTGCC	ATTTTCGTTG	CTTGCTGAC	TTTCATCGGC	60
AGCATGCAGG	CACAACAGGC	CAAAGATTAT	TTCAACTTTG	ACGAACGGGG	CGAGGCCTAC	120
TTCTCATTCA	AAGTGCCTGA	TAGGGCCGTT	CTACAAGAGC	TGGCTCTGAT	CATGTCCATC	180
GACGAGTTTG	ACCCCGTAAC	CAATGAAGCC	ATTGCCTATG	CCAGCGAAGA	GGAGTTCGAG	240
GCATTCCTGC	GCTATGGGCT	CAAGCCTACA	TTCTTGACTC	CTCCATCCAT	GCAGCGCGCT	300
GTCGAGATGT	TCGACTACCG	CTCAGGAGAA	AAATACGAAT	GGAAATGCTTA	CCCCACCTAT	360
GAAGCCTATA	TCAGCATGAT	GGAAGAGTTC	CAAACAAGT	ATCCATCACT	TTGTACTACT	420
TCCGTCAATG	GCAAGTCCGT	AAAGGATCGT	AAACTGATGA	TTTGCAAGCT	GACGTCTCT	480
GCCAATACAG	GGAAAAAGCC	TCGCGTGCTC	TATACTTCTA	CGATGCACGG	AGACGAAAACG	540
ACCGGATATG	TGGTACTGCT	CCGACTCATA	GACCATCTGC	TGTCGAACTA	CGAATCCGAT	600
CCGAGGATTA	AGAACATTCT	GGATAAAAACG	GAAGTATGGA	TCTGCCCTTT	GACCAATCCG	660
GACGGAGCAT	ACAGAGCCGG	AAACCACACC	GTACAAGGAG	CTACTCGCTA	CAATGCCAAC	720
AATGTCGATT	TGAACCGTAA	CTTCAAGGAT	GATGTAGCCG	GTGATCACCC	CGATGGAAAA	780
CCTTGGCAGC	CGGAGSCAAC	TGCATTCATG	GATTTGGAAG	GAACACCTC	TTTGTGTGCTC	840
GGTGCCAAATA	TACATGGAGG	AACAGAGGTG	GTGAACATATC	CATGGGATAA	TAAAAAGAA	900
AGACATGCAG	ACGATGAGTG	GTACAAACTG	ATCAGTCGCA	ACTACGCAGC	CGCTTGTGAC	960
AGTATTTCCG	CCAGCTACAT	GACCTCCGAA	ACCAATTCGG	GAATCATCAA	CGGTTCAGAC	1020
TGCTATGTAA	TTCCGGGAAG	TCGTCAGGAC	AATGCAAATT	ATTTCCATCG	TCTGCGAGAA	1080
ATTACCCCTG	AAATCAGCAA	CACGAAGTTG	GTGCCGGCCT	CTCAACTTCC	AAAGTATTGG	1140
AATCTGAACA	AAGAACTCT	GCTTGCTCTG	ATCSAAGAA	CCTTATACGG	CATCCATGGT	1200
ACAGTGACTT	CCGCTGCGAA	CGGACAGCCT	CTCAAAATGCC	AGATCTTGAT	AGAAAACCAT	1260
GACAAGCGCA	ACTCCGATGT	TTACTCCGAT	GCTACCACAG	GCTACTACGT	ACGTCCATATC	1320
AAAGCCGGCA	CTTATACGGT	GAAATACAAA	GCCGAGGGT	ATCCTGAGGC	AACTCGTACC	1380
ATTACGATCA	AGGACAAAGA	AACCGTCATC	ATGGACATTG	CATTGGGCAA	CTCGTTCCCT	1440
CTGCCTGTAC	CCGATTTTAC	AGCTTCTCCT	ATGACCATCT	CAGTAGGGCA	AAGCGTCCAA	1500
TTCCAAGATC	AAAGCACAAA	TAACCCACAG	AATTGGGAGT	GGACGTTTGA	AGGCGGACAG	1560
CCTGCCATGA	GTACAGAGCA	GAATCCGCTC	GTATCCTATA	GTATCCCGG	TCAGTACGAC	1620
GTTACGCTCA	AAGTGTGGAA	TGCAAGTGGT	TCCAACACGA	TTACGAAAAG	AAAATTCATC	1680
ACTGTCAATG	CCGTTATGCC	TGTAGCTGAA	TTGCTCGGTA	CCCCGACGGA	AATAGAAGAG	1740
GGCCAGACGG	TATCTTTCCA	AAACCAATCC	ACCAATGCCA	CCAACTACGT	ATGGATATTC	1800
GATGSCGGCA	CTCCCGCTAC	CAGTGAAGAC	GAAAACCGCA	CTGTGCTTTA	CAGCAAAGCC	1860
GGCCAATACG	ATGTCACGCT	CAAGCGGATC	AGTGCTTCCG	GTGAAACGGT	GAAGACGAAA	1920

GAAAAATACA	TCACTGTCAA	GAAAGCTCCG	GTCCCTGCTC	CGGTAGCCGA	CTTCGAAGGA	1980
ACACCTCGAA	AAGTAAAGAA	AGGCGAGACA	GTACTTTTCA	AAGACTTGTC	TACGAACAAT	2040
CCGACTTCAT	GGCTTTGGGT	GTTCGAAGGC	GGCTCTCCTG	CCACCAGCAC	GGAGCAAAAC	2100
CCGGTGSTCA	CCTACAATGA	AACAGGCAAG	TACGATGTCC	AGCTGACTGC	CACCAACGAG	2160
GGCGGAAGCA	ATGTGAAGAA	AGCAGAAGAC	TACATTGAGG	TTATCCTCGA	TGACAGTGTC	2220
GAGGACATAG	TGGCACAGAC	GGGTATCGTC	ATTGCTCCGC	AAAACGGAAC	GAAGCAGATC	2280
CTCATAGAAG	CCAACGCTGC	TATCAAAGCG	ATCGTTCTCT	ATGACATCAA	TGGACGGGTC	2340
GTAICTAAAA	CTACTCCGAA	TCAGCTCCGC	TCGACCGTAG	ATCTTTCCAT	CCTGCCCGAA	2400
GGAACTACA	CCATCAATAT	CAAAACGGAA	AAATCCGCTC	GCACGGAAAA	GATCCATATC	2460
GGG						2463

(2) INFORMACIÓN PARA LA SEQ ID NO: 170

(i) CARACTERÍSTICAS DE SECUENCIA:

(A) LONGITUD: 669 pares base

5 (B) TIPO: ácido nucleico

(C) CADENA: doble

(D) TOPOLOGÍA: circular

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(iii) HIPOTÉTICO: NO

10 (iv) ANTI-CODIFICANTE: NO

(vi) FUENTE ORIGINAL:

(A) ORGANISMO: PORYPHYROMONAS GINGIVALIS

(ix) CARACTERÍSTICA:

(A) NOMBRE/CLAVE: misc\_feature

(B) UBICACIÓN 1...669

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 170

```

ATGAAGAAAT CAAGTGTAGT AGCCTCAGTT TTGGCCGTGG CTCTCGTGT CGCCGGTTGC      60
GGACTGAACA ATATGGCAA AGGCGGCTT ATCGGCGCG GAGTAGGAGG TGCCATTGGT      120
GCCGGASTAG GTAACGTAGC CGGAAATACG GCTGTGGTG CCATCGTCGG TACTGCAGTC      180
GGTGGAGCAG CCGGTGCTCT CATCGGAAAG AAGATGGACA AGCAGAAAA AGAACTGGAG      240
GCCGCAGTAC CCGATGCTAC GATTCAGACA GTAAATGACG GAGAGGCTAT TCTGGTACT      300
TTCGATAGCG GTATCCTCTT TGCGACGAAC TCCAGCACTC TGAGTCCCAA CTCACGCACT      360
GCGCTGACGA AGTTTCTGTC AAACATGAAC AAAAACCCCG ACACGGATAT TCGTATCGTA      420
GCCATACGG ACAATACCGG CTCCGACAAG ATCAACGATC CTCTGTCTGA GAGACGTGCA      480

```

```

GCCAGCGTAT ATTCTTCTCT GAATTCCTCAG GGTGTGAGTA TGTCGGCGAT GGCAGCCGAA      540
GGGCGTGGGA GCCATGAACC GGTGCGAGAC AATAGCACAG TTGCCGGACG TTCGGCCAAC      600
CGCCGTGTGG AGGTTTATAT CTTGCCGAAT GCCAAGATGA TCGAACAAAGC ACAGCAAGGT      660
ACGCTGAAG                                     669

```

(2) INFORMACIÓN PARA LA SEQ ID NO: 171

5 (i) CARACTERÍSTICAS DE SECUENCIA:

(A) LONGITUD: 1011 pares base

(B) TIPO: ácido nucleico

(C) CADENA: doble

(D) TOPOLOGÍA: circular

10 (ii) TIPO DE MOLÉCULA: ADN (genómico)

(iii) HIPOTÉTICO: NO

(iv) ANTICODIFICANTE: NO

(vi) FUENTE ORIGINAL:

(A) ORGANISMO: PORYPHYROMONAS GINGIVALIS

15 (ix) CARACTERÍSTICA:

(A) NOMBRE/CLAVE: misc\_feature

(B) UBICACIÓN 1...1011

(::i) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 171

```

ATGTCGAAAA AATCGATCCT TCTGCTTTCG TGTCGCTGT GCTTCATTC TGCTACGAAG      60
GCTGTGACCC CCGTCAGAAA TGTGCGCAAT AGCCAAGTGA ACAGCAAAGC AAAGACCGAA      120
CGTACAAAGC CCTCGGACTC TGTACGGTAC ATTAGCAACA TGATTGCAGA TCGGCTGGAG      180
TTCGGCAACA AGATTTCTTC CGAAAAAGAG GTAAGAAAAG CCGAATATGA AAATCGGCTG      240
GCGATGGAAG CACTCAATTA CCCTGCCATA GATTTATATG GTGAAGATTG TTGGAGCGAG      300
TATGTAACC CTTFCGTCCG TGCAGGAACC GATGTCGAAA TTCGGAACCT CTATGACATT      360
GATTGCTCTT CGTTGCTGAT GCCCGTCGAA GATAAGCAGG TCACCTCTCA ATTTGGCTAC      420
CGTCGCGCTT TCGGACGGAT GCACTATGGT ATTGATCTTT CAGTGAATCG TGGCGATACG      480
ATACGAGCAG CCTTTGACGG GAAAGTTCGT GTACGCAGCT ATGAAGCGCG TGGCTATGGC      540
TACTACATAG TCTTGGCCCA TCCGAACGGA CTGGAGACTC TGTACGGACA CATGAGTCGC      600
CAATTGGTAG ACGAGAATCA GATCGTTCGA GCAGGACAAC CGATCGGATT AGGAGGCAGC      660
ACGGGTGAAA GCACCGGTCC TCATCTTCAC TTCGAGACCC GCTTCATGGG TATTCCCATC      720
AATCCGAGTA CCATTATAGA CTTGCGATAAC GGAGTGCCGC TCCGAGACAT TTACACATTC      780
AAACGAGGGA GCAATTCTCG CTATGCAAAA GCCTCTAAGA CTTCTTCTCG CTATGCAAAA      840
AAAGGGAAGA AAGGCAGACA AGCTTCTTCT CCTATGACCT ATAGAATCAA AAAAGGCGAT      900
ACTTTGGAAA CAATAGCCAA AAGGCACGGC ACTTCTGTTC AGAAACTCTG TGCTACCAAT      960
GGCATTGCCA AGAGTAAAT TTTGACTCCG GGCAAAGCCT TGAGGATCAA A                               1011

```

20 (2) INFORMACIÓN PARA LA SEQ ID NO: 261

(i) CARACTERÍSTICAS DE SECUENCIA:

ES 2 377 751 T3

(A) LONGITUD: 1668 pares base

(B) TIPO: ácido nucleico

(C) CADENA: doble

(D) TOPOLOGÍA: circular

5 (ii) TIPO DE MOLÉCULA: ADN (genómico)

(iii) HIPOTÉTICO: NO

(iv) ANTI-CODIFICANTE: NO

(vi) FUENTE ORIGINAL:

(A) ORGANISMO: PORYPHYROHONAS GINGIVALIS

10 (ix) CARACTERÍSTICA:

(A) NOMBRE/CLAVE: misc\_feature

(B) UBICACIÓN 1...1668

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 261

```

ATGAAAAAGC TTTTACAGGC TAAAGCCTTG ATTCTGGCAT TGGGACTCTT CCAACTGCCC      60
GCAATCGCCC AAACGCAAAT GCAAGCAGAC CGAACAAACG GTC AATTTGC AACAGAAGAG      120
ATGCAACGAG CATTCAGGA AACGAATCCC CCTGCGAGTC CTGTGCGTGC TATCGCTGAG      180
TACGAACGCT CTSCAGCCGT TTTGGTACGC TACCGSTTCG GTATCCCGAT GGAATTGATC      240
AAAGAGCTGG CCAAGAACGA CAAGGTGATT ACCATTGTGG CGAGTGAAG CCAAAAAAAC      300
ACCGTTATAA CCCAGTACAC CCAAAGCGST GTGAATCTCT CTAATTGCGA TTTTCATCATT      360
GGGAAAAGTG ACTCTTACTG GACACGCGAC TATACCGSTT GGTTCGCAAT GTACGATACG      420
AACAAAGTAG GTCTCGTGGA CTTTATTTAT AACCGCCCTC GTCCTAACGA TGATGAATTC      480
CCCAAATACG AAGCACAATA TCTGGGCATC GAGATGTTCG GGATGAAGCT CAAGCAGACC      540
GGTGGCAACT ACATGACGGA CGGATATGGA TCCGCTGTGC AGTCACATAT CGCATATACG      600
GAGAACTCCT CTCTGTCTCA AGCTCAAGTA AATCAAAGA TGAAGACTA TCTCGGCATC      660
ACACATCATG ATGTGTACA AGATCCGAAC GCGGAATATA TCAACCATGT GGACTGTTGG      720
GGCAAGTATT TGGCACCAGAA CAAAATCCTC ATCAGGAAAG TGCCCTGACAA TCACCCCTCAG      780
CACCAAGCCC TGGAGATAT GGCAGCCTAC TTCGAGCAC AGACCTGCGC ATGGGGAAAG      840
AAGTACGAGG TATATCGCGC TTTGGCCACC AATGAACAAC CGTACACGAA CTCTCTGATT      900
CTGAACAACA GGGTATTTGT TCCTGTCAAT GGCCCGCCTT CCGTGGACAA CGATGCTCTG      960
AACGTCTATA AGACGGCAAT GCCCGGTTAC GAAATTATAG GTGTCAAAGG GGCTTCAGGA     1020
ACACCTTGGT TAGGAACAGA TGCCCTGCAT TGTCGTACTC ACGAGGTAGC GGATAAGGGC     1080
TATCTCTATA TCAAGCACTA CCCGATACTG GCGGAACAGG CAGGCCCTGA TTATAAGATC     1140
GAAGCAGATG TCGTCTCATG CGCCAATGCT ACTATCTCGC CGSTACAATG TTACTATCGT     1200
ATCAATGTTT CCGGTAGCTT TAAGGCTGCT GATATGACGA TGGAAATCAAC AGGTCACTAT     1260
ACTTATAGCT TTACAGGTCT TAACAAGAAT GATAAGGTAG AATACTATAT CTCTGCGGCT     1320
GACAAATAGT GTCGCAAAGA GACTTATCCC TTTATCGGCG AACCTGATCC TTTCAAGTTT     1380
ACGTGTATGA ACGAAACCAA TACATGTACT GTGACCGGAG CTGCCAAAGC TCTTGTGCA     1440
TGGTTCAACG CCGGTCTGTC AGAACTGGCT GTTTCGGTAA GTTTGAATAT TGCCGGCACA     1500
TATCGGATAA AGCTTTATAA CACCGCAGGA GAAGAAGTCG CTSCAATGAC CAAGGAATTA     1560
GTAGCAGGGA CGAGTGTCTT CAGTATGGAT GTGTATTCTC AGSCTCCGGG CACATATGTT     1620
CTGGTTGTTG AAGSAAATGG AATCCGTGAG ACAATGAAAA TTCTCAA      1668
    
```

15 (7) INFORMACIÓN PARA LA SEQ ID NO: 262

(i) CARACTERÍSTICAS DE SECUENCIA:

(A) LONGITUD: 1284 pares base

(B) TIPO: ácido nucleico

(C) CADENA: doble

20 (D) TOPOLOGÍA: circular

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(iii) HIPOTÉTICO: NO

(iv) ANTI-CODIFICANTE: NO

(vi) FUENTE ORIGINAL:

(A) ORGANISMO: PORYPHYROMONAS GINGIVALIS

(ix) CARACTERÍSTICA:

(A) NOMBRE/CLAVE: misc\_feature

(B) UBICACIÓN 1...1284

5 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 262

```

ATGAAACTTT CATCTAAGAA AATCTTAGCA ATCATTGCAT TGCTGACGAT GGGACATGCT      60
GTGCAGGCAC AGTTTGTTCG GGCTCCCACC ACAGGATTC GCATGTCTGT CACTACAACC      120
AAGGCCGTAG GCGAAAAAAT CGAATTGTTG GTTCATTCCA TAGAGAAGAA AGGCATCTGG      180
ATCGATCTCA ATGSGGATGC CACTTACCAA CAAGGAGAGG AAATAACCGT ATTCGATGAG      240
GCATACCACG AATACACGAT CGGGACGCAA ACCCTCACTA TCTATGGTAA TACGACCCGA      300
TTGGGCTGTC GATCTACCGG TGCAACGGCT GTCGATGTAA CGAAAAACCC TAATCTGACC      360
TATCTCGCAT GCCCGAAAAA TAATCTGAAA TCATTGGACT TGACGCAAAA CCCAAAGCTG      420
CTGCGAGTTT GGTGCGACTC TAACGAAATA GAAAGTTTGG ACCTGAGTGG CAATCCGGCT      480
TTSATCATCC TCGSCTGTGA CAGGAATAAG CTGACTGAGC TGAAGACCGA TAACAACCCC      540
AAGTTGGCCT CTCTTTGGTG TTCTGATAAT AACCTGACGG AGTTGGAAC TCAATCCCAAT      600
CCTCGTCTCA ATGATCTTTG GTGCTTCGGT AATCGATCA CGAAACTCGA TCTGAGTGCC      660
AATCCTCTAT TGGTAACACT TTGGTGCACT GACAATGAGC TTTGACCTT GGATCTTTCC      720
AAGAATTCGG ACGTTGCTTA CCTTTGGTGT TCATCGAACA AACTTACATC CTTGAATCTG      780
TCGGGGGTGA AGGGACTGAG TGTTTGGTGT TGTCATTCCA ATCAGATCGC AGGTGAAGAA      840
ATGACGAAAG TGGTGAATGC TTTGCCACA CTATCTCCCG GCGCAGGCGC TCAGAGCAAG      900
TTCGTGTTG TAGACCTCAA GGACACTGAT GAGAAGAATA TCTGTACCGT AAAGGATGTG      960
GAAAAAGCTA AAAGTAAGAA CTGGCGAGTA TTTGACTTCA ACGGTGATTC TGACAATATG     1020
CTTCCATACG AAGGAAGTCC GACATCGAAC TTGGCAGTAG ATGCTCCCAC TGTCAGGATA     1080
TATCCCAATC CCGTAGGAAG ATATGCGCTC GTCGAGATCC CCGAGTCTCT TTTAGGGCAG     1140
GAAGCTGCTT TATACGATAT GAATGGGGTA AAAGTCTATA GTTTGCGGCT AGAGTCTCTT     1200
CGTCAGAAAC TTGACCTGAC ACATCTTCCC GACGGCACTT ATTTCTTCCG TCTCGATAAC     1260
TATACCACTA AGCTCATCAA ACAG
    
```

(2) INFORMACIÓN PARA LA SEQ ID NO: 265

(i) CARACTERÍSTICAS DE SECUENCIA:

(A) LONGITUD: 454 aminoácidos

10 (B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: proteína

(iii) HIPOTÉTICO: SI

(vi) FUENTE ORIGINAL:

15 (A) ORGANISMO: Porfiromonas gingivalis

(ix) CARACTERÍSTICA:

(A) NOMBRE/CLAVE: misc\_feature

(B) UBICACIÓN 1...454

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 265

ES 2 377 751 T3

Phe Cys Val Met Ala Lys Val Ile Lys Thr Lys Lys Gly Leu Ala Leu  
 1 5 10 15  
 Asn Leu Lys Gly Lys Pro Leu Pro Glu Met Leu Ala Glu Pro Ala Gln  
 20 25 30  
 Ser Pro Thr Tyr Ala Val Val Pro Asp Asp Phe Glu Gly Val Ile Pro  
 35 40 45  
 Lys Val Thr Ala Arg Pro Gly Asp Lys Val Arg Ala Gly Ser Ala Leu  
 50 55 60  
 Met His His Lys Ala Tyr Pro Glu Met Lys Phe Thr Ser Pro Val Ser  
 65 70 75 80  
 Gly Glu Val Ile Ala Val Asn Arg Gly Ala Lys Arg Lys Val Leu Ser  
 85 90 95  
 Ile Glu Val Lys Pro Asp Gly Leu Asn Glu Tyr Glu Ser Phe Pro Val  
 100 105 110  
 Gly Asp Pro Ser Ala Leu Ser Ala Glu Gln Ile Lys Glu Leu Leu Leu  
 115 120 125  
 Ser Ser Gly Met Trp Gly Phe Ile Lys Gln Arg Pro Tyr Asp Ile Val  
 130 135 140  
 Ala Thr Pro Asp Ile Ala Pro Arg Asp Ile Tyr Ile Thr Ala Asn Phe  
 145 150 155 160  
 Thr Ala Pro Leu Ala Pro Asp Phe Asp Phe Ile Val Arg Gly Glu Glu  
 165 170 175  
 Arg Ala Leu Gln Thr Ala Ile Asp Ala Leu Ala Lys Leu Thr Thr Gly  
 180 185 190  
 Lys Val Tyr Val Gly Leu Lys Pro Gly Ser Ser Leu Gly Leu His Asn  
 195 200 205  
 Ala Glu Ile Val Glu Val His Gly Pro His Pro Ala Gly Asn Val Gly  
 210 215 220  
 Val Leu Ile Asn His Thr Lys Pro Ile Asn Arg Gly Glu Thr Val Trp  
 225 230 235 240  
 Thr Leu Lys Ala Thr Asp Leu Ile Val Ile Gly Arg Phe Leu Leu Thr  
 245 250 255  
 Gly Lys Ala Asp Phe Thr Arg Met Ile Ala Met Thr Gly Ser Asp Ala  
 260 265 270  
 Ala Ala His Gly Tyr Val Arg Ile Met Pro Gly Cys Asn Val Phe Ala  
 275 280 285  
 Ser Phe Pro Gly Arg Leu Thr Ile Lys Glu Ser His Glu Arg Val Ile  
 290 295 300  
 Asp Gly Asn Val Leu Thr Gly Lys Lys Leu Cys Glu Lys Glu Pro Phe  
 305 310 315 320  
 Leu Ser Ala Arg Cys Asp Gln Ile Thr Val Ile Pro Glu Gly Asp Asp  
 325 330 335  
 Val Asp Glu Leu Phe Gly Trp Ala Ala Pro Arg Leu Asp Gln Tyr Ser  
 340 345 350  
 Met Ser Arg Ala Tyr Phe Ser Trp Leu Gln Gly Lys Asn Lys Glu Tyr  
 355 360 365  
 Val Leu Asp Ala Arg Ile Lys Gly Gly Glu Arg Ala Met Ile Met Ser  
 370 375 380  
 Asn Glu Tyr Asp Arg Val Phe Pro Met Asp Ile Tyr Pro Glu Tyr Leu  
 385 390 395 400  
 Leu Lys Ala Ile Ile Ala Phe Asp Ile Asp Lys Met Glu Asp Leu Gly  
 405 410 415  
 Ile Tyr Glu Val Ala Pro Glu Asp Phe Ala Thr Cys Glu Phe Val Asp  
 420 425 430  
 Thr Ser Lys Ile Glu Leu Gln Arg Ile Val Arg Glu Gly Leu Asp Met  
 435 440 445  
 Leu Tyr Lys Glu Met Asn  
 450

(2) INFORMACIÓN PARA LA SEQ ID NO: 301

(i) CARACTERÍSTICAS DE SECUENCIA:

(A) LONGITUD: 869 aminoácidos

5 (B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: proteína

(iii) HIPOTÉTICO: SI

(vi) FUENTE ORIGINAL:

(A) ORGANISMO: Porfiromonas gingivalis

5 (ix) CARACTERÍSTICA:

(A) NOMBRE/CLAVE: misc\_feature

(B) UBICACIÓN 1...869

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 301

```

Trp His Arg Asn Ile Phe Ile Phe Ala Ser Thr Phe Ser Pro Lys Asn
1      5      10      15
Met Leu Pro Leu Pro Tyr Arg Tyr Ala Lys Thr Glu His Leu Phe Leu
20      25      30
Ala Lys Gly Tyr Cys Lys Asn Pro Ile Thr Asn Ile Ile Leu Phe
35      40      45
Met Lys Lys Lys Asn Phe Leu Leu Leu Gly Ile Phe Val Ala Leu Leu
50      55      60
Thr Phe Ile Gly Ser Met Gln Ala Gln Gln Ala Lys Asp Tyr Phe Asn
65      70      75      80
Phe Asp Glu Arg Gly Glu Ala Tyr Phe Ser Phe Lys Val Pro Asp Arg
85      90      95
Ala Val Leu Gln Glu Leu Ala Leu Ile Met Ser Ile Asp Glu Phe Asp
100      105      110
Pro Val Thr Asn Glu Ala Ile Ala Tyr Ala Ser Glu Glu Glu Phe Glu
115      120      125
Ala Phe Leu Arg Tyr Gly Leu Lys Pro Thr Phe Leu Thr Pro Pro Ser
130      135      140
Met Gln Arg Ala Val Glu Met Phe Asp Tyr Arg Ser Gly Glu Lys Tyr
145      150      155      160
Glu Trp Asn Ala Tyr Pro Thr Tyr Glu Ala Tyr Ile Ser Met Met Glu
165      170      175
Glu Phe Gln Thr Lys Tyr Pro Ser Leu Cys Thr Thr Ser Val Ile Gly
180      185      190
Lys Ser Val Lys Asp Arg Lys Leu Met Ile Cys Lys Leu Thr Ser Ser
195      200      205
Ala Asn Thr Gly Lys Lys Pro Arg Val Leu Tyr Thr Ser Thr Met His
210      215      220
Gly Asp Glu Thr Thr Gly Tyr Val Val Leu Leu Arg Leu Ile Asp His
225      230      235      240
Leu Leu Ser Asn Tyr Glu Ser Asp Pro Arg Ile Lys Asn Ile Leu Asp
245      250      255

```

Lys Thr Glu Val Trp Ile Cys Pro Leu Thr Asn Pro Asp Gly Ala Tyr  
 260 265 270  
 Arg Ala Gly Asn His Thr Val Gln Gly Ala Thr Arg Tyr Asn Ala Asn  
 275 280 285  
 Asn Val Asp Leu Asn Arg Asn Phe Lys Asp Asp Val Ala Gly Asp His  
 290 295 300  
 Pro Asp Gly Lys Pro Trp Gln Pro Glu Ala Thr Ala Phe Met Asp Leu  
 305 310 315 320  
 Glu Gly Asn Thr Ser Phe Val Leu Gly Ala Asn Ile His Gly Gly Thr  
 325 330 335  
 Glu Val Val Asn Tyr Pro Trp Asp Asn Lys Lys Glu Arg His Ala Asp  
 340 345 350  
 Asp Glu Trp Tyr Lys Leu Ile Ser Arg Asn Tyr Ala Ala Ala Cys Gln  
 355 360 365  
 Ser Ile Ser Ala Ser Tyr Met Thr Ser Glu Thr Asn Ser Gly Ile Ile  
 370 375 380  
 Asn Gly Ser Asp Trp Tyr Val Ile Arg Gly Ser Arg Gln Asp Asn Ala  
 385 390 395 400  
 Asn Tyr Phe His Arg Leu Arg Glu Ile Thr Leu Glu Ile Ser Asn Thr  
 405 410 415  
 Lys Leu Val Pro Ala Ser Gln Leu Pro Lys Tyr Trp Asn Leu Asn Lys  
 420 425 430  
 Glu Ser Leu Leu Ala Leu Ile Glu Glu Ser Leu Tyr Gly Ile His Gly  
 435 440 445  
 Thr Val Thr Ser Ala Ala Asn Gly Gln Pro Leu Lys Cys Gln Ile Leu  
 450 455 460  
 Ile Glu Asn His Asp Lys Arg Asn Ser Asp Val Tyr Ser Asp Ala Thr  
 465 470 475 480  
 Thr Gly Tyr Tyr Val Arg Pro Ile Lys Ala Gly Thr Tyr Thr Val Lys  
 485 490 495  
 Tyr Lys Ala Glu Gly Tyr Pro Glu Ala Thr Arg Thr Ile Thr Ile Lys  
 500 505 510  
 Asp Lys Glu Thr Val Ile Met Asp Ile Ala Leu Gly Asn Ser Val Pro  
 515 520 525  
 Leu Pro Val Pro Asp Phe Thr Ala Ser Pro Met Thr Ile Ser Val Gly  
 530 535 540  
 Glu Ser Val Gln Phe Gln Asp Gln Thr Thr Asn Asn Pro Thr Asn Trp  
 545 550 555 560  
 Glu Trp Thr Phe Glu Gly Gly Gln Pro Ala Met Ser Thr Glu Gln Asn  
 565 570 575  
 Pro Leu Val Ser Tyr Ser His Pro Gly Gln Tyr Asp Val Thr Leu Lys  
 580 585 590  
 Val Trp Asn Ala Ser Gly Ser Asn Thr Ile Thr Lys Glu Lys Phe Ile  
 595 600 605  
 Thr Val Asn Ala Val Met Pro Val Ala Glu Phe Val Gly Thr Pro Thr  
 610 615 620  
 Glu Ile Glu Glu Gly Gln Thr Val Ser Phe Gln Asn Gln Ser Thr Asn  
 625 630 635 640  
 Ala Thr Asn Tyr Val Trp Ile Phe Asp Gly Gly Thr Pro Ala Thr Ser  
 645 650 655  
 Glu Asp Glu Asn Pro Thr Val Leu Tyr Ser Lys Ala Gly Gln Tyr Asp  
 660 665 670  
 Val Thr Leu Lys Ala Ile Ser Ala Ser Gly Glu Thr Val Lys Thr Lys  
 675 680 685  
 Glu Lys Tyr Ile Thr Val Lys Lys Ala Pro Val Pro Ala Pro Val Ala  
 690 695 700  
 Asp Phe Glu Gly Thr Pro Arg Lys Val Lys Lys Gly Glu Thr Val Thr  
 705 710 715 720  
 Phe Lys Asp Leu Ser Thr Asn Asn Pro Thr Ser Trp Leu Trp Val Phe  
 725 730 735  
 Glu Gly Gly Ser Pro Ala Thr Ser Thr Glu Gln Asn Pro Val Val Thr  
 740 745 750  
 Tyr Asn Glu Thr Gly Lys Tyr Asp Val Gln Leu Thr Ala Thr Asn Glu  
 755 760 765  
 Gly Gly Ser Asn Val Lys Lys Ala Glu Asp Tyr Ile Glu Val Ile Leu  
 770 775 780  
 Asp Asp Ser Val Glu Asp Ile Val Ala Gln Thr Gly Ile Val Ile Arg  
 785 790 795 800  
 Pro Gln Asn Gly Thr Lys Gln Ile Leu Ile Glu Ala Asn Ala Ala Ile  
 805 810 815  
 Lys Ala Ile Val Leu Tyr Asp Ile Asn Gly Arg Val Val Leu Lys Thr  
 820 825 830  
 Thr Pro Asn Gln Leu Arg Ser Thr Val Asp Leu Ser Ile Leu Pro Glu  
 835 840 845



ES 2 377 751 T3

(A) LONGITUD: 342 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: proteína

5 (iii) HIPOTÉTICO: SI

(vi) FUENTE ORIGINAL:

(A) ORGANISMO: Porfiromonas gingivalis

(ix) CARACTERÍSTICA:

(A) NOMBRE/CLAVE: misc\_feature

10 (B) UBICACIÓN 1...342

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 310

```

Asn Arg Asn Arg Asn Met Ser Lys Lys Ser Ile Leu Leu Leu Cys Cys
1      5      10      15
Ser Leu Cys Phe Ile Ser Ala Thr Lys Ala Val Thr Pro Val Arg Asn
20      25      30
Val Arg Asn Ser Gln Val Asn Ser Lys Ala Lys Thr Glu Arg Thr Lys
35      40      45
Pro Ser Asp Ser Val Arg Tyr Ile Ser Asn Met Ile Ala Asp Arg Leu
50      55      60
Glu Phe Arg Asn Lys Ile Ser Ser Glu Lys Glu Val Arg Lys Ala Glu
65      70      75      80
Tyr Glu Asn Arg Leu Ala Met Glu Ala Leu Asn Tyr Pro Ala Ile Asp
85      90      95
Leu Tyr Gly Glu Asp Ser Trp Ser Glu Tyr Val Asn Pro Phe Val Gly
100     105     110
Ala Gly Thr Asp Val Glu Ile Pro Asn Ser Tyr Asp Ile Asp Cys Ser
115     120     125
Ser Phe Val Met Pro Val Glu Asp Lys Gln Val Thr Ser Gln Phe Gly
130     135     140
Tyr Arg Arg Arg Phe Gly Arg Met His Tyr Gly Ile Asp Leu Ser Val
145     150     155     160
Asn Arg Gly Asp Thr Ile Arg Ala Ala Phe Asp Gly Lys Val Arg Val
165     170     175
Arg Ser Tyr Glu Ala Arg Gly Tyr Gly Tyr Tyr Ile Val Leu Arg His
180     185     190
Pro Asn Gly Leu Glu Thr Val Tyr Gly His Met Ser Arg Gln Leu Val
195     200     205
Asp Glu Asn Gln Ile Val Arg Ala Gly Gln Pro Ile Gly Leu Gly Gly
210     215     220
Ser Thr Gly Arg Ser Thr Gly Pro His Leu His Phe Glu Thr Arg Phe
225     230     235     240
Met Gly Ile Pro Ile Asn Pro Ser Thr Ile Ile Asp Phe Asp Asn Gly
245     250     255
Val Pro Leu Arg Asp Ile Tyr Thr Phe Lys Arg Gly Ser Asn Ser Arg
260     265     270
Tyr Ala Lys Ala Ser Lys Thr Ser Ser Arg Tyr Ala Lys Lys Gly Lys
275     280     285
Lys Gly Arg Gln Ala Ser Ser Pro Met Thr Tyr Arg Ile Lys Lys Gly
290     295     300
Asp Thr Leu Glu Thr Ile Ala Lys Arg His Gly Thr Ser Val Gln Lys
305     310     315     320
Leu Cys Ala Thr Asn Gly Ile Gly Lys Ser Lys Ile Leu Thr Pro Gly

```

```

325
Lys Ala Leu Arg Ile Lys
340

```

330

335

(2) INFORMACIÓN PARA LA SEQ ID NO: 382

(i) CARACTERÍSTICAS DE SECUENCIA:

(A) LONGITUD: 563 aminoácidos

(B) TIPO: aminoácido

5 (D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: proteína

(iii) HIPOTÉTICO: SI

(vi) FUENTE ORIGINAL:

(A) ORGANISMO: Porfiromonas gingivalis

10 (ix) CARACTERÍSTICA:

(A) NOMBRE/CLAVE: misc\_feature

(B) UBICACIÓN 1...563

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 382

ES 2 377 751 T3

His His Lys Thr Tyr Gln Thr Met Lys Lys Leu Leu Gln Ala Lys Ala  
 1 5 10 15  
 Leu Ile Leu Ala Leu Gly Leu Phe Gln Leu Pro Ala Ile Ala Gln Thr  
 20 25 30  
 Gln Met Gln Ala Asp Arg Thr Asn Gly Gln Phe Ala Thr Glu Glu Met  
 35 40 45  
 Gln Arg Ala Phe Gln Glu Thr Asn Pro Pro Ala Gly Pro Val Arg Ala  
 50 55 60  
 Ile Ala Glu Tyr Glu Arg Ser Ala Ala Val Leu Val Arg Tyr Pro Phe  
 65 70 75 80  
 Gly Ile Pro Met Glu Leu Ile Lys Glu Leu Ala Lys Asn Asp Lys Val  
 85 90 95  
 Ile Thr Ile Val Ala Ser Glu Ser Gln Lys Asn Thr Val Ile Thr Gln  
 100 105 110  
 Tyr Thr Gln Ser Gly Val Asn Leu Ser Asn Cys Asp Phe Ile Ile Ala  
 115 120 125  
 Lys Thr Asp Ser Tyr Trp Thr Arg Asp Tyr Thr Gly Trp Phe Ala Met  
 130 135 140  
 Tyr Asp Thr Asn Lys Val Gly Leu Val Asp Phe Ile Tyr Asn Arg Pro  
 145 150 155 160  
 Arg Pro Asn Asp Asp Glu Phe Pro Lys Tyr Glu Ala Gln Tyr Leu Gly  
 165 170 175  
 Ile Glu Met Phe Gly Met Lys Leu Lys Gln Thr Gly Gly Asn Tyr Met  
 180 185 190  
 Thr Asp Gly Tyr Gly Ser Ala Val Gln Ser His Ile Ala Tyr Thr Glu  
 195 200 205  
 Asn Ser Ser Leu Ser Gln Ala Gln Val Asn Gln Lys Met Lys Asp Tyr  
 210 215 220  
 Leu Gly Ile Thr His His Asp Val Val Gln Asp Pro Asn Gly Glu Tyr  
 225 230 235 240  
 Ile Asn His Val Asp Cys Trp Gly Lys Tyr Leu Ala Pro Asn Lys Ile  
 245 250 255  
 Leu Ile Arg Lys Val Pro Asp Asn His Pro Gln His Gln Ala Leu Glu  
 260 265 270  
 Asp Met Ala Ala Tyr Phe Ala Ala Gln Thr Cys Ala Trp Gly Thr Lys  
 275 280 285  
 Tyr Glu Val Tyr Arg Ala Leu Ala Thr Asn Glu Gln Pro Tyr Thr Asn  
 290 295 300  
 Ser Leu Ile Leu Asn Asn Arg Val Phe Val Pro Val Asn Gly Pro Ala  
 305 310 315 320  
 Ser Val Asp Asn Asp Ala Leu Asn Val Tyr Lys Thr Ala Met Pro Gly  
 325 330 335  
 Tyr Glu Ile Ile Gly Val Lys Gly Ala Ser Gly Thr Pro Trp Leu Gly  
 340 345 350  
 Thr Asp Ala Leu His Cys Arg Thr His Glu Val Ala Asp Lys Gly Tyr  
 355 360 365  
 Leu Tyr Ile Lys His Tyr Pro Ile Leu Gly Glu Gln Ala Gly Pro Asp  
 370 375 380  
 Tyr Lys Ile Glu Ala Asp Val Val Ser Cys Ala Asn Ala Thr Ile Ser  
 385 390 395 400  
 Pro Val Gln Cys Tyr Tyr Arg Ile Asn Gly Ser Gly Ser Phe Lys Ala  
 405 410 415  
 Ala Asp Met Thr Met Glu Ser Thr Gly His Tyr Thr Tyr Ser Phe Thr  
 420 425 430  
 Gly Leu Asn Lys Asn Asp Lys Val Glu Tyr Tyr Ile Ser Ala Ala Asp  
 435 440 445  
 Asn Ser Gly Arg Lys Glu Thr Tyr Pro Phe Ile Gly Glu Pro Asp Pro  
 450 455 460  
  
 Phe Lys Phe Thr Cys Met Asn Glu Thr Asn Thr Cys Thr Val Thr Gly  
 465 470 475 480  
 Ala Ala Lys Ala Leu Arg Ala Trp Phe Asn Ala Gly Arg Ser Glu Leu  
 485 490 495  
 Ala Val Ser Val Ser Leu Asn Ile Ala Gly Thr Tyr Arg Ile Lys Leu  
 500 505 510  
 Tyr Asn Thr Ala Gly Glu Glu Val Ala Ala Met Thr Lys Glu Leu Val  
 515 520 525  
 Ala Gly Thr Ser Val Phe Ser Met Asp Val Tyr Ser Gln Ala Pro Gly  
 530 535 540  
 Thr Tyr Val Leu Val Val Glu Gly Asn Gly Ile Arg Glu Thr Met Lys  
 545 550 555 560  
 Ile Leu Lys

(2) INFORMACIÓN PARA LA SEQ ID NO: 383

(i) CARACTERÍSTICAS DE SECUENCIA:

(A) LONGITUD: 437 aminoácidos

5 (B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: proteína

(iii) HIPOTÉTICO: SI

(vi) FUENTE ORIGINAL:

5 (A) ORGANISMO: Porfiromonas gingivalis

(ix) CARACTERÍSTICA:

(A) NOMBRE/CLAVE: misc\_feature

(B) UBICACIÓN 1...437

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 383

```

Thr Thr Asn Arg Lys Pro Asn Thr Asn Met Lys Leu Ser Ser Lys Lys
1 5 10 15
Ile Leu Ala Ile Ile Ala Leu Leu Thr Met Gly His Ala Val Gln Ala
20 25 30
Gln Phe Val Pro Ala Pro Thr Thr Gly Ile Arg Met Ser Val Thr Thr
35 40 45
Thr Lys Ala Val Gly Glu Lys Ile Glu Leu Leu Val His Ser Ile Glu
50 55 60
Lys Lys Gly Ile Trp Ile Asp Leu Asn Gly Asp Ala Thr Tyr Gln Gln
65 70 75 80
Gly Glu Glu Ile Thr Val Phe Asp Glu Ala Tyr His Glu Tyr Thr Ile
85 90 95
Gly Thr Gln Thr Leu Thr Ile Tyr Gly Asn Thr Thr Arg Leu Gly Cys
100 105 110
Arg Ser Thr Gly Ala Thr Ala Val Asp Val Thr Lys Asn Pro Asn Leu
115 120 125
Thr Tyr Leu Ala Cys Pro Lys Asn Asn Leu Lys Ser Leu Asp Leu Thr
130 135 140
Gln Asn Pro Lys Leu Leu Arg Val Trp Cys Asp Ser Asn Glu Ile Glu
145 150 155 160
Ser Leu Asp Leu Ser Gly Asn Pro Ala Leu Ile Ile Leu Gly Cys Asp
165 170 175
Arg Asn Lys Leu Thr Glu Leu Lys Thr Asp Asn Asn Pro Lys Leu Ala
180 185 190
Ser Leu Trp Cys Ser Asp Asn Asn Leu Thr Glu Leu Glu Leu Ser Ala
195 200 205
Asn Pro Arg Leu Asn Asp Leu Trp Cys Phe Gly Asn Arg Ile Thr Lys
210 215 220
Leu Asp Leu Ser Ala Asn Pro Leu Leu Val Thr Leu Trp Cys Ser Asp
225 230 235 240
Asn Glu Leu Ser Thr Leu Asp Leu Ser Lys Asn Ser Asp Val Ala Tyr
245 250 255
Leu Trp Cys Ser Ser Asn Lys Leu Thr Ser Leu Asn Leu Ser Gly Val
260 265 270
Lys Gly Leu Ser Val Leu Val Cys His Ser Asn Gln Ile Ala Gly Glu
275 280 285
Glu Met Thr Lys Val Val Asn Ala Leu Pro Thr Leu Ser Pro Gly Ala
290 295 300
Gly Ala Gln Ser Lys Phe Val Val Val Asp Leu Lys Asp Thr Asp Glu
305 310 315 320

```

10

```

Lys Asn Ile Cys Thr Val Lys Asp Val Glu Lys Ala Lys Ser Lys Asn
325 330 335
Trp Arg Val Phe Asp Phe Asn Gly Asp Ser Asp Asn Met Leu Pro Tyr
340 345 350
Glu Gly Ser Pro Thr Ser Asn Leu Ala Val Asp Ala Pro Thr Val Arg
355 360 365
Ile Tyr Pro Asn Pro Val Gly Arg Tyr Ala Leu Val Glu Ile Pro Glu
370 375 380
Ser Leu Leu Gly Gln Glu Ala Ala Leu Tyr Asp Met Asn Gly Val Lys
385 390 395 400
Val Tyr Ser Phe Ala Val Glu Ser Leu Arg Gln Asn Ile Asp Leu Thr
405 410 415
His Leu Pro Asp Gly Thr Tyr Phe Phe Arg Leu Asp Asn Tyr Thr Thr
420 425 430
Lys Leu Ile Lys Gln
435

```

(2) INFORMACIÓN PARA LA SEQ ID NO: 386

(i) CARACTERÍSTICAS DE SECUENCIA:

(A) LONGITUD: 451 aminoácidos

(B) TIPO: aminoácido

5 (D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: proteína

(iii) HIPOTÉTICO: SI

(vi) FUENTE ORIGINAL:

(A) ORGANISMO: Porfiromonas gingivalis

10 (ix) CARACTERÍSTICA:

(A) NOMBRE/CLAVE: misc\_feature

(B) UBICACIÓN 1... 451

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 386

ES 2 377 751 T3

Met Ala Lys Val Ile Lys Thr Lys Lys Gly Leu Ala Leu Asn Leu Lys  
 1 5 10 15  
 Gly Lys Pro Leu Pro Glu Met Leu Ala Glu Pro Ala Gln Ser Pro Thr  
 20 25 30  
 Tyr Ala Val Val Pro Asp Asp Phe Glu Gly Val Ile Pro Lys Val Thr  
 35 40 45  
 Ala Arg Pro Gly Asp Lys Val Arg Ala Gly Ser Ala Leu Met His His  
 50 55 60  
 Lys Ala Tyr Pro Glu Met Lys Phe Thr Ser Pro Val Ser Gly Glu Val  
 65 70 75 80  
 Ile Ala Val Asn Arg Gly Ala Lys Arg Lys Val Leu Ser Ile Glu Val  
 85 90 95  
 Lys Pro Asp Gly Leu Asn Glu Tyr Glu Ser Phe Pro Val Gly Asp Pro  
 100 105 110  
 Ser Ala Leu Ser Ala Glu Gln Ile Lys Glu Leu Leu Leu Ser Ser Gly  
 115 120 125  
 Met Trp Gly Phe Ile Lys Gln Arg Pro Tyr Asp Ile Val Ala Thr Pro  
 130 135 140  
 Asp Ile Ala Pro Arg Asp Ile Tyr Ile Thr Ala Asn Phe Thr Ala Pro  
 145 150 155 160  
 Leu Ala Pro Asp Phe Asp Phe Ile Val Arg Gly Glu Glu Arg Ala Leu  
 165 170 175  
 Gln Thr Ala Ile Asp Ala Leu Ala Lys Leu Thr Thr Gly Lys Val Tyr  
 180 185 190  
 Val Gly Leu Lys Pro Gly Ser Ser Leu Gly Leu His Asn Ala Glu Ile  
 195 200 205  
 Val Glu Val His Gly Pro His Pro Ala Gly Asn Val Gly Val Leu Ile  
 210 215 220  
 Asn His Thr Lys Pro Ile Asn Arg Gly Glu Thr Val Trp Thr Leu Lys  
 225 230 235 240  
 Ala Thr Asp Leu Ile Val Ile Gly Arg Phe Leu Leu Thr Gly Lys Ala  
 245 250 255  
 Asp Phe Thr Arg Met Ile Ala Met Thr Gly Ser Asp Ala Ala Ala His  
 260 265 270  
 Gly Tyr Val Arg Ile Met Pro Gly Cys Asn Val Phe Ala Ser Phe Pro  
 275 280 285  
 Gly Arg Leu Thr Ile Lys Glu Ser His Glu Arg Val Ile Asp Gly Asn  
 290 295 300  
 Val Leu Thr Gly Lys Lys Leu Cys Glu Lys Glu Pro Phe Leu Ser Ala  
 305 310 315 320  
 Arg Cys Asp Gln Ile Thr Val Ile Pro Glu Gly Asp Asp Val Asp Glu  
 325 330 335  
 Leu Phe Gly Trp Ala Ala Pro Arg Leu Asp Gln Tyr Ser Met Ser Arg  
 340 345 350  
 Ala Tyr Phe Ser Trp Leu Gln Gly Lys Asn Lys Glu Tyr Val Leu Asp  
 355 360 365  
 Ala Arg Ile Lys Gly Gly Glu Arg Ala Met Ile Met Ser Asn Glu Tyr  
 370 375 380

Asp Arg Val Phe Pro Met Asp Ile Tyr Pro Glu Tyr Leu Leu Lys Ala  
 385 390 395 400  
 Ile Ile Ala Phe Asp Ile Asp Lys Met Glu Asp Leu Gly Ile Tyr Glu  
 405 410 415  
 Val Ala Pro Glu Asp Phe Ala Thr Cys Glu Phe Val Asp Thr Ser Lys  
 420 425 430  
 Ile Glu Leu Gln Arg Ile Val Arg Glu Gly Leu Asp Met Leu Tyr Lys  
 435 440 445  
 Glu Met Asn  
 450

(2) INFORMACIÓN PARA LA SEQ ID NO: 426

(j) CARACTERÍSTICAS DE SECUENCIA:

(A) LONGITUD: 821 aminoácidos

5 (B) TIPO: aminoácido

(D) TOPOLOGÍA: linear

(ii) TIPO DE MOLÉCULA: proteína

(iii) HIPOTÉTICO: SI

(vi) FUENTE ORIGINAL:

(A) ORGANISMO: Porfiromonas gingivalis

(ix) CARACTERÍSTICA:

(A) NOMBRE/CLAVE: misc\_feature

(B) UBICACIÓN 1...821

5 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 426

```

Met Lys Lys Lys Asn Phe Leu Leu Leu Gly Ile Phe Val Ala Leu Leu
1      5      10      15
Thr Phe Ile Gly Ser Met Gln Ala Gln Gln Ala Lys Asp Tyr Phe Asn
20      25      30
Phe Asp Glu Arg Gly Glu Ala Tyr Phe Ser Phe Lys Val Pro Asp Arg
35      40      45
Ala Val Leu Gln Glu Leu Ala Leu Ile Met Ser Ile Asp Glu Phe Asp
50      55      60
Pro Val Thr Asn Glu Ala Ile Ala Tyr Ala Ser Glu Glu Glu Phe Glu
65      70      75      80
Ala Phe Leu Arg Tyr Gly Leu Lys Pro Thr Phe Leu Thr Pro Pro Ser
85      90      95
Met Gln Arg Ala Val Glu Met Phe Asp Tyr Arg Ser Gly Glu Lys Tyr
100     105     110
Glu Trp Asn Ala Tyr Pro Thr Tyr Glu Ala Tyr Ile Ser Met Met Glu
115     120     125
Glu Phe Gln Thr Lys Tyr Pro Ser Leu Cys Thr Thr Ser Val Ile Gly
130     135     140
Lys Ser Val Lys Asp Arg Lys Leu Met Ile Cys Lys Leu Thr Ser Ser
145     150     155     160
Ala Asn Thr Gly Lys Lys Pro Arg Val Leu Tyr Thr Ser Thr Met His
165     170     175
Gly Asp Glu Thr Thr Gly Tyr Val Val Leu Leu Arg Leu Ile Asp His
180     185     190
Leu Leu Ser Asn Tyr Glu Ser Asp Pro Arg Ile Lys Asn Ile Leu Asp
195     200     205
Lys Thr Glu Val Trp Ile Cys Pro Leu Thr Asn Pro Asp Gly Ala Tyr
210     215     220
Arg Ala Gly Asn His Thr Val Gln Gly Ala Thr Arg Tyr Asn Ala Asn
225     230     235     240
Asn Val Asp Leu Asn Arg Asn Phe Lys Asp Asp Val Ala Gly Asp His
245     250     255
Pro Asp Gly Lys Pro Trp Gln Pro Glu Ala Thr Ala Phe Met Asp Leu
260     265     270
Glu Gly Asn Thr Ser Phe Val Leu Gly Ala Asn Ile His Gly Gly Thr
275     280     285
Glu Val Val Asn Tyr Pro Trp Asp Asn Lys Lys Glu Arg His Ala Asp
290     295     300
Asp Glu Trp Tyr Lys Leu Ile Ser Arg Asn Tyr Ala Ala Ala Cys Gln
305     310     315     320

```

ES 2 377 751 T3

Ser Ile Ser Ala Ser Tyr Met Thr Ser Glu Thr Asn Ser Gly Ile Ile  
 325 330 335  
 Asn Gly Ser Asp Trp Tyr Val Ile Arg Gly Ser Arg Gln Asp Asn Ala  
 340 345 350  
 Asn Tyr Phe His Arg Leu Arg Glu Ile Thr Leu Glu Ile Ser Asn Thr  
 355 360 365  
 Lys Leu Val Pro Ala Ser Gln Leu Pro Lys Tyr Trp Asn Leu Asn Lys  
 370 375 380  
 Glu Ser Leu Leu Ala Leu Ile Glu Glu Ser Leu Tyr Gly Ile His Gly  
 385 390 395 400  
 Thr Val Thr Ser Ala Ala Asn Gly Gln Pro Leu Lys Cys Gln Ile Leu  
 405 410 415  
 Ile Glu Asn His Asp Lys Arg Asn Ser Asp Val Tyr Ser Asp Ala Thr  
 420 425 430  
 Thr Gly Tyr Tyr Val Arg Pro Ile Lys Ala Gly Thr Tyr Thr Val Lys  
 435 440 445  
 Tyr Lys Ala Glu Gly Tyr Pro Glu Ala Thr Arg Thr Ile Thr Ile Lys  
 450 455 460  
 Asp Lys Glu Thr Val Ile Met Asp Ile Ala Leu Gly Asn Ser Val Pro  
 465 470 475 480  
 Leu Pro Val Pro Asp Phe Thr Ala Ser Pro Met Thr Ile Ser Val Gly  
 485 490 495  
 Glu Ser Val Gln Phe Gln Asp Gln Thr Thr Asn Asn Pro Thr Asn Trp  
 500 505 510  
 Glu Trp Thr Phe Glu Gly Gly Gln Pro Ala Met Ser Thr Glu Gln Asn  
 515 520 525  
 Pro Leu Val Ser Tyr Ser His Pro Gly Gln Tyr Asp Val Thr Leu Lys  
 530 535 540  
 Val Trp Asn Ala Ser Gly Ser Asn Thr Ile Thr Lys Glu Lys Phe Ile  
 545 550 555 560  
 Thr Val Asn Ala Val Met Pro Val Ala Glu Phe Val Gly Thr Pro Thr  
 565 570 575  
 Glu Ile Glu Glu Gly Gln Thr Val Ser Phe Gln Asn Gln Ser Thr Asn  
 580 585 590  
 Ala Thr Asn Tyr Val Trp Ile Phe Asp Gly Gly Thr Pro Ala Thr Ser  
 595 600 605  
 Glu Asp Glu Asn Pro Thr Val Leu Tyr Ser Lys Ala Gly Gln Tyr Asp  
 610 615 620  
 Val Thr Leu Lys Ala Ile Ser Ala Ser Gly Glu Thr Val Lys Thr Lys  
 625 630 635 640  
 Glu Lys Tyr Ile Thr Val Lys Lys Ala Pro Val Pro Ala Pro Val Ala  
 645 650 655  
 Asp Phe Glu Gly Thr Pro Arg Lys Val Lys Lys Gly Glu Thr Val Thr  
 660 665 670  
 Phe Lys Asp Leu Ser Thr Asn Asn Pro Thr Ser Trp Leu Trp Val Phe  
 675 680 685  
 Glu Gly Gly Ser Pro Ala Thr Ser Thr Glu Gln Asn Pro Val Val Thr  
 690 695 700  
 Tyr Asn Glu Thr Gly Lys Tyr Asp Val Gln Leu Thr Ala Thr Asn Glu  
 705 710 715 720  
 Gly Gly Ser Asn Val Lys Lys Ala Glu Asp Tyr Ile Glu Val Ile Leu  
 725 730 735  
 Asp Asp Ser Val Glu Asp Ile Val Ala Gln Thr Gly Ile Val Ile Arg  
 740 745 750  
 Pro Gln Asn Gly Thr Lys Gln Ile Leu Ile Glu Ala Asn Ala Ala Ile  
 755 760 765  
 Lys Ala Ile Val Leu Tyr Asp Ile Asn Gly Arg Val Val Leu Lys Thr  
 770 775 780  
 Thr Pro Asn Gln Leu Arg Ser Thr Val Asp Leu Ser Ile Leu Pro Glu  
 785 790 795 800  
 Gly Ile Tyr Thr Ile Asn Ile Lys Thr Glu Lys Ser Ala Arg Thr Glu  
 805 810 815  
 Lys Ile His Ile Gly  
 820

(2) INFORMACIÓN PARA LA SEQ ID NO: 434

(i) CARACTERÍSTICAS DE SECUENCIA:

(A) LONGITUD: 223 aminoácidos

5 (B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: proteína

(iii) HIPOTÉTICO: SI

(vi) FUENTE ORIGINAL:

(A) ORGANISMO: Porfiromonas gingivalis

(ix) CARACTERÍSTICA:

(A) NOMBRE/CLAVE: misc\_feature

(B) UBICACIÓN 1...223

5 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 434

```

Met Lys Lys Ser Ser Val Val Ala Ser Val Leu Ala Val Ala Leu Val
1      5      10      15
Phe Ala Gly Cys Gly Leu Asn Asn Met Ala Lys Gly Gly Leu Ile Gly
20      25      30
Ala Gly Val Gly Gly Ala Ile Gly Ala Gly Val Gly Asn Val Ala Gly
35      40      45
Asn Thr Ala Val Gly Ala Ile Val Gly Thr Ala Val Gly Gly Ala Ala
50      55      60
Gly Ala Leu Ile Gly Lys Lys Met Asp Lys Gln Lys Lys Glu Leu Glu
65      70      75      80
Ala Ala Val Pro Asp Ala Thr Ile Gln Thr Val Asn Asp Gly Glu Ala
85      90      95
Ile Leu Val Thr Phe Asp Ser Gly Ile Leu Phe Ala Thr Asn Ser Ser
100     105     110
Thr Leu Ser Pro Asn Ser Arg Thr Ala Leu Thr Lys Phe Ala Ala Asn
115     120     125
Met Asn Lys Asn Pro Asp Thr Asp Ile Arg Ile Val Gly His Thr Asp
130     135     140
Asn Thr Gly Ser Asp Lys Ile Asn Asp Pro Leu Ser Glu Arg Arg Ala
145     150     155     160
Ala Ser Val Tyr Ser Phe Leu Asn Ser Gln Gly Val Ser Met Ser Arg
165     170     175
Met Ala Ala Glu Gly Arg Gly Ser His Glu Pro Val Ala Asp Asn Ser
180     185     190
Thr Val Ala Gly Arg Ser Ala Asn Arg Arg Val Glu Val Tyr Ile Leu
195     200     205
Pro Asn Ala Lys Met Ile Glu Gln Ala Gln Gln Gly Thr Leu Lys
210     215     220
    
```

(2) INFORMACIÓN PARA LA SEQ ID NO: 435

(i) CARACTERÍSTICAS DE SECUENCIA:

(A) LONGITUD: 337 aminoácidos

10 (B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: proteína

(iii) HIPOTÉTICO: SI

(vi) FUENTE ORIGINAL:

15 (A) ORGANISMO: Porfiromonas gingivalis

(ix) CARACTERÍSTICA:

(A) NOMBRE/CLAVE: misc\_feature

(B) UBICACIÓN 1...337

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 435

ES 2 377 751 T3

```

Met Ser Lys Lys Ser Ile Leu Leu Leu Cys Cys Ser Leu Cys Phe Ile
1 5 10 15
Ser Ala Thr Lys Ala Val Thr Pro Val Arg Asn Val Arg Asn Ser Gln
20 25 30
Val Asn Ser Lys Ala Lys Thr Glu Arg Thr Lys Pro Ser Asp Ser Val
35 40 45
Arg Tyr Ile Ser Asn Met Ile Ala Asp Arg Leu Glu Phe Arg Asn Lys
50 55 60
Ile Ser Ser Glu Lys Glu Val Arg Lys Ala Glu Tyr Glu Asn Arg Leu
65 70 75 80
Ala Met Glu Ala Leu Asn Tyr Pro Ala Ile Asp Leu Tyr Gly Glu Asp
85 90 95
Ser Trp Ser Glu Tyr Val Asn Pro Phe Val Gly Ala Gly Thr Asp Val
100 105 110
Glu Ile Pro Asn Ser Tyr Asp Ile Asp Cys Ser Ser Phe Val Met Pro
115 120 125
Val Glu Asp Lys Gln Val Thr Ser Gln Phe Gly Tyr Arg Arg Arg Phe
130 135 140
Gly Arg Met His Tyr Gly Ile Asp Leu Ser Val Asn Arg Gly Asp Thr
145 150 155 160
Ile Arg Ala Ala Phe Asp Gly Lys Val Arg Val Arg Ser Tyr Glu Ala
165 170 175
Arg Gly Tyr Gly Tyr Tyr Ile Val Leu Arg His Pro Asn Gly Leu Glu
180 185 190
Thr Val Tyr Gly His Met Ser Arg Gln Leu Val Asp Glu Asn Gln Ile
195 200 205
Val Arg Ala Gly Gln Pro Ile Gly Leu Gly Gly Ser Thr Gly Arg Ser
210 215 220
Thr Gly Pro His Leu His Phe Glu Thr Arg Phe Met Gly Ile Pro Ile
225 230 235 240
Asn Pro Ser Thr Ile Ile Asp Phe Asp Asn Gly Val Pro Leu Arg Asp
245 250 255
Ile Tyr Thr Phe Lys Arg Gly Ser Asn Ser Arg Tyr Ala Lys Ala Ser
260 265 270
Lys Thr Ser Ser Arg Tyr Ala Lys Lys Gly Lys Lys Gly Arg Gln Ala
275 280 285
Ser Ser Pro Met Thr Tyr Arg Ile Lys Lys Gly Asp Thr Leu Glu Thr
290 295 300
Ile Ala Lys Arg His Gly Thr Ser Val Gln Lys Leu Cys Ala Thr Asn
305 310 315 320
Gly Ile Gly Lys Ser Lys Ile Leu Thr Pro Gly Lys Ala Leu Arg Ile
325 330 335
Lys

```

(2) INFORMACIÓN PARA LA SEQ ID NO: 525

(i) CARACTERÍSTICAS DE SECUENCIA:

(A) LONGITUD: 556 aminoácidos

5 (B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: proteína

(iii) HIPOTÉTICO: SI

(vi) FUENTE ORIGINAL:

10 (A) ORGANISMO: Porfiromonas gingivalis

(ix) CARACTERÍSTICA:

(A) NOMBRE/CLAVE: misc\_feature

(B) UBICACIÓN 1...556

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 525

ES 2 377 751 T3

Met Lys Lys Leu Leu Gln Ala Lys Ala Leu Ile Leu Ala Leu Gly Leu  
 1 5 10 15  
 Phe Gln Leu Pro Ala Ile Ala Gln Thr Gln Met Gln Ala Asp Arg Thr  
 20 25 30  
 Asn Gly Gln Phe Ala Thr Glu Glu Met Gln Arg Ala Phe Gln Glu Thr  
 35 40 45  
 Asn Pro Pro Ala Gly Pro Val Arg Ala Ile Ala Glu Tyr Glu Arg Ser  
 50 55 60  
 Ala Ala Val Leu Val Arg Tyr Pro Phe Gly Ile Pro Met Glu Leu Ile  
 65 70 75 80  
 Lys Glu Leu Ala Lys Asn Asp Lys Val Ile Thr Ile Val Ala Ser Glu  
 85 90 95  
 Ser Gln Lys Asn Thr Val Ile Thr Gln Tyr Thr Gln Ser Gly Val Asn  
 100 105 110  
 Leu Ser Asn Cys Asp Phe Ile Ile Ala Lys Thr Asp Ser Tyr Trp Thr  
 115 120 125

Arg Asp Tyr Thr Gly Trp Phe Ala Met Tyr Asp Thr Asn Lys Val Gly  
 130 135 140  
 Leu Val Asp Phe Ile Tyr Asn Arg Pro Arg Pro Asn Asp Asp Glu Phe  
 145 150 155 160  
 Pro Lys Tyr Glu Ala Gln Tyr Leu Gly Ile Glu Met Phe Gly Met Lys  
 165 170 175  
 Leu Lys Gln Thr Gly Gly Asn Tyr Met Thr Asp Gly Tyr Gly Ser Ala  
 180 185 190  
 Val Gln Ser His Ile Ala Tyr Thr Glu Asn Ser Ser Leu Ser Gln Ala  
 195 200 205  
 Gln Val Asn Gln Lys Met Lys Asp Tyr Leu Gly Ile Thr His His Asp  
 210 215 220  
 Val Val Gln Asp Pro Asn Gly Glu Tyr Ile Asn His Val Asp Cys Trp  
 225 230 235 240  
 Gly Lys Tyr Leu Ala Pro Asn Lys Ile Leu Ile Arg Lys Val Pro Asp  
 245 250 255  
 Asn His Pro Gln His Gln Ala Leu Glu Asp Met Ala Ala Tyr Phe Ala  
 260 265 270  
 Ala Gln Thr Cys Ala Trp Gly Thr Lys Tyr Glu Val Tyr Arg Ala Leu  
 275 280 285  
 Ala Thr Asn Glu Gln Pro Tyr Thr Asn Ser Leu Ile Leu Asn Asn Arg  
 290 295 300  
 Val Phe Val Pro Val Asn Gly Pro Ala Ser Val Asp Asn Asp Ala Leu  
 305 310 315 320  
 Asn Val Tyr Lys Thr Ala Met Pro Gly Tyr Glu Ile Ile Gly Val Lys  
 325 330 335  
 Gly Ala Ser Gly Thr Pro Trp Leu Gly Thr Asp Ala Leu His Cys Arg  
 340 345 350  
 Thr His Glu Val Ala Asp Lys Gly Tyr Leu Tyr Ile Lys His Tyr Pro  
 355 360 365  
 Ile Leu Gly Glu Gln Ala Gly Pro Asp Tyr Lys Ile Glu Ala Asp Val  
 370 375 380  
 Val Ser Cys Ala Asn Ala Thr Ile Ser Pro Val Gln Cys Tyr Tyr Arg  
 385 390 395 400  
 Ile Asn Gly Ser Gly Ser Phe Lys Ala Ala Asp Met Thr Met Glu Ser  
 405 410 415  
 Thr Gly His Tyr Thr Tyr Ser Phe Thr Gly Leu Asn Lys Asn Asp Lys  
 420 425 430  
 Val Glu Tyr Tyr Ile Ser Ala Ala Asp Asn Ser Gly Arg Lys Glu Thr  
 435 440 445  
 Tyr Pro Phe Ile Gly Glu Pro Asp Pro Phe Lys Phe Thr Cys Met Asn  
 450 455 460  
 Glu Thr Asn Thr Cys Thr Val Thr Gly Ala Ala Lys Ala Leu Arg Ala  
 465 470 475 480  
 Trp Phe Asn Ala Gly Arg Ser Glu Leu Ala Val Ser Val Ser Leu Asn  
 485 490 495  
 Ile Ala Gly Thr Tyr Arg Ile Lys Leu Tyr Asn Thr Ala Gly Glu Glu  
 500 505 510  
 Val Ala Ala Met Thr Lys Glu Leu Val Ala Gly Thr Ser Val Phe Ser  
 515 520 525  
 Met Asp Val Tyr Ser Gln Ala Pro Gly Thr Tyr Val Leu Val Val Glu  
 530 535 540  
 Gly Asn Gly Ile Arg Glu Thr Met Lys Ile Leu Lys  
 545 550 555

(2) INFORMACIÓN PARA LA SEQ ID NO: 526

(i) CARACTERÍSTICAS DE SECUENCIA:

(A) LONGITUD: 428 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

5 (ii) TIPO DE MOLÉCULA: proteína

(iii) HIPOTÉTICO: SI

(vi) FUENTE ORIGINAL:

(A) ORGANISMO: Porfiromonas gingivalis

(ix) CARACTERÍSTICA:

10 (A) NOMBRE/CLAVE: misc\_feature

(B) UBICACIÓN 1...428

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 526

```

Met Lys Leu Ser Ser Lys Lys Ile Leu Ala Ile Ile Ala Leu Leu Thr
1      5      10      15
Met Gly His Ala Val Gln Ala Gln Phe Val Pro Ala Pro Thr Thr Gly
20      25      30
Ile Arg Met Ser Val Thr Thr Thr Lys Ala Val Gly Glu Lys Ile Glu
35      40      45
Leu Leu Val His Ser Ile Glu Lys Lys Gly Ile Trp Ile Asp Leu Asn
50      55      60
Gly Asp Ala Thr Tyr Gln Gln Gly Glu Glu Ile Thr Val Phe Asp Glu
65      70      75      80
Ala Tyr His Glu Tyr Thr Ile Gly Thr Gln Thr Leu Thr Ile Tyr Gly
95
Asn Thr Thr Arg Leu Gly Cys Arg Ser Thr Gly Ala Thr Ala Val Asp
100      105      110
Val Thr Lys Asn Pro Asn Leu Thr Tyr Leu Ala Cys Pro Lys Asn Asn
115      120      125
Leu Lys Ser Leu Asp Leu Thr Gln Asn Pro Lys Leu Leu Arg Val Trp
130      135      140
Cys Asp Ser Asn Glu Ile Glu Ser Leu Asp Leu Ser Gly Asn Pro Ala
145      150      155      160
Leu Ile Ile Leu Gly Cys Asp Arg Asn Lys Leu Thr Glu Leu Lys Thr
165      170      175
Asp Asn Asn Pro Lys Leu Ala Ser Leu Trp Cys Ser Asp Asn Asn Leu
180      185      190
Thr Glu Leu Glu Leu Ser Ala Asn Pro Arg Leu Asn Asp Leu Trp Cys
195      200      205
Phe Gly Asn Arg Ile Thr Lys Leu Asp Leu Ser Ala Asn Pro Leu Leu
210      215      220
Val Thr Leu Trp Cys Ser Asp Asn Glu Leu Ser Thr Leu Asp Leu Ser
225      230      235      240
Lys Asn Ser Asp Val Ala Tyr Leu Trp Cys Ser Ser Asn Lys Leu Thr
245      250      255
Ser Leu Asn Leu Ser Gly Val Lys Gly Leu Ser Val Leu Val Cys His
260      265      270
Ser Asn Gln Ile Ala Gly Glu Glu Met Thr Lys Val Val Asn Ala Leu
275      280      285
Pro Thr Leu Ser Pro Gly Ala Gly Ala Gln Ser Lys Phe Val Val Val
290      295      300
Asp Leu Lys Asp Thr Asp Glu Lys Asn Ile Cys Thr Val Lys Asp Val
305      310      315      320
Glu Lys Ala Lys Ser Lys Asn Trp Arg Val Phe Asp Phe Asn Gly Asp
325      330      335
Ser Asp Asn Met Leu Pro Tyr Glu Gly Ser Pro Thr Ser Asn Leu Ala
340      345      350
Val Asp Ala Pro Thr Val Arg Ile Tyr Pro Asn Pro Val Gly Arg Tyr
355      360      365
Ala Leu Val Glu Ile Pro Glu Ser Leu Leu Gly Gln Glu Ala Ala Leu
370      375      380
Tyr Asp Met Asn Gly Val Lys Val Tyr Ser Phe Ala Val Glu Ser Leu
385      390      395      400
Arg Gln Asn Ile Asp Leu Thr His Leu Pro Asp Gly Thr Tyr Phe Phe
405      410      415
Arg Leu Asp Asn Tyr Thr Thr Lys Leu Ile Lys Gln
420      425

```

Referencias.

1. Lipman DJ, Pearson WR. 1985. Rapid and sensitive protein similarity searches. *Science* 277:1435-1441.
2. Horton, P. and Nakai, K. (1996). A probabilistic classification system for predicting the cellular localization sites of proteins. *Intellig. Syst. Mol. Biol.* 4: 109-115.
- 5 3. Nakai K, Kanehisa M. 1991. Expert systems for predicting protein localization sites in Gram-negative bacteria. *Proteins: Structure, Function, and Genetics* 11:95-110.
4. Nielsen H, Engelbrecht J, Brunak S and von Heijne G. 1997. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Engineering* 10, 1-6.
- 10 5. Claros MG and G von Heijne. (1994). TopPred II: an improved software for membrane protein structure predictions. *Comput. Appl. Biosci.* 10: 685-686.
6. Borodovsky M, Rudd KE, and EV Koonin. (1994). Intrinsic and extrinsic approaches for detecting genes in a bacterial genome. *Nucleic Acids Res.* 22: 4756-4767.
7. Struvye M, Moons M, Tommassen J. 1991. Carboxy-terminal phenylalanine is essential for the correct assembly of a bacterial outer membrane protein *J. Mol. Biol.* 218:141-148.
- 15 8. Aduse-Opoku J, Slaney JM, Rangarajan M, Muir J, Young KA, Curtis MA. 1997. The Tla receptor protein of *Porfiromonas gingivalis* W50: a homolog of the RI precursor (PrpRI) is an outer membrane receptor required for growth on low levels of hemin. *J. Bacteriol.* 179:4778-4788.
9. Needleman SB, Munsch CD. 1970. A general method applicable to the search of similarity in the amino acid sequence of two proteins. *J. Molec. Biol.* 48: 443-453.

20

**REIVINDICACIONES**

1. Un polipéptido *Porfiromonas gingivalis* antigénico aislado, el polipéptido comprende:
  - (i) la secuencia de aminoácidos de la SEQ ID NO. 426;
  - (ii) la secuencia de aminoácidos de los residuos 24 a 821 o 27 a 821 de la SEQ ID NO. 426; o
  - (iii) una secuencia de aminoácidos por lo menos 85 % idéntica a una secuencia de aminoácidos de (i).
2. Un polipéptido como se reivindica en la reivindicación 1, que comprende una secuencia de aminoácidos por lo menos 95 % idéntica a una secuencia de aminoácidos de (i).
3. Un polipéptido como se reivindica en la reivindicación 1 o reivindicación 2, en donde la secuencia de aminoácidos es la SEQ ID NO. 426.
4. Una molécula de ADN aislada, la molécula de ADN comprende una secuencia de nucleótido que codifica el polipéptido como se reivindica en la reivindicación 1, 2 o 3.
5. Una molécula de ADN aislada como se reivindica en la reivindicación 4 en la que la molécula de ADN comprende la secuencia de nucleótido de la SEQ ID NO. 162.
6. Un vector de expresión recombinante que comprende la molécula de ADN como se reivindica en la reivindicación 4 o reivindicación 5 ligada operativamente a un elemento regulador de transcripción.
7. Una célula aislada que comprende el vector de expresión recombinante como se reivindica en la reivindicación 6.
8. Un método para producir un polipéptido *P. gingivalis* que comprende cultivar la célula como se reivindica en la reivindicación 7 bajo condiciones que permiten la expresión del polipéptido.
9. Una composición farmacéutica que comprende:
  - (i) una cantidad efectiva de (a) por lo menos un polipéptido como se reivindica en la reivindicación 1, 2 o 3, (b) por lo menos una molécula de ADN como se reivindica en la reivindicación 4 o reivindicación 5, y/o (c) un polipéptido de por lo menos 40 aminoácidos que tiene una secuencia contigua de por lo menos 40 aminoácidos idénticos a una secuencia de aminoácidos contigua de la SEQ ID NO. 426, en donde el polipéptido es capaz de elevar una respuesta inmune contra *P. gingivalis*; y
  - (ii) un portador farmacéuticamente aceptable.
10. Una composición como se reivindica en la reivindicación 9 o reivindicación 10 en la que el portador farmacéuticamente aceptable es un adyuvante.
11. Una composición como se reivindica en la reivindicación 9 o reivindicación 10 para uso en elevar una respuesta inmune dirigida a *P. gingivalis* en un sujeto.
12. Una composición para uso como se reivindica en la reivindicación 11 para tratar un sujeto para infección *P. gingivalis*.
13. Una composición para uso como se reivindica en la reivindicación 12, en donde el tratamiento es un tratamiento profiláctico o un tratamiento terapéutico.
14. Un anticuerpo elevado contra un polipéptido como se reivindica en la reivindicación 1, 2 o 3.
15. Un anticuerpo como se reivindica en la reivindicación 14 en el que el anticuerpo es policlonal o monoclonal.
16. Una composición que comprende por lo menos un anticuerpo como se reivindica en la reivindicación 14 o reivindicación 15.
17. Una composición como se reivindica en la reivindicación 16 que se adapta para uso oral.

18. Un método para detectar la presencia de una molécula de ADN como se reivindica en la reivindicación 4 en una muestra que comprende:

- 5 (a) poner en contacto una muestra con la sonda de nucleótido que comprende por lo menos 18 nucleótidos y que tiene una secuencia contigua de por lo menos 18 nucleótidos idénticos a una secuencia de nucleótido contigua de la SEQ ID NO. 37 bajo condiciones en las que se puede formar un híbrido entre la sonda y la molécula de ADN como se reivindica en la reivindicación 4 en la muestra; y
- (b) detectar el híbrido formado en la etapa (a), en donde la detección de un híbrido indica la presencia de la molécula de ADN como se reivindica en la reivindicación 4 en la muestra.

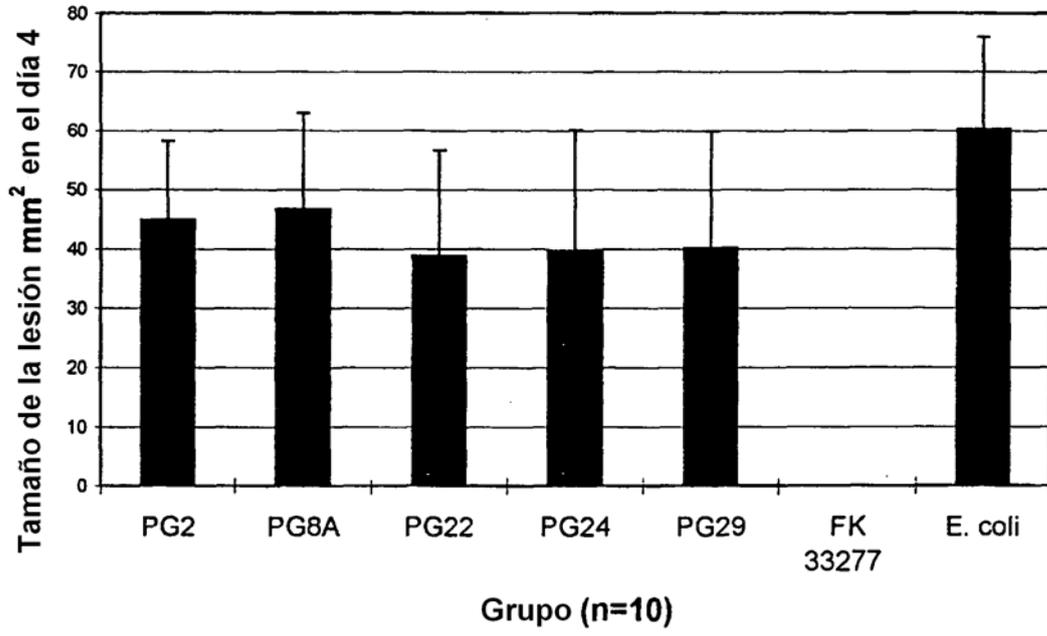


Figura 1

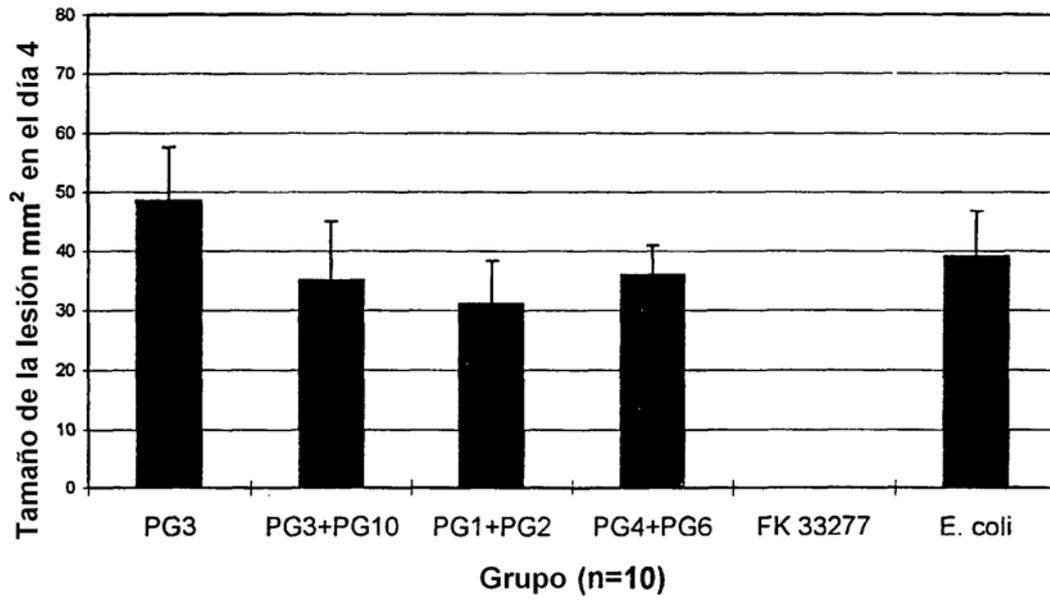


Figura 2