UNIVERSIDAD NACIONAL DE COLOMBIA

Análisis de la variabilidad genética y las fuerzas evolutivas que actúan sobre la proteína del cuello de roptrias 4 de *Plasmodium vivax* (*Pv*RON4), a partir de aislados Colombianos obtenidos entre 2007-2015

Sindy Paola Buitrago Puentes

Universidad Nacional de Colombia

Facultad de ciencias

Maestría en Ciencias - Microbiología

Bogotá. Colombia

2016

Análisis de la variabilidad genética y las fuerzas evolutivas que actúan sobre la proteína del cuello de roptrias 4 de *Plasmodium vivax* (*Pv*RON4), a partir de aislados Colombianos obtenidos entre 2007-2015

Sindy Paola Buitrago Puentes

Tesis presentada como requisito parcial para optar al título de:

Magister en Ciencias-Microbiología

Director:

Diego Edison Garzón Ospina B.Sc., M.Sc.

Codirector:

Manuel Alfonso Patarroyo M.D., Dr.Sc.

Línea de Investigación:

Biología Molecular de Agentes Infecciosos

Grupo de Investigación:

Biología Molecular e Inmunología

Universidad Nacional de Colombia

Facultad de Ciencias

Maestría en Ciencias - Microbiología

Bogotá, Colombia

2016

“No es la especie más fuerte la que sobrevive, ni la más inteligente,
sino la más receptiva al cambio”.

Charles Darwin

Agradecimientos

El esfuerzo de una persona por llegar a cumplir sus sueños, no solo se ve impulsado por su deseo y la voluntad de hacerlo, sino también por los actos y consejos de aquellas personas que diariamente se convierten en una pieza invaluable en nuestras vidas. De esta manera, agradezco primeramente a Dios, quien siempre nos acompaña y a mi familia, las personas más importantes de mi vida, por su apoyo incondicional.

Agradezco a mi Director de tesis Diego Edison Garzón, por su guía, enseñanza y apoyo durante el desarrollo del trabajo. Al Doctor Manuel Alfonso Patarroyo, Co-director de trabajo, por su acompañamiento en este proceso y en general por abrirme las puertas de la FIDIC. Al grupo funcional de Biología Molecular e Inmunología, donde he podido afianzar y adquirir nuevos conocimientos, gracias a la colaboración y el apoyo de mis compañeros y pasantes, personas con un alto grado de compromiso y gusto por la investigación, a quienes admiro.

Agradezco a los docentes de la Universidad Nacional, en especial a Marta Fontanilla, Daniel Uribe y María Teresa Reguero, y a Socorrito Prieto, Secretaria del Posgrado, por su acompañamiento durante la Maestría y su preocupación por el bienestar de sus estudiantes. En general, a todos mis compañeros, quienes hicieron parte en este proceso de formación y aprendizaje.

Finalmente, agradezco especialmente a mi hermano Ramiro Hernando Buitrago por instruirme desde muy pequeña en este camino de agrado y pasión por la ciencia, compartiendo conmigo el respeto por aquellas personas que sacrifican parte de sus vidas para que el mundo adquiera un tanto más de conocimiento.

Resumen

El diseño de una vacuna contra *Plasmodium vivax* se ha centrado en la selección de antígenos involucrados en el proceso de invasión del parásito que presenten bajo polimorfismo, para así evitar las respuestas inmunes alelo-específicas. La proteína de cuello de roptrias-4 (RON4) ha sido caracterizada como un miembro indispensable del *tight junction*, un complejo proteico esencial en la invasión a los hepatocitos y/o eritrocitos. Sin embargo, poco se conoce acerca de su variabilidad genética.

En este estudio, se analizó la diversidad genética y haplotípica del locus *pvron4* a partir de 73 secuencias de ADN obtenidas de aislados clínicos colombianos, colectados entre los años 2007 y 2015. El efecto de la selección natural, así como las relaciones evolutivas entre especies emparentadas, fueron caracterizados mediante enfoques de genética de poblaciones y evolución molecular, con el fin de explicar el patrón de variación observado en el locus. Los resultados mostraron que, aunque *pvron4* posee una baja diversidad genética a nivel de secuencia, un número variable de repeticiones en tándem hacia la región N-terminal genera un alto polimorfismo en tamaño. La región central y C-terminal de la proteína, son altamente conservadas intra- e inter-especies, como consecuencia de la selección purificante. Por lo tanto, estas regiones parecen tener una restricción funcional o estructural, lo que las convierte en regiones prometedoras a incluir en el diseño de una vacuna altamente eficiente, basada en subunidades.

Palabras clave: *Plasmodium vivax*, roptrias, diversidad genética, repeticiones en tándem, *pvron4*, selección natural, restricción funcional.

Abstract

Designing a vaccine against *Plasmodium vivax* has been focused on selecting antigens involved in parasite mechanisms to invade red blood cells; such antigens must have domains with low polymorphism to avoid inducing allele-specific immune responses. The rhoptry neck protein 4 (RON4) has been characterized as a vital member of the tight junction, a protein complex which is essential for invasion of hepatocytes and/or erythrocytes; however, little is known about this locus' genetic diversity.

In this study, the genetic and haplotype diversity of the *pvron4* locus were analyzed from 73 DNA sequences obtained from Colombian *P. vivax* clinical isolates collected between 2007 and 2015. The effect of natural selection and the evolutionary relationships between closely related species were characterized by population genetics and molecular evolutionary approaches, to explain the pattern of genetic variation observed in the locus. The results showed that, although *pvron4* had a low genetic diversity at sequence level, a variable amount of tandem repeats at the N-terminal region led to an extensive size polymorphism. On the other hand, the central and C-terminal regions are highly conserved at intra- and inter-species level, as a consequence of a purifying selection effect. Thus, these regions seem to be under functional or structural constraint, representing promising candidates when designing a highly efficient subunit-based vaccine against *P. vivax*.

Key words: *Plasmodium vivax*, rhoptry, genetic diversity, tandem repeats, *pvron4*, natural selection, functional restriction.

Contenido

Pág.

Resumen ..... **IX**

Lista de figuras ..... **XV**

Lista de tablas ..... **XVII**

Lista de símbolos y abreviaturas ..... **XVIII**

Introducción **1**

1. Planteamiento del problema **5**

2. Justificación **7**

3. Hipótesis **9**

4. Objetivos **11**
4.1 Objetivo general 11
4.2 Objetivos específicos 11

5. Marco teórico **13**
5.1 Ciclo de vida del parásito 15
5.2 Mecanismo molecular de invasión al eritrocito 17
5.3 *Tight Junction* 19
5.4 RON4 en *Plasmodium vivax* (*Pv*RON4) 21

6. Metodología **23**
6.1 Tipo de investigación 23
6.2 Población de estudio 23
6.3 Verificación de la integridad del ADN 23
6.4 Genotipificación de las muestras por PCR-RFLPs 24
6.5 Diseño de cebadores 24
6.6 Amplificación por PCR 24
6.7 Ligación, transformación y clonación del fragmento purificado 25
6.8 Extracción del ADN plasmídico y secuenciación 25
6.9 Análisis de la diversidad genética y el efecto de la selección natural sobre el locus *pvron-4* 26
6.10 Análisis de selección positiva episódica sobre el locus *pvron-4* 27
6.11 Uso efectivo de codones 28
6.12 Desequilibrio de ligamiento y recombinación 28

6.13 Diferenciación y estructura genética poblacional del locus *pvron4* 28
6.14 Predicción de dominios putativos y del potencial antigénico de *pvron4* 29

7. Resultados 31
7.1 Análisis y genotipificación de las muestras 31
7.2 Diversidad genética en el locus *pvron4* 32
7.3 Efecto de la selección en el locus *pvron4* 36
7.4 Uso efectivo de codones 38
7.5 Desequilibrio de ligamiento y recombinación 39
7.6 Estructura genética de la población Colombiana respecto al locus *pvron4* 39
7.7 Predicción de dominios putativos y del potencial antigénico de *pvron4* 41

8. Discusión 45

9. Conclusiones y recomendaciones 51
9.1 Conclusiones 51
9.2 Recomendaciones 51

10. Bibliografía 53

11. Anexos 61
11.1 Anexo 1. Variación en el número de repeticiones en tándem en el locus *pvron4*.....61
11.2 Anexo 2. Alineamiento de los 32 haplotipos identificados para *pvron4*.....63
11.3 Anexo 3. Publicación 87

Lista de figuras

Pág.

Figura 5-1. Proporción de casos de malaria por *P. vivax* 14
Figura 5-2. Porcentaje de casos producidos por *P. falciparum* y *P. vivax* en el 2015. 15
Figura 5-3. Ciclo de vida de *Plasmodium* spp**.** 16
Figura 5-4. Esquema del merozoito de *P. vivax*..... 17
Figura 5-5. Modelo de la secuencia de interacciones ligando-receptor durante la invasión de *P. falciparum* al eritrocito 19
Figura 5-6. Modelo del *Tight Junction* propuesto en *Toxoplasma gondii*. 20
Figura 5-7. Esquema del gen que codifica la proteína *Pv*RON4 en la cepa VCG-I..... 21
Figura 7-8. Productos amplificados del gen *pvmsp3-α*. 31
Figura 7-9. Productos de la digestión enzimatica del gen *pvmsp3α* tratado con las enzimas *Alu-I* y *Hha-I*..... 32
Figura 7-10. Productos de amplificación del gen *pvron4*. 32
Figura 7-11. Productos purificados del gen *pvron4*..... 33
Figura 7-12. Representación esquemática del CDS de *pvron4* y ventana deslizante para la tasa ω. 35
Figura 7-13. Secuencia de aminoácidos de los Haplotipos 6 y 17 del locus *pvron4*..... 35
Figura 7-14. Linajes sujetos a selección diversificante episódica en *pvron4*..... 38
Figura 7-15. Regresión lineal entre el desequilibrio de ligamiento y la distancia nucleotídica. ... 39
Figura 7-16. *Median–joining network* por regiones colombianas**.**..... 41
Figura 7-17. *Median –joining network* para las poblaciones parasitarias por departamentos..... 42
Figura 7-18. Regiones potencialmente antigénicas de *Pv*RON4. 43

Lista de tablas

Pág.

Tabla 1. Estimadores de diversidad genética calculados a partir del polimorfismo de nucleótido simple (SNP) y la longitud de secuencia en *pvron4*. 36
Tabla 2. Pruebas de neutralidad, desequilibrio de ligamiento y recombinación para el gen *pvron4* en la población colombiana de *P. vivax*..... 36
Tabla 3. Evaluación de la sustitución intra-especie (d_N-d_S), divergencia (K_N-K_S) e índice de neutralidad de la prueba MK. 37
Tabla 4. Estadístico F_{ST} inter-poblaciones para *pvron4*..... 40
Tabla 5. Estadístico F_{ST} interpoblacional para *pvron4* 40

Lista de símbolos y abreviaturas

Símbolo	Término
$\mathbf{d_N}$	Tasa de substitución no-sinónima por sitio no-sinónimo
$\mathbf{d_S}$	Tasa de substitución sinónima por sitio sinónimo
$\mathbf{F_{SC}}$	Estadístico F, usado para estimar la proporción de la variabilidad genética encontrada entre poblaciones dentro de grupos
$\mathbf{F_{ST}}$	Índice de Fijación de Wright usado para estimar la proporción de variabilidad genética encontrada entre poblaciones
H	Número de haplotipos
Hd	Diversidad haplotípica
$\mathbf{K_N}$	Número promedio de divergencias no-sinónimas por sitio no-sinónimo
$\mathbf{K_S}$	Número promedio de divergencias sinónimas por sitio sinónimo
N'c	Numero efectivo de codones (ENC) prima
Ps	Número de sitios parsimoniosos informativos
Rm	Número mínimo de eventos de recombinación
R^2	Desequilibrio de ligamiento
S	Número de sitios singleton
Ss	Número de sitios segregantes
Za	Promedio de R^2 sobre todas las comparaciones entre sitios polimórficos adyacentes
ZnS	Promedio de R^2 sobre todas las comparaciones
ZZ	Za-ZnS. Estadístico usado para detectar recombinación intragénica
$\boldsymbol{\theta_w}$	Estimador de diversidad de Watterson
$\boldsymbol{\pi}$	Diversidad nucleotídica por sitio

Abreviatura	Término
AMA1	Antígeno apical de membrana-1 (*Apical membrane antigen 1*)
AMOVA	Análisis de varianza molecular
BepiPred	Predictor de Epítopes B
CBI	Índice de sesgo de codón (*Codon Bias Index*)
CSP	Proteína Circumsporozoito del esporozoito

ENC	Número efectivo de codones
FEL	Test de Probabilidad de efectos fijos (*Fixed effects likelihood*)
FUBAR	Test de Aproximación bayesiana rápida sin restricciones para inferir la selección (*Fast, Unconstrained Bayesian AppRoximation for inferring selection*)
GARD	Algoritmo genético para la detección de recombinación (*Genetic Algorithm Recombination Detection*)
IEDB	*Immune epitope database and analysis resource*
IFEL	Test de probabilidad de efectos fijos internos (*Internal fixed effects likelihood*)
InDels	Inserciones /Deleciones
LD	desequilibrio de ligamiento
MEME	Modelo de efectos mixtos de evolución (*Mixed effects model of evolution*)
MJ	Unión en movimiento (*Moving junction*)
NCBI	*National Center for Biotechnology Information*
PCR-RFLP	Reacción en cadena de la polimerasa-Polimorfismo de longitud de fragmentos de restricción
***pfron4*:**	Gen que codifica para la proteína del cuello de las roptrias 4 en *Plasmodium falciparum*
pvmsp-3α	Gen que codifica para la subunidad alfa de la proteína de superficie del merozoito 3 en *Plasmodium vivax*
pvrap1	Gen que codifica para la proteína asociada a roptrias 1 en *Plasmodium vivax*
pvrap2	Gen que codifica para la proteína asociada a roptrias 2 en *Plasmodium vivax*
pvron4	Gen que codifica para la proteína del cuello de las roptrias 4 en *Plasmodium vivax*
***Pv*RON4**	Proteína del cuello de roptrias 4 de *Plasmodium vivax*
RDP	Programa para la detección de recombinación
REL	Test de Probabilidad de efectos aleatorios (*Random effects likelihood*)
RON2	Proteína del cuello de roptrias 2
RONs	Proteínas del cuello de las roptrias
Sal-I	Salvador –I
SLAC	Test de probabilidad individual de conteo ancestro (*Single likelihood ancestor counting*)
SNP	Polimorfismo de nucleótido simple (*Single Nucleotide Polymorphism*)
***Tg*RON4**	Proteína del cuello de las roptrias 4 en *Toxoplasma gondii*
VCG-1	Cepa de laboratorio de *P. vivax* del departamento del Guaviare, Colombia (Vivax Colombia Guaviare-I)

Introducción

La Malaria es una de las enfermedades parasitarias de alto impacto en la salud pública alrededor del mundo. Esta enfermedad es causada por *Apicomplexos* del género *Plasmodium*, un grupo de parásitos intracelulares obligados que invaden células tanto del hígado (hepatocitos) como de la sangre (eritrocitos y/o reticulocitos), transmitiéndose a través de su hospedero definitivo, el mosquito hembra del genero *Anopheles*. Estos parásitos se encuentran ampliamente distribuidos sobre las regiones tropicales y subtropicales del mundo [1-3]. Los síntomas clínicos de la enfermedad abarcan desde dolor de cabeza, fiebre, escalofríos y anemia, hasta la muerte a causa de la malaria cerebral o falla renal secundaria a la enfermedad.

Anualmente, se reportan cerca de 300 millones de casos clínicos y más de 400 mil muertes por malaria alrededor del mundo [4], siendo *Plasmodium falciparum* el responsable de la gran mayoría de estas muertes, afectando principalmente a la población africana. De las cinco especies capaces de causar malaria en humanos (*P. falciparum*, *P. vivax*, *P. ovale*, *P. malariae* y *P. knowlesi*), esta especie es la más patogénica, razón por la cual se ha convertido en el principal blanco de una vacuna. Sin embargo, los crecientes reportes de malaria severa a causa de *Plasmodium vivax*, las frecuentes recaídas de la enfermedad en pacientes tratados y la generación de resistencia a los fármacos, convierten a esta especie en un problema de salud pública creciente [5, 6].

Plasmodium vivax es la especie de mayor distribución global [2, 7-9], ya que el mosquito transmisor es capaz de resistir climas más fríos, con alta prevalencia en los continentes de Asia y América del Sur; cerca de 2.500 millones de individuos permanecen en riesgo de infección por las especies más prevalentes de malaria (*P. falciparum* y *P. vivax*) [4, 8, 10]. La capacidad que tiene *P. vivax* de permanecer como formas latentes en el hígado (hipnozoitos), así como la producción de gametos infectivos para nuevas hembras de *Anopheles*, aún antes de causar síntomas, dificultan enormemente el control de la enfermedad causada por esta especie de parásito. Es por esto que la investigación y el desarrollo de nuevas estrategias de control, como lo es el desarrollo una vacuna antimalárica, son necesarias para disminuir su incidencia y asegurar su posterior erradicación.

Uno de los mayores obstáculos para el desarrollo de una vacuna antimalárica eficaz, es el polimorfismo de los antígenos del parásito, el cual se asocia con la evasión a la respuesta inmune del hospedero [8, 10, 11]. Aunque un importante número de antígenos de *P. vivax* ha sido caracterizado en los diferentes estadios de su ciclo de vida, sólo algunos de ellos presentan características que los convierten en buenos candidatos a vacuna. Para la fase sanguínea de la infección, un potencial candidato debe ser una proteína que sea expresada en las últimas horas del ciclo intraeritrocítico, necesaria o indispensable en la invasión a las células hospederas y capaz de despertar una respuesta inmune, comprobada bajo condiciones naturales o experimentales. Estructuralmente, dichas proteínas suelen tener una secuencia señal, dominios transmembranales o anclaje GPI y regiones o dominios conservados [12].

A la fecha, la mayoría de los antígenos caracterizados corresponden a proteínas del estadio infectivo en sangre o merozoito [12]. La alta homología entre proteínas de *P. falciparum* y *P. viva*x, así como la secuencia de eventos similar en los mecanismos de invasión de estas especies, con otros *Apicomplexa* como *T. gondii*, ha permitido identificar potenciales candidatos y caracterizar procesos importantes como el *tight junction* (TJ o unión estrecha), un paso decisivo en el éxito de la invasión. El TJ es una estructura proteica formada entre el antígeno de micronemas AMA1 y proteínas del cuello de roptrias (RONs). Esta estructura permite la formación de un anillo conectivo a través del cual el parásito entra a la célula huésped impulsado por su motor actina-miosina [13, 14]. El complejo de proteínas RON (caracterizado en *P. falciparum*) está constituido por las proteínas RON2, RON4 y RON5, de las cuales se conoce que RON2 (expuesta en la membrana de la célula huésped) une directamente a AMA1, la cual está anclada a la membrana del parásito. Aunque el mecanismo en cuanto a función e interacción entre las demás proteínas el complejo no es claro, éstas son consideradas importantes blancos para bloquear la invasión. Sin embargo, el conocimiento actual sobre su potencial como candidatos a vacuna es insuficiente.

Estudios de co-localización y modelos de invasión descritos para *Plasmodium* spp y *T. gondii* han conducido a establecer la participación convincente de RON4 en el TJ [13, 15, 16]. Así mismo, su expresión en las formas invasivas del parásito (el esporozoito y el merozoito) [14, 17-19], la habilidad para provocar una respuesta inmune en infecciones naturales [19] y la naturaleza conservada de la proteína, específicamente hacia la región C-terminal (inferida por análisis comparativos entre secuencias de aminoácidos de *Pf*RON4 y *Tg*RON4) [20], sugieren que esta proteína juega un papel importante para el parásito y podría así ser evaluada como candidato a vacuna. Recientemente, el ortólogo de la proteína *Pf*RON4 fue caracterizado en la cepa VCG-1 de

Plasmodium vivax [18]. En esta especie, RON4 (*Pv*RON4) es codificada por un gen de aproximadamente de 2.657 pares de bases (pb), expresado durante las últimas horas del ciclo intra-eritrocítico y es secretada desde el cuello de las roptrias. Esta proteína contiene un péptido señal, un dominio de baja complejidad conformado por dos tipos de repeticiones en tándem, un dominio hélice alfa de doble espiral y cinco cisteínas conservadas hacia el extremo C-terminal [18]; esta última región parece estar altamente conservada entre *P. vivax* y especies del parásito que infectan primates del viejo mundo [21].

Teniendo en cuenta la participación potencial de RON4 en la invasión y dado que la diversidad genética de los antígenos de parásito es un obstáculo para el diseño de una vacuna completamente eficaz contra *P. vivax*, en este estudio se caracterizó la variabilidad genética del locus *pvron4* y los procesos evolutivos que generan su patrón de variación, con el fin de determinar dominios conservados que puedan ser incluidos en el desarrollo de una vacuna multiantígeno contra la malaria producida por *P. vivax*.

1. Planteamiento del problema

La diversidad genética de los organismos les permite responder a cambios en el medio ambiente en el cual estos se desarrollan. Por lo tanto, esta diversidad puede permitir la adaptación bajo ciertas circunstancias. En Malaria, la diversidad genética de los antígenos parasitarios de *P. falciparum* ha generado una disminución en la eficacia de las vacunas desarrolladas. Este alto polimorfismo suele ser consecuencia de presiones selectivas impuestas por el sistema inmune del hospedero, convirtiendo la variabilidad genética en un mecanismo de evasión, debido a que variantes de un determinado antígeno provocarán respuestas inmunes alelo–específicas, que permiten el escape del parásito a las defensas del individuo infectado. Por lo tanto, este polimorfismo es uno de los principales obstáculos en el desarrollo de vacunas [12]. Ya que *P. vivax* presenta casi el doble de diversidad genética con respecto a *P. falciparum*, el desarrollo de una vacuna contra este parásito podría encontrar un obstáculo aún mayor. De acuerdo a lo anterior, una vacuna efectiva contra *P. vivax* basada en subunidades dependería de la inclusión de regiones o dominios conservados de los antígenos candidatos, lo que hace necesario evaluar la diversidad genética de los potenciales antígenos candidatos. Considerando que RON4 es un componente indispensable del TJ [14, 16] y un potencial candidato a vacuna contra *P. vivax* [15, 17], el presente estudio pretende describir la variabilidad genética de este locus y las fuerzas evolutivas que han modulado su variación, los cuales, a la fecha son desconocidos.

2. Justificación

Plasmodium vivax es la especie causante de malaria más ampliamente distribuida en el mundo. El estudio de la biología funcional de este parásito se ha visto limitado por su preferencia por infectar reticulocitos o glóbulos rojos inmaduros, los que se hallan en baja proporción en sangre. Esta preferencia por células inmaduras impide el mantenimiento de un cultivo *in vitro* y por ende la caracterización funcional de la mayoría de los antígenos parasitarios. Diversos enfoques computacionales han sido efectuados para superar este obstáculo y así evaluar aspectos de la biología, comportamiento poblacional, distribución y resistencia de esta especie parasitaria.

Dentro de las principales características que hacen de esta especie una gran amenaza para la salud pública se encuentran: su capacidad para permanecer en formas latentes en las células del hígado, la producción de estadios sexuales en corto tiempo y la reciente detección de cepas resistentes a los fármacos antimaláricos (resistencia que ha sido atribuida a la acumulación de polimorfismos en su genoma). En consecuencia, el aumento de la tasa de morbilidad, las frecuentes recaídas en pacientes tratados y el reporte de casos de malaria por *P. vivax* cada vez más complicados [6, 22], han hecho necesario ahondar esfuerzos e incrementar la investigación en esta especie, con el fin de diseñar una vacuna antimalárica completamente efectiva.

Dado que las manifestaciones clínicas de la malaria ocurren en el estadio intra-eritrocítico (o estadio sanguíneo de la infección), parte del desarrollo de vacunas basadas en subunidades contra estos parásitos se ha enfocado en las formas parasitarias que infectan los glóbulos rojos, denominadas merozoitos [23]. Es así como antígenos del merozoito están siendo caracterizados, con el fin de entender su modo de acción e inducir una respuesta inmune contra ellos, lo que evitaría la interacción parásito-hospedero, reduciendo así la carga del parásito en el individuo y llevando a un descenso en la transmisión de la enfermedad.

Sin embargo, para inducir una respuesta inmune efectiva, es necesario dirigirla hacia las regiones más conservadas del antígeno, para evitar así respuestas alelo-específicas. En este sentido, los

candidatos potenciales a vacuna deben ser aquellos que, además de cumplir un rol esencial en la invasión, posean regiones o dominios conservados. Recientemente, varios complejos proteicos que median el proceso de invasión del merozoito a la célula hospedera han sido descritos en otros *Apicomplexa* como *Toxoplasma gondii* y *P. falciparum*. Ortólogos de las proteínas que participan en estas interacciones han sido descritos en *P. vivax* [23], dentro de los cuales se encuentra la proteína del cuello de las roptrias 4 (RONs) [15, 17]. La caracterización de esta proteína ha permitido identificar su participación convincente en la formación del TJ [15] y ensayos de antigenicidad bajo infecciones naturales han demostrado que puede inducir una respuesta inmune [18]. Sin embargo, la variabilidad genética de esta proteína se desconoce, razón por la cual es necesario evaluar su variabilidad genética y determinar su potencial como blanco vacunal.

3. Hipótesis

La proteína del cuello de roptrias 4 de *Plasmodium vivax* (*Pv*RON4) posee dominios altamente conservados, que se han mantenido por selección natural.

4. Objetivos

4.1 Objetivo general

Evaluar la variabilidad genética del locus que codifica para la proteína del cuello de las roptrias 4 de *Plasmodium vivax* (*Pv*RON4) y las fuerzas evolutivas que generan el patrón de variación observado, a partir de aislados clínicos colombianos provenientes de las tres regiones de mayor transmisibilidad en Colombia, obtenidos entre los años 2007-2015.

4.2 Objetivos específicos

- Determinar la presencia de diferentes genotipos de *Plasmodium vivax* a partir de aislados clínicos colombianos provenientes de la Costa Pacífica (departamentos de Chocó y Nariño), Urabá/Bajo Cauca/Sur de Córdoba (departamentos de Córdoba y Antioquia), y la región de Orinoquia-Amazonia (departamentos de Amazonas, Guainía), entre los años 2007 y 2015.

- Obtener fragmentos de ADN del gen *pvron4* a partir de aislados clínicos colombianos colectados entre los años 2007 y 2015.

- Evaluar el grado de diferenciación genética en el locus *pvron4* entre las poblaciones parasitarias procedentes de la Costa Pacífica, Urabá/Bajo Cauca/Sur de Córdoba, y la región de Orinoquia-Amazonia.

- Determinar señales de selección natural en el locus *pvron4* a partir de las secuencias obtenidas de las poblaciones colombianas y a partir de aquellas disponibles en bases de datos.

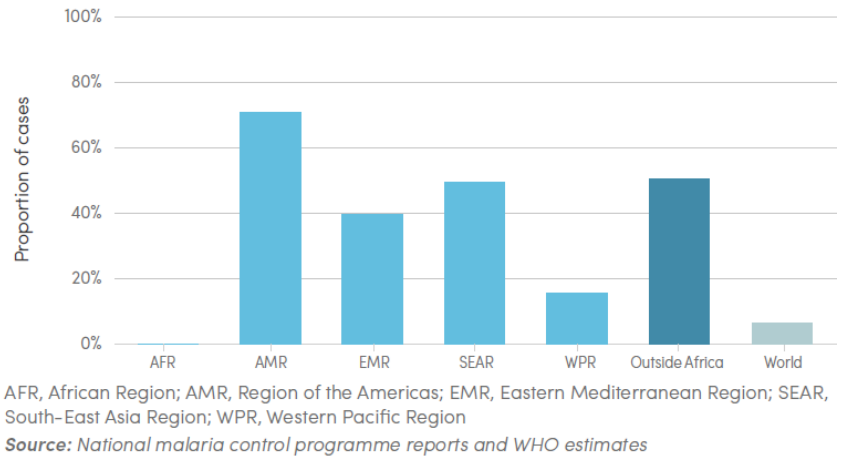
5. Marco teórico

Anualmente, las enfermedades parasitarias cobran un alto número de vidas humanas alrededor del mundo. Para el año 2013, se reportó más de un millón de muertes, de las cuales, más del 85% fueron causadas por parásitos de la malaria [24]. Esta es una enfermedad infecciosa causada por *Apicomplexas* del género *Plasmodium* que afecta a cerca de 97 países alrededor del mundo [3, 25, 26]. A pesar de los constantes esfuerzos por controlarla en regiones endémicas, esta enfermedad permanece como un problema de salud pública global. Durante el 2015, ésta llego a causar 300 millones de casos, con un promedio estimado de 438.000 muertes [3].

A la fecha, se han descrito cinco especies de *Plasmodium* capaces de infectar seres humanos (*P. falciparum, P. vivax, P. ovale, P. malariae* y *P. knowlesi*) [27], de las cuales *P. falciparum* y *P. vivax* representan la mayor amenaza. *P. falciparum* es el responsable de la mayoría (91%) de las muertes relacionadas con la malaria en África subsahariana, donde los niños menores de 5 años son la población más vulnerable [3, 24]. La severidad de la infección por esta especie, la ha clasificado como la especie más patogénica [28]. Por su parte, el vector transmisor de *P. vivax* tiene la característica de soportar climas más fríos, además, el parásito puede permanecer latente (como hipnozoitos) en las células del hígado y generar formas sexuales en corto tiempo, convirtiendo al huésped en un potencial transmisor, aún sin presentar los síntomas de la enfermedad. Todas estas características le han permitido a *P. vivax* tener una mayor distribución con respecto a otras especies, siendo ésta la responsable de altas tasas de morbilidad. Aunque puede ocurrir en África, la alta frecuencia de individuos negativos para el antígeno Duffy del glóbulo rojo, el cual media la invasión del parásito, hace que la mayoría de los individuos en este continente sean resistentes a la infección [3], por lo que es más prevalente fuera de África.

Para el 2015, la malaria causada por *P. vivax* ha sido responsable de cerca de 13.8 millones de casos y alrededor de 14.900 muertes, con alta prevalencia en el continente Americano, el sudeste asiático y la región del Mediterráneo Oriental (Figura 5-1). Según la Organización Mundial de la Salud (OMS) (2015), la etapa de hipnozoito es la principal causa de múltiples recaídas, aumentando la incidencia de *P. vivax*. Sumado a esto, las limitaciones en la adquisición del tratamiento con fármacos como la primaquina para controlarlo, así como la aparición de cepas resistentes a los medicamentos antimalaricos como la cloroquina, hacen que su incidencia decrezca más lentamente que la de *P. falciparum* en las regiones donde las dos especies coexisten. Así, *P. vivax* puede persistir como el principal causante de la malaria y representa el principal desafío para la eliminación de la malaria [29].

Figura 5-1. Proporción de casos de malaria por *P. vivax*



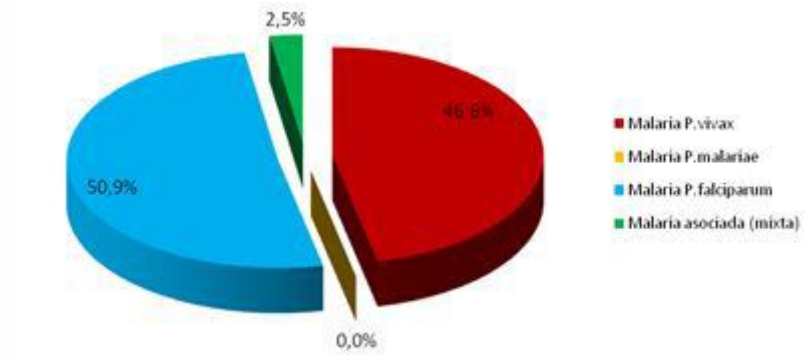
Tomado de [29]. Los casos fueron estimados en cada región durante el año 2015.

En América, el número de casos clínicos de infecciones por *Plasmodium* supera el millón, de los cuales, el 65% corresponden a *P. vivax*. En este continente, un tercio de la población se encuentra en riesgo y los países que aportan la mayoría de casos son Colombia y Brasil con el 68% [26], siendo parte de los 24 países en donde el tratamiento con cloroquina ha presentado falencias [29].

Según la OMS [30], en Colombia, el 22% de la población está en riesgo de contraer la enfermedad. Del total de casos de malaria reportados para el 2015 (50.243), el 46,6% correspondió a infecciones causadas por *P. vivax* y el 50,9% a *P. falciparum* [31] (Figura 5-2). Por otro lado, la infección por *P. malariae* ocurre en focos dispersos a lo largo de la costa Pacífica y no existe transmisión reportada de las dos especies restantes (*P. knowlesi* y *P. ovale*). A pesar de los reportes

oficiales anteriores, una muy alta incidencia de infección por *P. malariae* ha sido descrita recientemente en el Departamento del Amazonas [32]. Infecciones mixtas o infecciones simultáneas por dos especies, usualmente *P. vivax* y *P. falciparum*, pueden ocurrir en bajas proporciones [33]. Sin embargo, éstas representan un desafío para el diagnóstico y el tratamiento de la enfermedad, debido a que recientes estudios de genética poblacional en la región de Córdoba, un departamento endémico para la malaria en Colombia, han revelado la aparición de cepas de *P. vivax* resistentes a fármacos antimaláricos, que deberían ser usados exclusivamente en pacientes infectados con *P. falciparum* [34, 35].

Figura 5-2. Porcentaje de casos producidos por *P. falciparum* y *P. vivax* en el 2015.



Tomado de [31].

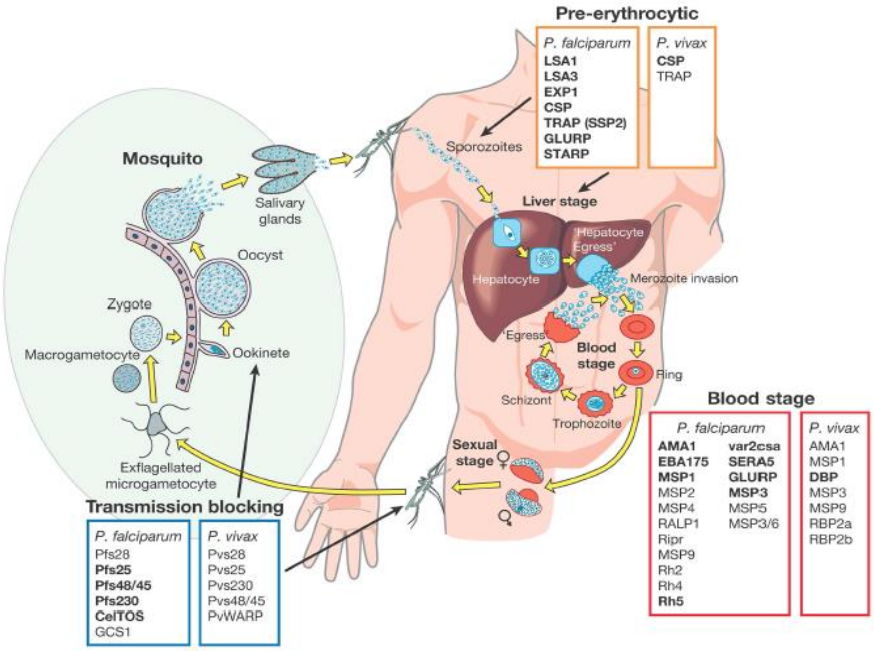
En determinadas regiones de Colombia, existe una tendencia a la transmisión de la malaria influenciada en gran medida por el fenómeno del Niño, que propicia las condiciones ecológicas para la infección, siendo tres los focos de mayor transmisibilidad: 1) el foco de Urabá - Bajo Cauca – Sur de Córdoba; 2) el foco de la Costa Pacífica y 3) el foco de la Orinoquia-Amazonia [25]. Esta transmisión focalizada y la variabilidad de la incidencia de una región a otra, durante épocas y entre años, son las razones por las cuales la malaria en Colombia se considera de transmisión inestable [25].

5.1 Ciclo de vida del parásito

Plasmodium es un género de parásitos intracelulares obligados, clasificados dentro del filo *Apicomplexa* y el orden *Haemosporidia.* Estos protozoarios se caracterizan por presentar un conjunto de orgánulos en su extremo apical conocidos como roptrias, micronemas y gránulos densos, los cuales liberan su contenido (proteico y también lipídico) durante la invasión a las células hospederas [23, 36].

Para completar su ciclo de vida, *Plasmodium* spp. requiere de dos hospederos (Figura 5-3); un hospedero invertebrado que es el mosquito hembra del género *Anopheles*, en el cual se lleva a cabo la fase sexual, y un hospedero vertebrado, particularmente mamífero, en el que se lleva a cabo la fase asexual. Durante estas fases, el parásito expone diferentes estadios de desarrollo que implican cambios morfológicos y distintos patrones de expresión de proteínas necesarios para mediar la interacción y la invasión a las diferentes células del hospedero [37, 38]. De esta manera, la fase asexual se puede dividir en exo-eritrocítica e intra-eritrocítica [36].

Figura 5-3. Ciclo de vida de *Plasmodium* spp**.**



Tomado de [12]. En cada estadio parasitario han sido descritos varios antígenos involucrados en el proceso de invasión. Estos son blancos potenciales para el desarrollo de una vacuna contra *P. falciparum* o *P. vivax*.

La etapa exo-eritrocitica inicia con la picadura de un mosquito hembra a un hospedero sano. En ese momento, el mosquito trasmite esporozoitos desde sus glándulas salivales al torrente sanguíneo del hospedero, a través de su probóscide [39]. Los esporozoitos se transportan rápidamente al hígado, donde pueden llegar a invadir de 100-300 hepatocitos [28]. Una vez en el hepatocito, se desarrolla un esquizonte tisular que se reproduce asexualmente, generando alrededor de 30.000 merozoitos. Los hepatocitos infectados con las formas maduras de esquizontes mueren y se rompen, liberando los merosomas (sacos en los que van contenidos miles de merozoitos). Posteriormente, cada merozoito puede llegar a invadir un eritrocito o un reticulocito, siendo este último la célula diana

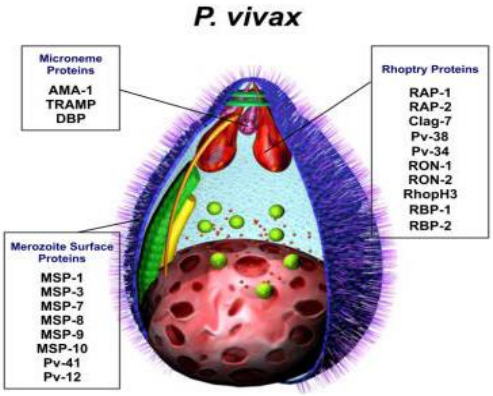
de *P. vivax* [12, 40]. Tanto *P. vivax* como *P. ovale* pueden residir en el tejido hepático por un largo periodo de tiempo (meses) en un estadio conocido como hipnozoito [39].

En la etapa intra-eritrocítica, el merozoito se internaliza y desarrolla dentro de los glóbulos rojos. Allí, el parásito atraviesa por los estadios de anillo, trofozoito y esquizonte. Tras un par de días desde la invasión, el esquizonte madura y el eritrocito infectado se rompe liberando en promedio 16 nuevos merozoitos, que invadirán nuevos eritrocitos. Esta etapa es considerada la más crítica, debido a que el parásito entra en un ciclo de desarrollo, liberación e invasión, causando todas las manifestaciones clínicas de la enfermedad [39, 40]. Alternativamente, durante la etapa intra-eritrocítica, el merozoito se puede diferenciar en gametocito masculino o femenino, que posteriormente serán tomados por el mosquito durante la ingesta de sangre. Una vez en el intestino del insecto, los gametocitos se fusionan para formar el cigoto y éste a su vez, se diferencia en un ooquinete invasivo, que atraviesa el intestino medio y se transforma en un ooquiste, a partir del cual se liberan nuevos esporozoitos. Finalmente, estos esporozoitos migran a través de la hemolinfa hacia las glándulas salivales del mosquito, donde maduran y se almacenan hasta su transmisión a un hospedero sano en la siguiente ingesta de alimento [12, 36, 40].

5.2 Mecanismo molecular de invasión al eritrocito

El estadio parasitario que se desarrolla dentro de los eritrocitos es el merozoito (Figura 5-4). Al igual que el esporozoito (forma invasiva de los hepatocitos), este estadio posee orgánulos apicales conocidos como roptrias, micronemas y gránulos densos, los cuales secretan sus contenidos durante el proceso de invasión [23].

Figura 5-4. Esquema del merozoito de *P. vivax*.



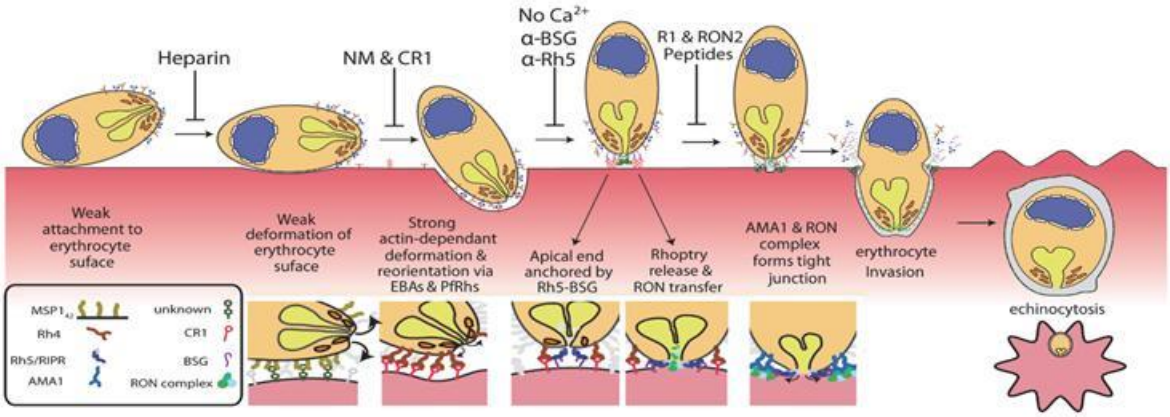
Tomado de [23]. En el esquema se observan algunas de las proteínas involucradas en el proceso de invasión estudiadas como potenciales candidatos a vacuna.

La exposición del merozoito al sistema inmune durante su recorrido por el torrente sanguíneo, a causa del repetitivo ciclo de invasión a los eritrocitos, es la principal razón por la que la gran mayoría de los candidatos a vacuna descritos pertenecen a este estadio. Es así como proteínas del merozoito, tanto de superficie, como aquellas secretadas desde sus organelos, son evaluadas como potenciales candidatos a vacuna, debido a que las respuesta inmune generada hacia ellas, podría inhibir la invasión [12].

Varios modelos, tanto para *T. gondii* como para *P. falciparum*, han sido descritos para explicar el proceso de invasión. Recientemente, Weiss y cols., [13] propusieron una secuencia de interacciones receptor-ligando que ocurren en *P. falciparum* durante su proceso de invasión al eritrocito, la cual se ha considerado como un modelo general para este género. En resumen, este proceso involucra tres pasos importantes: 1) un contacto inicial, 2) la reorientación del extremo apical del merozoito hacia la superficie de la célula diana y 3) la internalización del merozoito dentro de una vacuola parasitófora hacia el interior del eritrocito (Figura 5-5).

El contacto inicial entre el parásito y el eritrocito está mediado por una familia de proteínas de superficie denominadas MSPs (del inglés, *Merozoite Surface Proteins*) que reconocen receptores del eritrocito como Banda 3 y otros aún desconocidos, induciendo una deformación débil del eritrocito [14, 41]. Posteriormente, se establecen interacciones alternas de ligandos como las mediadas por las proteínas de la familia EBA (del inglés, *Erythrocyte Binding Antigen*) y las PfRh (del inglés, *Plasmodium falciparum Reticulocyte binding-like proteins*), que se unen a receptores específicos como la glicoforina y la basigina, respectivamente, induciendo una deformación fuerte del eritrocito y la reorientación del parásito [14]. Pese a que estas proteínas son funcionalmente redundantes, la proteína Rh5, perteneciente a la familia de las PfRhs, parece desempeñar un papel importante en la secreción de los contenidos de los micronemas y las roptrias. Aunque no se tiene total claridad respecto a su función, se plantea que la interacción con basigina, media la formación de una apertura o poro por donde se secretan los contenidos de los orgánulos apicales hacia el interior de la célula huésped, permitiendo la internalización del parásito [13]. En *P. vivax*, las proteínas que se encuentran mediando este paso son las RBPs (del inglés, *Reticulocyte binding proteins*) y la proteína de unión a Duffy o DBP (del inglés, *Duffy binding protein*).

Figura 5-5. Modelo de la secuencia de interacciones ligando-receptor durante la invasión de *P. falciparum* al eritrocito



Tomado de [13].

Dentro de las proteínas secretadas por los orgánulos apicales del parásito, se encuentra la proteína de micronemas AMA1 (del inglés, *Apical Membrane Antigen-1*) y las proteínas del cuello de roptrias o RONs. La interacción entre AMA1 y el complejo de proteínas RONs constituyen un anillo conectivo entre el merozoito y el eritrocito denominado *Tight Junction* (TJ) o *Moving Junction* (MJ), a través del cual, el merozoito ingresa al eritrocito impulsado por el motor actina-miosina [14, 42]. El TJ es una estructura importante en la formación de la vacuola parasitófora e induce el cierre de la misma, constituyendo un paso decisivo para una invasión exitosa [13].

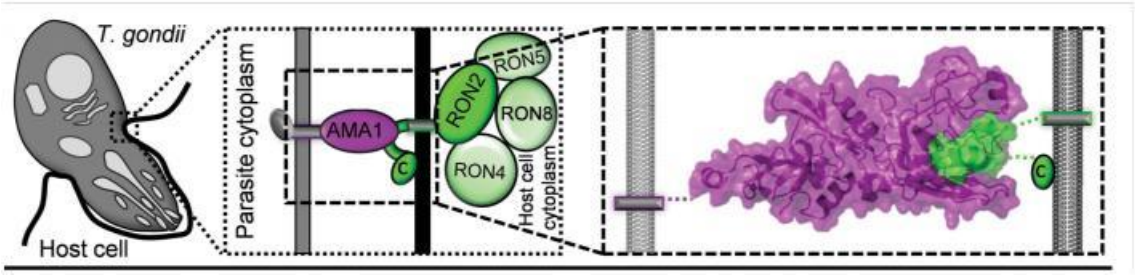
5.3 *Tight Junction*

La proteína AMA1 es una proteína antigénica, considerada un potencial candidato a vacuna debido al papel que cumple en las diferentes especies de *Plasmodium* durante el proceso de invasión y dado que se expresa en dos estadios críticos del ciclo de vida del parásito, el esporozoito y el merozoito [43, 44]. Este antígeno es una proteína de membrana integral constituida por una secuencia señal, un ectodominio N-terminal dividido en tres dominios estructurales (DI, DII, DIII) definidos por ocho enlaces disulfuro [45] y una hélice transmembranal, seguida por un pequeño dominio citoplásmico C-terminal. Los dominios DI y DII son estructuralmente similares entre sí y cumplen diversas funciones de adhesión [14, 22].

El complejo de proteínas RON, caracterizado en Apicomplexas como *T. gondii* y *P. falciparum*, está constituido por las proteínas RON2/4/5/8. Una vez secretadas desde las roptrias, éstas se insertan dentro de la membrana de la célula hospedera. RON4/5/8 forman un complejo que soporta

a RON2, la única proteína con al menos una hélice transmembranal, para facilitar la exposición de un gran ectodominio hacia el exterior del eritrocito, el cual actúa como receptor para AMA1 [42, 46]. Tras la secreción desde los micronemas, la proteína AMA1 se ancla a la membrana del merozoito, exponiendo el ectodominio hacia el exterior celular, donde el dominio I y II de AMA1 conforman el bolsillo de unión a RON2 [13, 14, 16, 18, 36, 42, 47, 48]. De esta manera, se establece el TJ y los parásitos *Plasmodium* y *Toxoplasma* codifican su propio ligando y receptor (Figura 5-6).

Figura 5-6. Modelo del *Tight Junction* propuesto en *Toxoplasma gondii*.



La proteína AMA1 se ancla a la membrana del parásito y establece una unión con la región próxima al C-terminal de RON2, que forma un bucle. El complejo RON2/4/5/8 se encuentra insertado dentro de la membrana de la célula hospedera [42].

Aunque no se tiene claridad acerca de la interacción entre las proteínas del complejo RON, se sabe que la proteína RON8 pueden no ser indispensable en especies de *Plasmodium* [14]. Por el contrario, RON4 resulta ser la única proteína que ha sido claramente identificada y localizada en la interacción con AMA1 después de RON2, tanto en *T. gondii* como en especies de *Plasmodium* [16, 49, 50]. El análisis de las secuencias de aminoácidos de RON4 en estos *Apicomplexa* indica que esta proteína posee dominios altamente conservados en su extremo C-terminal, que pueden estar asociados a una restricción funcional, importante en términos evolutivos [20].

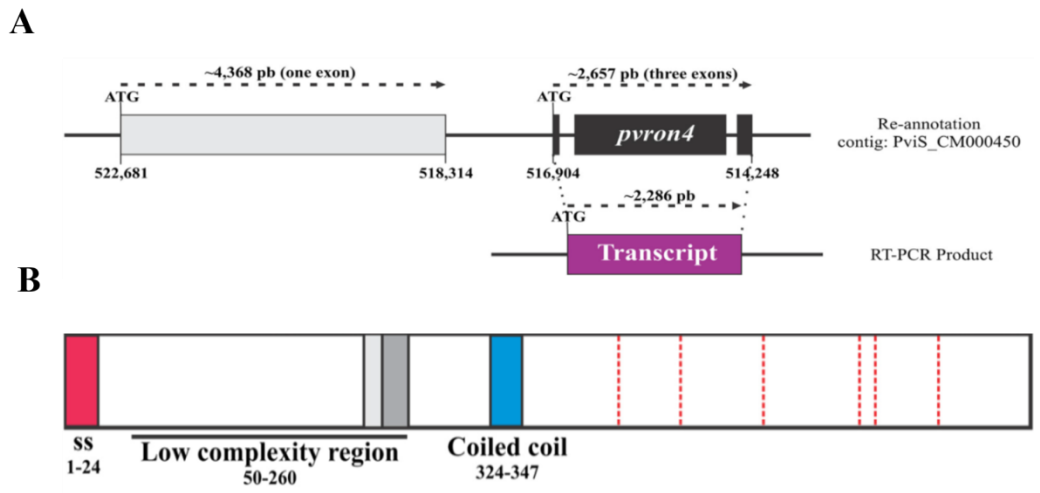
Luego de AMA1-RON2, RON4 es el segundo antígeno más estudiados del complejo [15]. Su importancia como potencial candidato a vacuna radica en los siguientes aspectos: es esencial en la invasión por merozoitos y esporozoitos a sus respectivas células diana [16, 49, 50], estudios de antigenicidad han demostrado que la proteína *Pf*RON4 induce respuesta inmune en infecciones naturales, lo que sugiere que en algún momento de la invasión, es detectable por el sistema inmune [18, 51].

5.4 RON4 en *Plasmodium vivax* (*Pv*RON4)

En *Plasmodium vivax*, *Pv*RON4 ha sido caracterizada a partir de la cepa Vivax Colombia Guaviare-1 (VCG-I) como una proteína de 733 aminoácidos y una masa molecular de ~80 kDa. Está constituida por una secuencia señal, un motivo α-helicoidal de doble espiral, una región variable de baja complejidad que contiene dos repeticiones en tándem (TR), y seis cisteínas conservadas hacia el carboxi-terminal (Figura 5-7). Carece de dominios transmembranales y anclaje GPI, lo que sugiere que se trata de una proteína soluble que se expresa en el extremo apical de los esquizontes maduros [18].

El gen *pvron4* que codifica para la proteína *Pv*RON4, posee una longitud aproximada de 2.657 pb y está constituido por 3 exones y 2 intrones. El exón 2 varía en longitud entre las cepas Salvador-1 (Sal-I) y VCG-I, debido a que codifica para la región de baja complejidad. Estas regiones variables son específicas de cada especie e incluso de diferentes cepas del parásito [52]. De acuerdo con Grau & Behr [53], estas regiones pueden actuar como blancos para anticuerpos antígeno-específicos, y a menudo funcionan como mitógenos de células B y células T gama-delta. Tienden a ser regiones altamente inmunogénicas y, por lo tanto, generan una "pantalla de humo inmunológica", ya que inducen respuestas inmunitarias ineficaces [53, 54].

Figura 5-7. Esquema del gen que codifica la proteína *Pv*RON4 en la cepa VCG-I



A. Estructura del gen *pvron4* obtenida tras la re-anotación del *contig* PviS_CM000450. B. Representación esquemática de la proteína *Pv*RON4, constituida por un péptido señal (rojo), una región α-helicoidal de doble espiral (azul) y una región de repeticiones en tándem (gris). Las líneas rojas punteadas indican las cisteínas conservadas entre las secuencias consenso [18].

En un análisis previo, en el que se evaluaron las señales de selección natural en 59 antígenos de *P. vivax*, utilizando únicamente 5 secuencias de referencia [21], se encontró que *pvron4* está dentro de los 8 genes con menor diversidad genética. Adicionalmente, este gen posee una región variable (dentro de la especie) y divergente (entre especies) entre los residuos 255 a 292, lo que sugiere que es blanco del sistema inmune. Por otro lado, posee un dominio altamente conservado entre especies hacia el extremo C-terminal, que podría indicar una restricción funcional [21].

6. Metodología

6.1 Tipo de investigación

Este es un trabajo descriptivo, en el que se evaluó la variabilidad genética del locus *pvron4* que codifica para la proteína de cuello de roptrias 4 de *Plasmodium vivax* (*Pv*RON4).

6.2 Población de estudio

El ADN genómico de *P. vivax* fue obtenido a partir de 73 muestras de sangre colectadas entre los años 2007 a 2015, provenientes de la Costa Pacífica (departamentos de Chocó y Nariño), Urabá/Bajo Cauca/Sur de Córdoba (departamentos de Córdoba y Antioquia), y la región de Orinoquia-Amazonia (departamentos de Amazonas, Guainía), regiones con la mayor transmisibilidad de malaria por *P. vivax* en Colombia [25]. Los pacientes fueron notificados del propósito de la investigación y dieron su consentimiento informado. Todos los procedimientos fueron aprobados por el Comité de Ética de la Fundación Instituto de Inmunología de Colombia y por el Comité de Ética de la Universidad del Rosario.

6.3 Verificación de la integridad del ADN

La presencia e integridad del ADN parasitario se verificó través de la amplificación por PCR de la subunidad ribosomal 18S de *P. vivax*. Los productos de PCR fueron evaluados por electroforesis en gel de agarosa al 1,5%, con el marcador de peso Hyperladder I (Bioline Inc. USA).

6.4 Genotipificación de las muestras por PCR-RFLPs

A partir de 100 muestras de ADN parasitario obtenidas de 3 diferentes regiones de Colombia, se seleccionaron aquellas que presentaron infección por cepa única. Para esto, se llevó a cabo la amplificación por PCR del marcador *pvmsp-3α,* un gen altamente polimórfico. Muestras que presentaron más de una banda de amplificación fueron excluidas, mientras que aquellas que presentaban una sola banda fueron tipificadas por PCR-RFLPs. Con el fin de evaluar la presencia de genotipos distintos del parasito en las poblaciones estudiadas, la tipificación de las muestras fue llevada a cabo a partir de la digestión del fragmento *pvmsp-3α* con las enzimas de restricción *Alu-I* y *Hha-I* (BioLabs Inc. New Egland). Los resultados fueron analizados por electroforesis en gel de agarosa al 3%, utilizando el marcador de peso Hyperladder V (Bioline Inc. USA).

6.5 Diseño de cebadores

El diseño de los cebadores para la amplificación y la secuenciación del gen *pvron4* se realizó con el software Generunner v3.05 (Hastings software, Inc.) y el servidor en línea Oligoanalizer v3.1 (IDT® Integrated DNA Technologies). Para esto, se utilizó la secuencia genómica reportada de la cepa de referencia Sal-I para este gen (GenBank, número de acceso AAKM01000005.1). Las regiones seleccionadas como cebadores tenían un tamaño entre 18 y 25 pb, presentaron ΔG positivos al analizar la formación de *harpins*, con temperaturas de formación por debajo de 0 o no superior a 20 °C. La temperatura de anillamiento (Tm) de los cebadores de amplificación fue de 56.4 y 52.6 °C y la de los cebadores utilizados en la secuenciación fue de 50.1 y 51.7 °C. Cada cebador al extremo 3' inicio por una citosina o una guanina.

6.6 Amplificación por PCR

La amplificación del locus *pvron4* se llevó a cabo utilizando los cebadores *pvron4*dir 5' CACAGTGCAACCATGTCTCG 3', y *pvron4*rev 5' GCAAGCTAATTTCACAAGTCTTC 3'. La reacción de PCR se realizó mediante una *Touchdown* PCR utilizando la enzima KAPA-HiFi HotStart Readymix (kapabiosystems®) y ADN genómico de la cepa VCG-I como control positivo, en reacciones de 25µL, con el siguiente perfil térmico: un paso de denaturación inicial a 95°C por 5 min, 10 ciclos de 20 seg a 98°C, 15 seg a 68°C (la temperatura disminuyó un grado por ciclo), y 1 min a 72°C, seguido por 35 ciclos de 20 seg a 98°C, 15 seg a 60°C, 1 min a 72°C y un paso de extensión final de 5 min a 72°C.

Los productos amplificados fueron verificados por electroforesis en gel de agarosa al 1,5%, con el marcador de peso Hyperladder I. Posteriormente, se realizó la purificación de los productos utilizando el kit de purificación *UltraClean PCR Clean-up* (MOBIO®), siguiendo las recomendaciones del fabricante. Los fragmentos purificados fueron confirmados por electroforesis en gel de agarosa al 1,5%, con el marcador de peso Hyperladder I.

6.7 Ligación, transformación y clonación del fragmento purificado

Los productos de PCR purificados fueron ligados al vector de clonación pGEM T-easy, el cual fue usado para la transformación de células *Escherichia coli* de la cepa JM109. Las bacterias (colonias) recombinantes fueron seleccionadas por el método de alfa complementación y su capacidad de crecer en presencia de ampicilina. Éstas fueron confirmadas por PCR utilizando la enzima MangoTaq™ ADN Polimerasa y cebadores internos para el gen (*intdir*: 5' TGTGGGTGGCGAGAGTG 3', e *intrev*: 5' ATATTTCCATTGCTGTACTAGG 3') bajo las siguientes condiciones térmicas: un paso de 5 min a 95°C, 35 ciclos de 20 seg a 98°C, 15 seg a 60°C y un paso de extensión final por 5 min a 72°C.

6.8 Extracción del ADN plasmídico y secuenciación

Las colonias recombinantes confirmadas por PCR fueron tomadas para realizar inóculos en medio LB suplementado con ampicilina. Los inóculos se incubaron toda la noche a 37°C en agitación constante. Posteriormente, se centrifugaron a 4.000 rpm durante 20 minutos y el pellet celular fue utilizado para la extracción del ADN plasmídico utilizando el *kit UltraClean® 6 Minute Mini Plasmid Prep* (MOBIO®), de acuerdo con las recomendaciones del fabricante. La calidad y concentración del ADN plasmídico extraído se verificó por electroforesis en gel de agarosa al 1,5%, con el marcador de peso Hyperladder I.

ADN plasmídico procedente de dos colonias recombinantes por muestra, clonados a partir de reacciones de PCR independientes, fueron enviados a secuenciar. La secuenciación se realizó bidireccionalmente utilizando el *kit BigDye Terminator* (MACROGEN, Seúl, Corea del Sur), con los cebadores universales T7 y SP6 y un juego de cebadores internos (*pvron4dseq*: 5'

CACTAGAAAAGCTAAACATAAACC 3', y *pvron4rseq*: 5' ACTCCAATGGTCCTCAAC 3') diseñados previamente para la secuenciación.

6.9 Análisis de la diversidad genética y el efecto de la selección natural sobre el locus *pvron4*

Los electroferogramas obtenidos por secuenciación fueron ensamblados usando el software CLC DNA workbench v3 (CLC bio, Cambridge, MA, USA). 73 secuencias consenso obtenidas en este estudio, fueron analizadas junto con 7 secuencias de referencia (de las cepas Salvador-I (Sal-I), Mauitania-I, India-VII, Brazil-I, Corea del Norte, Ctg y P.vivax01) y 13 secuencias ortólogas (de las especies *Plasmodium cynomolgi*, *Plasmodium knowlesi*, *Plasmodium inui*, *Plasmodium fragile*, *Plasmodium coatneyi*, *Plasmodium reichenowi*, *Plasmodium falciparum*, *Plasmodium bergei*, *Plasmodium yoelii*, *Plasmodium chabaudi*, *Plasmodium vinckei*, *Plasmodium gaboni*, y *Plasmodium gallinaceum* (Sanger institute: gal28a.d000001405.Contig1/gal28a.d000000110.Contig1)), obtenidas de las bases de datos GenBank y PlasmoDB.

Las secuencias de ADN de los aislados colombianos fueron utilizadas para identificar regiones repetitivas con el algoritmo T-REKS [55]. Adicionalmente, a partir de estas secuencias, se infirió las secuencias deducidas de aminoácidos, con las cuales se realizó un alineamiento múltiple utilizando el algoritmo MUSCLE [56]. Este alineamiento fue editado manualmente para asegurar la alineación correcta de las regiones repetitivas utilizado el software GeneDoc v.2.6.002 [57]. A partir de este alineamiento, se infirió el alineamiento de las secuencias de ADN, manteniendo el marco de lectura, mediante el uso del software Pal2Nal [58].

A partir del alineamiento de ADN obtenido de las secuencias colombianas y las secuencias de referencia de *P. vivax*, se estimó la diversidad genética usando el software DnaSP v5 [59]. Los estimadores de diversidad tomados en cuenta fueron: el número de sitios segregantes (S), el número de sitios singleton (Ss), el número de sitios parsimoniosos (P), el número de haplotipos (H) y la diversidad nucleotídica (Hd), los cuales se calcularon teniendo en cuenta el polimorfismo de secuencia (SNPs: *Single Nucleotide Polimorphism*) y polimorfismo de longitud (InDels: *Insertions/ Deletions*),

Para evaluar desviaciones de la neutralidad, se utilizaron únicamente las secuencias colombianas, mediante el cálculo de los estadísticos de Tajima [60], Fu & Li [61], Fu [62] y Fay & Wu [63], teniendo en cuenta simulaciones de coalescencia, para obtener los intervalos de confianza y su significancia estadística. Las regiones repetitivas o con gaps fueron excluidas de los análisis.

Adicionalmente, se calculó el número promedio de sustituciones sinónimas por sitio sinónimo (d_S) y el número promedio de sustituciones no-sinónimas por sitio no-sinónimo (d_N) para identificar selección a nivel intra-específico, utilizando el método modificado de Nei-Gojobori [64], con el software MEGA v.6 [65] Por otra parte, para determinar señales de selección natural a lo largo de la historia evolutiva de *Plasmodium* (usando las secuencias de *P. vivax* junto con las secuencias ortólogas de parásitos filogenéticamente cercanos) se calculó el número promedio de divergencias sinónimas por sitio sinónimo (K_S) y el número promedio de divergencias no-sinónimas por sitio no-sinónimo (K_N), utilizando el método modificado de Nei-Gojobori con la corrección de Jukes-Cantor [66]. Las diferencias entre las tasas de sustitución intra- e inter-especie fueron determinadas aplicando la prueba de Fisher (*Fisher Exact Test*, recomendada cuando el número de sustituciones sinónimas y/o no-sinónimas es menor a 10) o la prueba Z incorporada en el software MEGA v.6. Adicionalmente, se utilizó el test de McDonald-Kretiman (MK) [67] para evaluar desviaciones de la neutralidad usando un servidor web [68]. Finalmente, la tasa evolutiva ($\omega = d_N/d_S$ y/o K_N/K_S) se analizó mediante una ventana deslizante para evaluar el efecto de la selección a lo largo del gen. Sitios individuales bajo selección fueron identificados calculando las tasas de sustituciones sinónimas y no-sinónimas por codón mediante los métodos SLAC, FEL, REL, IFEL [69], MEME [70] y FUBAR [71], implementados en el servidor en línea Datamonkey [72]. Para estos análisis, las regiones repetitivas o con gaps no fueron tenidas en cuenta.

6.10 Análisis de selección positiva episódica sobre el locus *pvron4*

Para evaluar la existencia de linajes bajo selección diversificante episódica en la historia evolutiva de *Plasmodium* para el locus *pvron4*, se utilizó el método *Branch-site REL* (*Branch-site Random effects likelihood*) [73]. A partir de 14 secuencias ortólogas de diferentes especies del género, se realizó un alineamiento múltiple usando el algoritmo MUSCLE y a partir de éste, se infirió el mejor modelo evolutivo por medio del software MEGA. Posteriormente, se infirió una filogenia con el método de máxima verosimilitud usando el modelo JFF+G+F. Esta se utilizó como

referencia para analizar la selección positiva linaje-especifica con el método *Branch-site REL* en el paquete HyPhy [74], usando el alineamiento de ADN inferido por Pal2Nal.

6.11 Uso efectivo de codones

El número efectivo de codones (ENC) se estimó utilizando el software ENCprime [75] y DnaSP. Esta es una medida de presión selectiva a nivel traduccional, que conlleva a la pérdida o ganancia de función de la proteína [75]. Esta prueba compara el uso de cada codón con una distribución nula (uso uniforme de codones sinónimos). Valores ENC cercanos a 61 indica que todos los codones sinónimos para cada aminoácido son usados equitativamente, mientras valores cercanos a 0 sugieren un sesgo o uso preferencial de codones [76]. La significancia estadística se puede ver afectada por la longitud del gen o la recombinación [76], por esta razón, se utilizó la medida de la desviación del uso uniforme de codones o CBI (del inglés *codón bias index*), la cual toma un rango de valores de 0 (uso uniforme de codones sinónimos) a 1 (sesgo máximo de codones) [77] .

6.12 Desequilibrio de ligamiento y recombinación

El desequilibrio de ligamiento (LD) fue evaluado mediante el cálculo del estimador Z_{Ns} [78]. Una regresión lineal entre éste y la distancia nucleotídica fue realizada para evaluar si la recombinación intragénica ocurre en *pvron4*, utilizando el software DnaSp v.5. La recombinación fue evaluada por el estimador ZZ [79], por el número mínimo de eventos de recombinación (Rm) [80], por el algoritmo GARD [81] y por el programa RDP v.3 [82].

6.13 Diferenciación y estructura genética poblacional del locus *pvron4*

El grado de diferenciación genética (o heterogeneidad inter-poblacional) en el locus *pvron4* entre las regiones de donde se colectaron muestras (Costa Pacífica, Urabá/Bajo-Cauca/Sur de Córdoba y la región de la Orinoquia-Amazonia), se estimó mediante un Análisis de la Varianza Molecular (AMOVA) y el índice de fijación de Wright (F_{ST}), utilizando el programa Arlequin v.3.1 [83]. Adicionalmente, para evaluar los posibles caminos mutacionales que dan origen a los haplotipos de *pvron4*, la distribución de éstos y sus frecuencias, se construyó una red mediana de unión utilizando el programa NETWORK v.5. Este método expresa múltiples caminos evolutivos plausibles en forma de ciclos, unidos por vectores interpretados como secuencias ancestrales extintas [84].

6.14 Predicción de dominios putativos y del potencial antigénico de *pvron4*

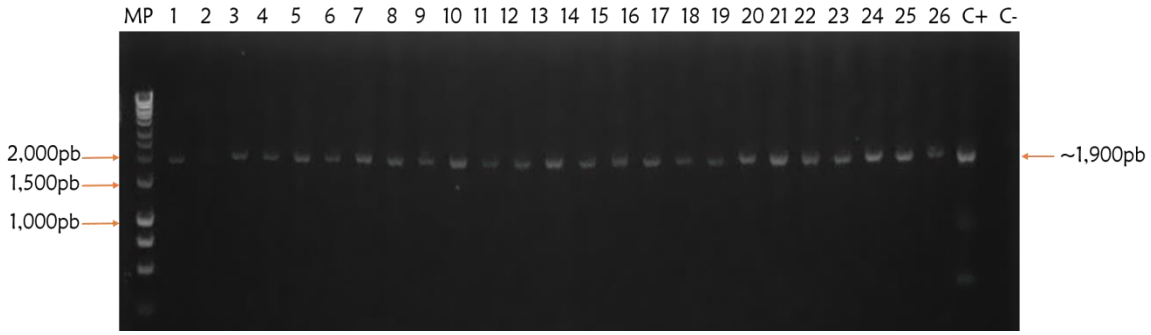
El algoritmo *Blastp* del servidor web NCBI fue utilizado para identificar dominios putativos en *Pv*RON4 utilizando la secuencia Sal-1 como referencia. Con el fin de evaluar el potencial antigénico de *Pv*RON4, se predijo parámetros como su antigenicidad [85], así como su hidrofilicidad [86], la disponibilidad al solvente de la proteína [87] y sus potenciales epítopes B lineales [88], utilizando la herramienta BepiPred en la base de datos IEDB (*Immune epitope database and analysis resource*). Estas pruebas fueron realizadas con dos diferentes haplotipos de *Pv*RON4 diferenciados por el número de repeticiones: el haplotipo 6 (con 1 copia de GEHGEHAEHGE) y el haplotipo 17 (con 7 copias de la repetición), con el fin de determinar si el *repeat* y su longitud cumple una función en el comportamiento antigénico de la proteína.

7. Resultados

7.1 Análisis y genotipificación de las muestras

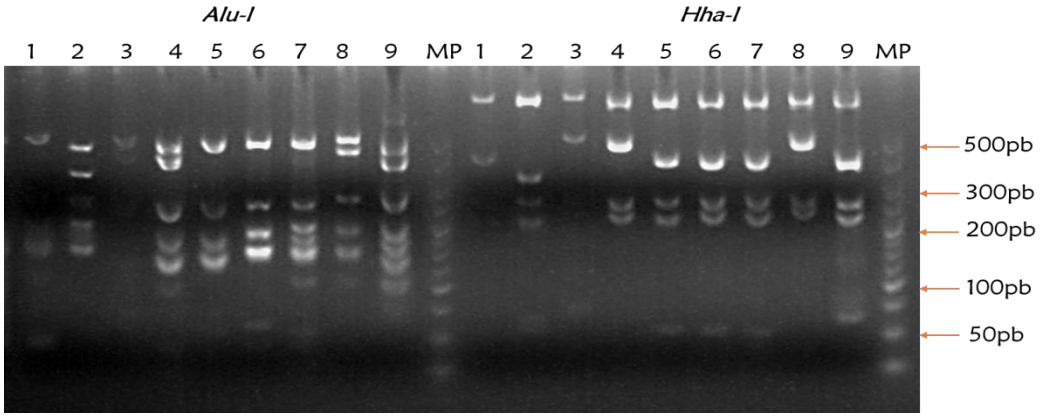
73 secuencias del locus *pvron4* fueron obtenidas a partir de muestras colectadas de la población colombiana (24 de la Amazonia-Orinoquia, 21 de la Costa Pacífica, 28 del Urabá/Bajo Cauca/Sur de Córdoba). La presencia de ADN genómico del parasito fue previamente confirmada mediante la amplificación de un fragmento de la subunidad ribosomal 18S de *P. vivax*. Las muestras positivas para esta especie fueron tipificadas por RFLPs del gen *pvmsp3-α*, confirmando la presencia de infecciones únicas (Fig. 7-8) y distintos genotipos parasitarios (Fig. 7-9). Posteriormente, se amplificó el gen *pvron4* (Fig. 7-10) a partir de las muestras seleccionadas y los productos fueron purificados (Fig. 7-11) y enviados a secuenciar.

Figura 7-8. Productos amplificados del gen *pvmsp3-α*.



Todas las muestras analizadas en este estudio correspondieron a infecciones únicas. El análisis por RFLPs del gen *pvmsp3α* confirmó los resultados de genotipificación obtenidos previamente con *pvmsp1*. Posición 1: marcador de peso molecular (MP) Hyperladder I, posición 2-26: infecciones únicas, C+: Control positivo, ADN de la cepa VCG-I, C-: Control negativo.

Figura 7-9. Productos de la digestión enzimatica del gen *pvmsp3α* tratado con las enzimas *Alu-I* y *Hha-I*.

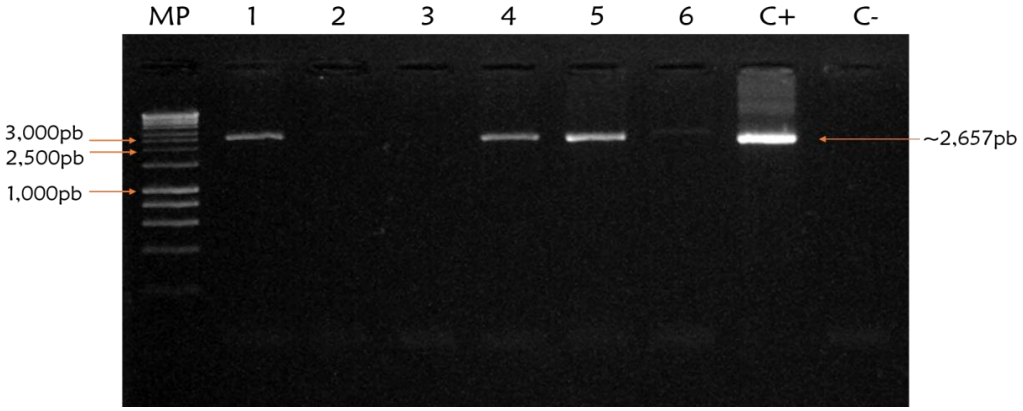


El patrón de bandeo obtenido tras la digestión enzimática evidenció la presencia de distintos genotipos. Posición 1-9: gen *pvmsp3α* digerido con las enzimas respectivas. Marcador de peso molecular (MP) Hyperladder V. Las mismas muestras fueron tratadas con las dos enzimas.

7.2 Diversidad genética en el locus *pvron4*

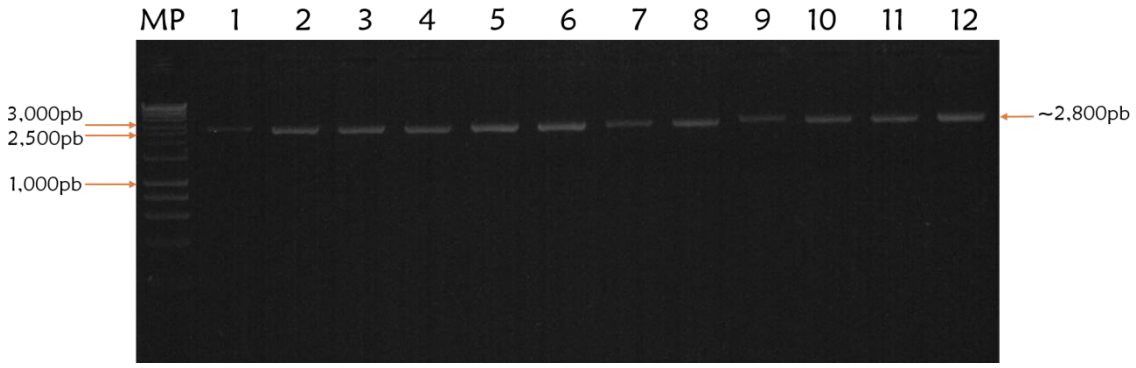
pvron4 fue anotado inicialmente a partir de la cepa VCG-I como un gen de 2.657pb [18], sin embargo, el locus analizado a partir de las muestras colombianas, presentó una variación en longitud como consecuencia de dos repeticiones en tándem ubicadas hacia el extremo 5' del gen (Fig. 7-12). Estas repeticiones están constituidas por copias de la secuencia nucleotídica GTGGCGAGA, que codifica para los aminoácidos GES (1-3 veces) y una secuencia más larga CGGAGAGCACGGTGAACACGCTGAACATGGGGAGCA, que codifica para la secuencia GEHGEHAEHGE, la cual se repite entre 1-7 veces (Figura 7-13 y Anexo 1).

Figura 7-10. Productos de amplificación del gen *pvron4*.



Posición 1: Marcador de peso molecular (MP) Hyperladder I. Posiciones 2-6: productos de PCR de muestras provenientes del departamento de Amazonas, C+: Control positivo ADN de la cepa VCG-I, C-: Control negativo.

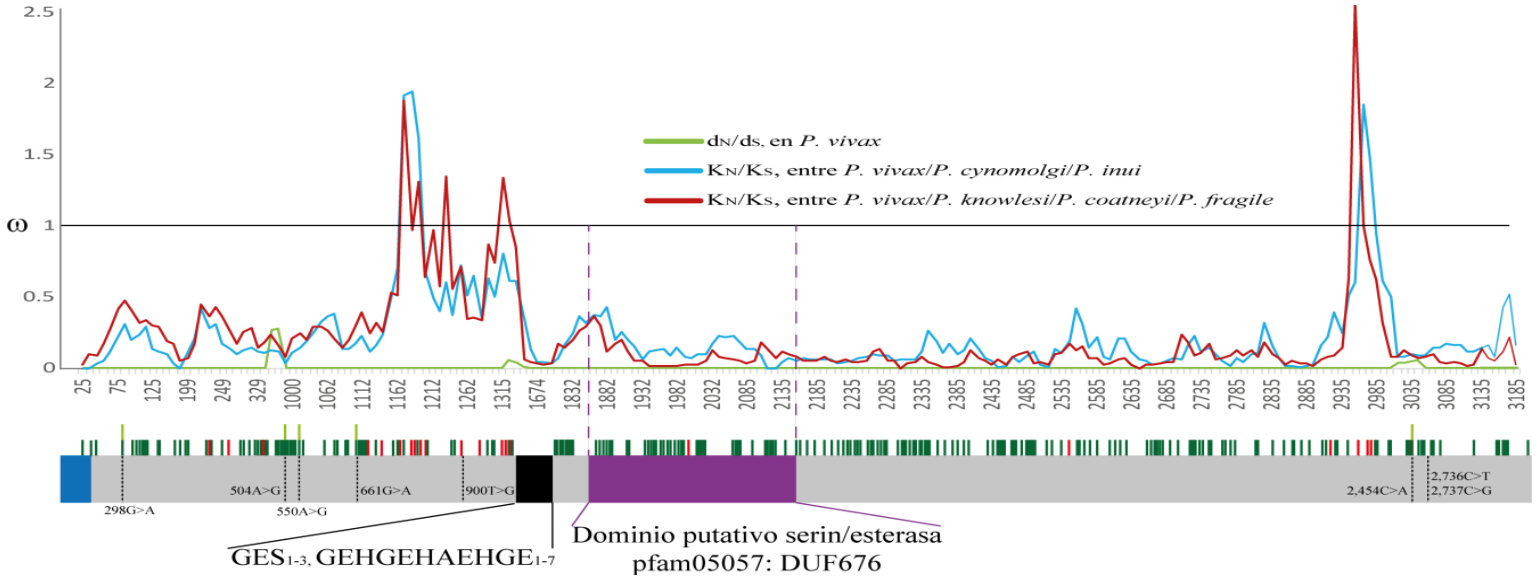
Figura 7-11. Productos purificados del gen *pvron4*.



Productos purificados representativos del total de muestras enviadas a secuenciar. Posición 1-12: gen *pvron4* purificado. Marcador de peso molecular (MP) Hyperladder I. Nótese que la longitud del fragmento varía entre muestras.

A nivel de secuencia, un bajo número de polimorfismos de nucleótido simple (SNPs) fue encontrado. A partir de 2.542 sitios analizados entre las secuencias colombianas y las de referencia, el número de SNPs varió entre 5 a 8 (Tabla 1). Los estimadores de diversidad genética mostraron valores bajos ($\theta_w = 4.7 \times 10^{-4}$ y $\pi = 4.1 \times 10^{-4}$) en la población colombiana. Estos valores se mantuvieron constantes cuando se analizaron las secuencias colombianas con las de referencia (Tabla 1). Así mismo, el número total de haplotipos identificados y la diversidad haplotípica en la población colombiana fueron bajos. Sin embargo, al analizar las inserciones/deleciones (InDels) en *pvron4*, se encontró un total de 21 haplotipos y valores altos en los estimadores de diversidad (Tabla 1). Teniendo en cuenta los SNPs y los InDels entre secuencias de referencia y las secuencias colombianas de *pvron4*, fueron identificados un total de 32 haplotipos (Anexo 2).

Figura 7-12. Representación esquemática del CDS de *pvron4* y ventana deslizante para la tasa ω.



Los valores ω de *ron4* dentro de *P. vivax* se muestran en verde a lo largo del gen. Así mismo, las tasas de divergencia entre las especies que infectan primates y *P. vivax* son mostradas en azul (para las especies evolutivamente cercanas a *P. vivax*) y rojo (más distantes). Una representación del gen *ron4* es dada bajo la ventana deslizante indicando el péptido señal (azul), la región repetitiva (negro) y el dominio putativo serin/esterasa (púrpura). Las posiciones de SNPs de *P. vivax* se indican con líneas discontinuas; los números corresponden a las posiciones del alineamiento del Anexo 2. Los sitios bajo selección purificante dentro de la especie son representados con líneas verde claras y verde oscuras entre especies, mientras los sitios seleccionados positivamente entre especies son mostrados por líneas rojas.

Figura 7-13. Secuencia de aminoácidos de los Haplotipos 6 y 17 del locus *pvron4*

```
                 *        20         *        40         *        60         *        80         *       100         *
H 6  : PHEVADKNGEDTSQEGSHAGALSPAHVGGES---GGEHGEHGEHAEHGEHGEHGE------------------------------------------------------------ :  52
H 17 : PHEVADKNGEDTSQEGSHAGALSPAHVGGESGESGGEHGEHGEHAEHGEHGEHGEHAEHGEHGEHGEHAEHGEHGEHGEHAEHGEHGEHGEHAEHGEHGEHGEHAEHGEHGEHGE : 115

         120         *       140         *       160
H 6  : ------------HGEHGEHAEQEEHQIYGTIHPEAEASILPALKE :  85
H 17 : HAEHGEHGEHGEHGEHGEHAEQEEHQIYGTIHPEAEASILPALKE : 160
```

El alineamiento de las dos clases de repeticiones en tándem y el número de éstas en el haplotipo 6 (corto) y el haplotipo 17 (largo) encontrados para el locus *pvron4*. La repetición GES se encuentra de 1-3 veces y la repetición GEHGEHAEHGEH se encuentra 1-7 veces en todos los haplotipos.

Tabla 1. Estimadores de diversidad genética calculados a partir del polimorfismo de nucleótido simple (SNP) y la longitud de secuencia en *pvron4*.

	Polimorfismo de nucleótido simple								Polimorfismo de longitud de secuencia				
n	**Sitio**	**Ss**	**S**	**Ps**	**H**	**Hd**	**θ_w**	**π**	**Sitios**	**N° InDels**	**H**	**Hd**	**π**
Secuencias totales													
80	2.542	8	3	5	11	0,67	$7,6 \times 10^{-4}$	$4,3 \times 10^{-4}$	408	21	21	0,82	$9,9 \times 10^{-4}$
Secuencias Colombianas													
73	2.464	5	0	5	8	0,65	$4,7 \times 10^{-4}$	$4,1 \times 10^{-4}$	289	15	15	0,78	$8,0 \times 10^{-4}$

Los estimadores de diversidad genética fueron calculados usando las secuencias obtenidas de bases de datos junto con secuencias obtenidas a partir de 73 aislados colombianos (secuencias totales), así como solo a partir de las secuencias colombianas. n: número de aislados analizados, sitios: total de sitios analizados, Ss: número de sitios segregantes, S: número de sitios singleton, Ps: número de sitios parsimonioso informativos, H: número de haplotipos, Hd: diversidad haplotípica, θw: estimador de Watterson, π: diversidad nucleotídica por sitio.

7.3 Efecto de la selección en el locus *pvron4*

No se encontraron valores estadísticamente significativos para *pvron4* utilizando los estimadores de Tajima, Fu & Li, Fay & Wu y Fu (Tabla 2). Así mismo, la prueba MK no mostró desviaciones de la neutralidad (Tabla 3). Por otra parte, la diferencia d_N-d_S (Tabla 3) evidenció que la tasa de sustitución sinónima fue más alta que la tasa de sustitución no-sinónima ($p = 0.000$, *Fisher exact test*), sugiriendo que *pvron4* se encuentra bajo selección purificante. Adicionalmente, a través de la ventana deslizante, se observaron valores de $\omega < 1$ a lo largo del gen (Fig. 7-12).

Tabla 2. Pruebas de neutralidad, desequilibrio de ligamiento y recombinación para el gen *pvron4* en la población colombiana de *P. vivax***.**

N	**Gen**	**Tajima**	**Fu and Li**		**Fay and Wu's H**	**Fu´s Fs**	**Z_{nS}**	**ZZ**	**RM**
		D	D*	F*					
Secuencias totales									
80	2.578	NP	NP	NP	NP	NP	$0,154^{\tau}$	-0,002	0
Secuencias Colombianas									
73	2.578	-0,037	-0,035	-0,020	-0,039	-0,028	$0,163^{\tau}$	-0,002	0

Las pruebas de neutralidad no arrojaron valores estadísticamente significativos. Z_{nS}: Promedio de R^2 de todos los pares de comparaciones; ZZ: diferencia Z_{nS}-Za. Rm: Mínimo número de eventos de recombinación. $p < 0.05\tau$. NA: prueba no aplicada, ya que no todas las secuencias pertenecían a la misma población.

Tabla 3. Evaluación de la sustitución intra-especie (d_N-d_S), divergencia (K_N-K_S) e índice de neutralidad de la prueba MK.

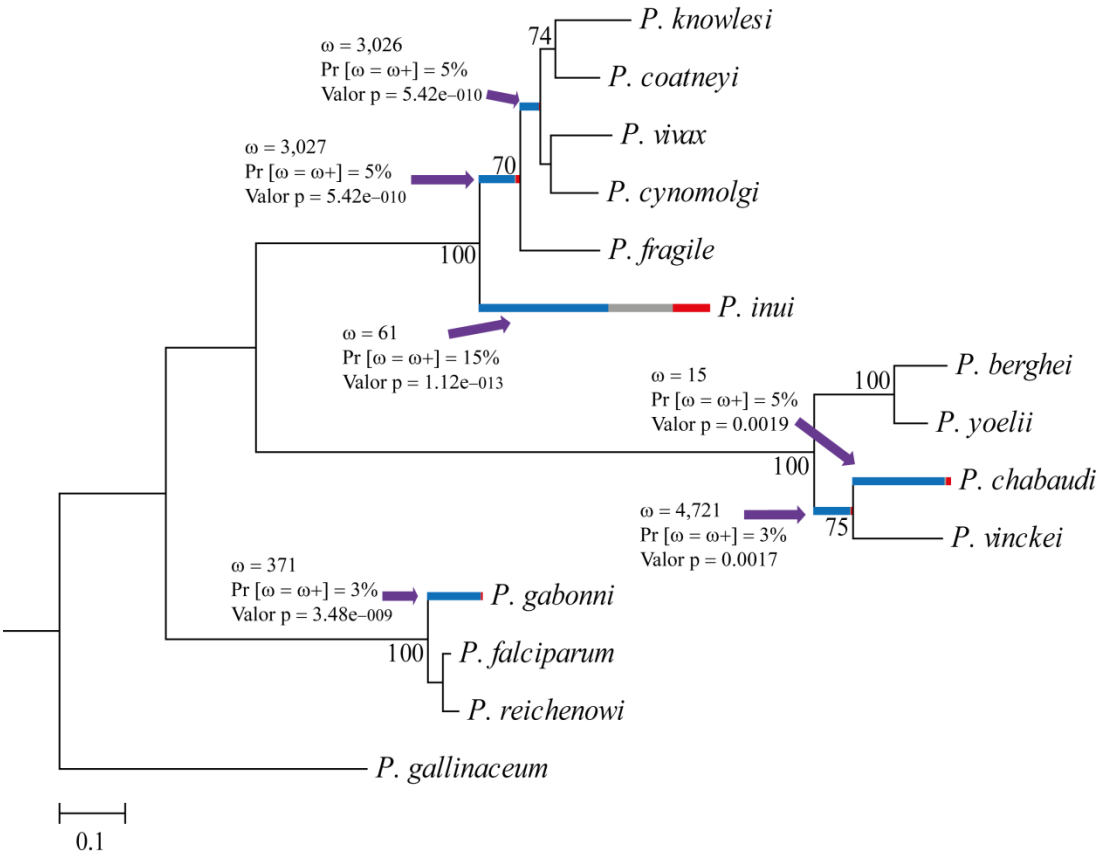
P. vivax	***P. vivax/Plasmodium* ssp**						
	P. knowlesi	***P. inui***	***P. coatneyi***	***P. cynomolgi***	***P. fragile***	***Pcyn/Pinu***	***Pkno/Pcoa/Pfra***
d_N-d_S				**K_N-K_S**			
	-0,011$^{\beta}$	-0,006$^{\beta}$	-0,006$^{\beta}$	-0,009$^{\beta}$	-0,010$^{\beta}$	-0,015$^{\beta}$	-0,025$^{\beta}$
-0.002*			**NI**				
	0,696	0,597	0,857	1,133	0,919		

Tasa de sustitución no-sinónima (d_N) y tasa de sustitución sinónima (d_S) dentro *P. vivax*. Divergencia no-sinónima (K_N) y sinónima (K_S) entre *P. vivax* y especies filogenéticamente cercanas del género. Índice de neutralidad (NI), estimado por el test de McDonald-Kreitman (MK) usando la corrección de Jukes Cantor. *: $p < 0,01$, $^{\beta}$: $p < 0,001$, Pcyn: *P. cynomolgi*, Pinu: *P. inui*, Pkno: *P. knowlesi*, Pcoa: *P. coatneyi*, Pfra: *P. fragile*.

Cuando se comparan especies filogenéticamente relacionadas, la ventana deslizante mostró valores de $\omega < 1$, con pocas regiones con un $\omega \geq 1$ (posiciones 1.172-1.325, 2.955- 2.975 y en la posición 2.955; Figura 7-12). La diferencia K_N-K_S (Tabla 3) mostró valores negativos sugiriendo que la selección purificante tiene un rol importante en la historia evolutiva de este locus en el género *Plasmodium*. Por otro lado, el cálculo de selección por codón basado en los métodos de máxima probabilidad (SCAL, FEL, IFEL, REL y FUBAR) y el método bayesiano (MEME), encontraron 4 sitios seleccionados negativamente (28, 112, 149 y 643) a lo largo del gen dentro de *P. vivax*, mientras ningún sitio bajo selección positiva (Figura 7-12). Estos métodos encontraron que entre especies hay 162 sitios bajo selección negativa y 21 bajo selección positiva (Figura 7-12).

A partir de secuencias ortólogas de 14 especies de *Plasmodium*, se infirió una filogenia que fue utilizada como marco de referencia para realizar la prueba *Branch-site REL*. Este método encontró 6 ramas (linajes) con evidencia de selección episódica (Figura 7-14). Tres corresponden a ramas ancestrales (internas) mientras las otras tres (externas) dieron origen a *P. inui*, *P. chabaudi* y *P. gabonni*. Adicionalmente, el método MEME encontró codones bajo selección episódica diversificante.

Figura 7-14. Linajes sujetos a selección diversificante episódica en *pvron4*



Las ramas (linajes) bajo selección positiva episódica fueron identificadas por el método *Branch-site REL*. La figura muestra los valores ω+, el porcentaje de sitios (Pr [ω = ω+]) y los valores p. La sombra de cada color en las ramas indica fuerza de la selección (en rojo se muestra ω+ > 15, azul ω- ≤ 1, y gris ω = 1). El tamaño de cada color representa el porcentaje de sitios en la clase correspondiente. Las ramas han sido clasificadas como sometidas a selección episódica diversificante por el valor p corregido para múltiples pruebas, utilizando el método de Bonferroni-Holm con p <0,002.

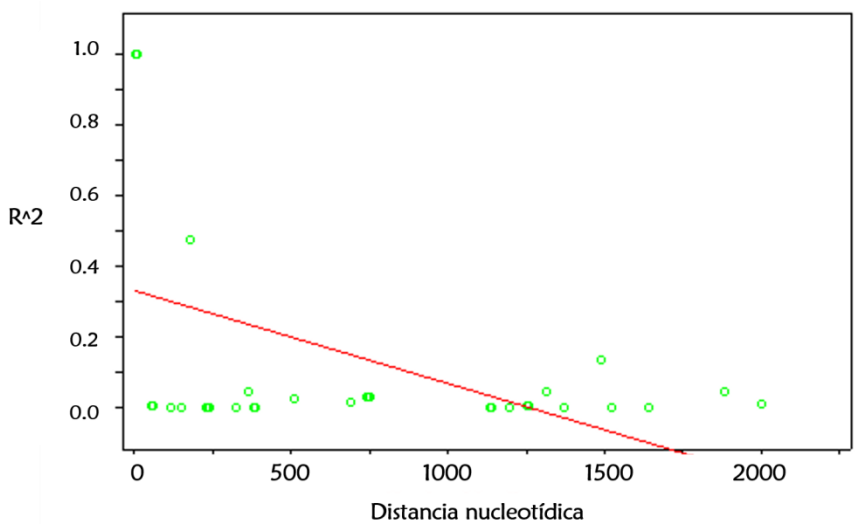
7.4 Uso efectivo de codones

Dada la alta conservación del locus *pvron4*, se evaluó la desviación del uso efectivo de codones, como una medida de selección a nivel traduccional. El ENC para *pvron4* estimado por ENCprime (N'c = 53, scaled X^2= 0,103) y por DnaSP (ENC = 55,4, scaled X^2= 0,159), mostró valores cercanos a 61, con un CBI de 0,162, sugiriendo que no existe un sesgo en el uso efectivo de codones, descartando una selección a nivel traduccional.

7.5 Desequilibrio de ligamiento y recombinación

El desequilibrio de ligamiento (LD) entre los polimorfismos de *pvron4* fue evaluado por el estimador Z_{nS}, el cual mostró un valor de 0,15, no significativo. La regresión lineal entre el LD y la distancia nucleotídica mostró una reducción del LD conforme la distancia aumenta, sugiriendo recombinación (Figura 7-15). Sin embargo, el estimador ZZ mostró un valor de -0,0019 y no se detectaron sitios mínimos de recombinación. Así mismo, los métodos GARD y el programa RDP no encontraron *break-points* ni *tracks* de recombinación, respectivamente.

Figura 7-15. Regresión lineal entre el desequilibrio de ligamiento y la distancia nucleotídica.



La línea roja indica la tendencia del desequilibrio de ligamiento (R^2) a disminuir, a medida que aumenta la distancia nucleotídica.

7.6 Estructura genética de la población Colombiana respecto al locus *pvron4*

Para evaluar la diversidad genética geoespacial de *pvron4* en Colombia, se calculó un AMOVA entre las tres regiones colombianas (Costa Pacífica [departamentos de Chocó y Nariño], Urabá/Bajo-Cauca/Sur de Córdoba [departamentos de Córdoba y Antioquia] y la región de Orinoquia-Amazonia [departamentos de Amazonas, Guainía]), así como el índice de fijación de Wright (F_{ST}) entre las diferentes poblaciones. El AMOVA mostró que existen diferencias estadísticamente significativas entre las subpoblaciones dentro de las regiones (F_{SC} = 0,06, p = 0,04). Las comparaciones entre pares de poblaciones (regiones) estimadas por el índice F_{ST} dio valores cercanos a 0, con una diferencia estadísticamente significativa entre las regiones de Urabá/Bajo-Cauca/Sur de Córdoba y la Orinoquía-Amazonia (Tabla 4). La subpoblación de Nariño

(costa pacífica) puede ser la responsable de las diferencias genéticas respecto a cada una de las demás subpoblaciones (Tabla 5).

Tabla 4. Estadístico F_{ST} inter-poblaciones para *pvron4*.

F_{ST}	Costa Pacifica	Urabá/Bajo-Cauca/Sur de Córdoba	Orinoquia-Amazonia
Costa Pacífica		0,44043	0,31836
Urabá/Bajo-Cauca/Sur de Córdoba	-0,00581		0,01465
Orinoquia-Amazonia	0,00542	**0,08128**	

El F_{ST} fue calculado para las poblaciones parasitarias en tres regiones de Colombia. Los valores cercanos a 0 indican baja diferenciación genética mientras valores cercanos a 1 indican alta diferenciación genética. Valores por debajo de la diagonal indican los valores F_{ST} y aquellos sobre la diagonal representan los valores p. Valores en negrita indican diferencias estadísticamente significativas p<0,02

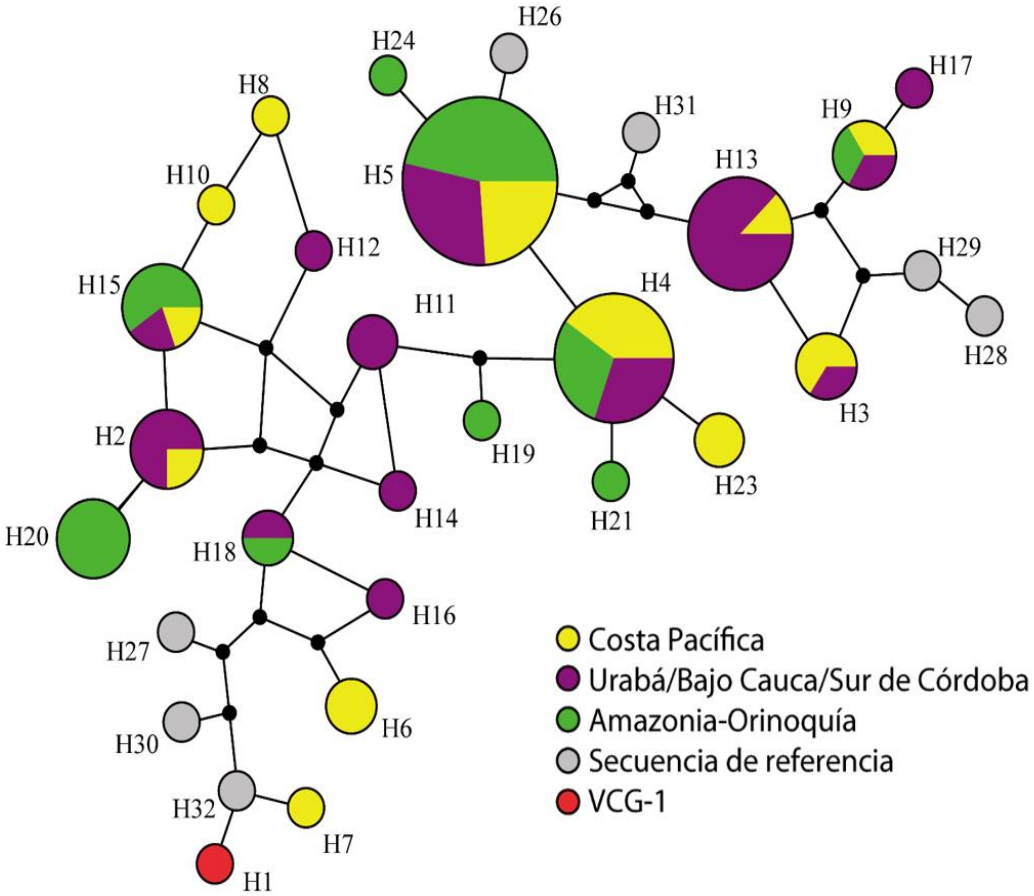
Tabla 5. Estadístico F_{ST} interpoblacional para *pvron4*

F_{ST}	Chocó	Nariño	Antioquia	Córdoba	Amazonas
Chocó		0,00195	0,22656	0,19336	0,05176
Nariño	**0,40062**		0,02930	0,05273	0,00879
Antioquia	0,04272	**0,28000**		0,43555	0,01172
Córdoba	0,02428	0,13422	-0,01859		0,07715
Amazonas	0,09516	**0,27552**	**0,22928**	0,05302	

El F_{ST} fue calculado para las subpoblaciones parasitarias (departamentos) en Colombia. Los valores por debajo de la diagonal corresponden al valor del F_{ST} (cercanos a 0 indican baja diferenciación genética y cercanos a 1 indican alta diferenciación genética), mientras los valores p se encuentran por encima de la diagonal. En negrita se muestran los valores estadísticamente significativos.

Para entender mejor la relación evolutiva entre los haplotipos de *pvron4*, se empleó un *Median Joining Network* para describir el conjunto de vías mutacionales potenciales que dan lugar a los 32 haplotipos disponibles para el locus. Network mostró que las distintas poblaciones del parásito comparten haplotipos, independiente de la región geográfica, los cuales se hallan relacionados por distintos caminos mutacionales y secuencias ancestrales (*median vectors*) (Figura 7-16 y 7-17). Dentro de los haplotipos más frecuentes, se encuentran: H5 (53,1%), seguido por H4 (31,2%), H13 (25%), H3 y H20 (12,5% cada uno), Cabe resaltar la presencia de haplotipos únicos en las poblaciones de Nariño (H7, H8 y H10) y Amazonas (H21, H22, H2) que pueden ser considerados alelos raros o privados (Anexo 2 y Anexo 4).

Figura 7-16. *Median–joining network* por regiones colombianas**.**



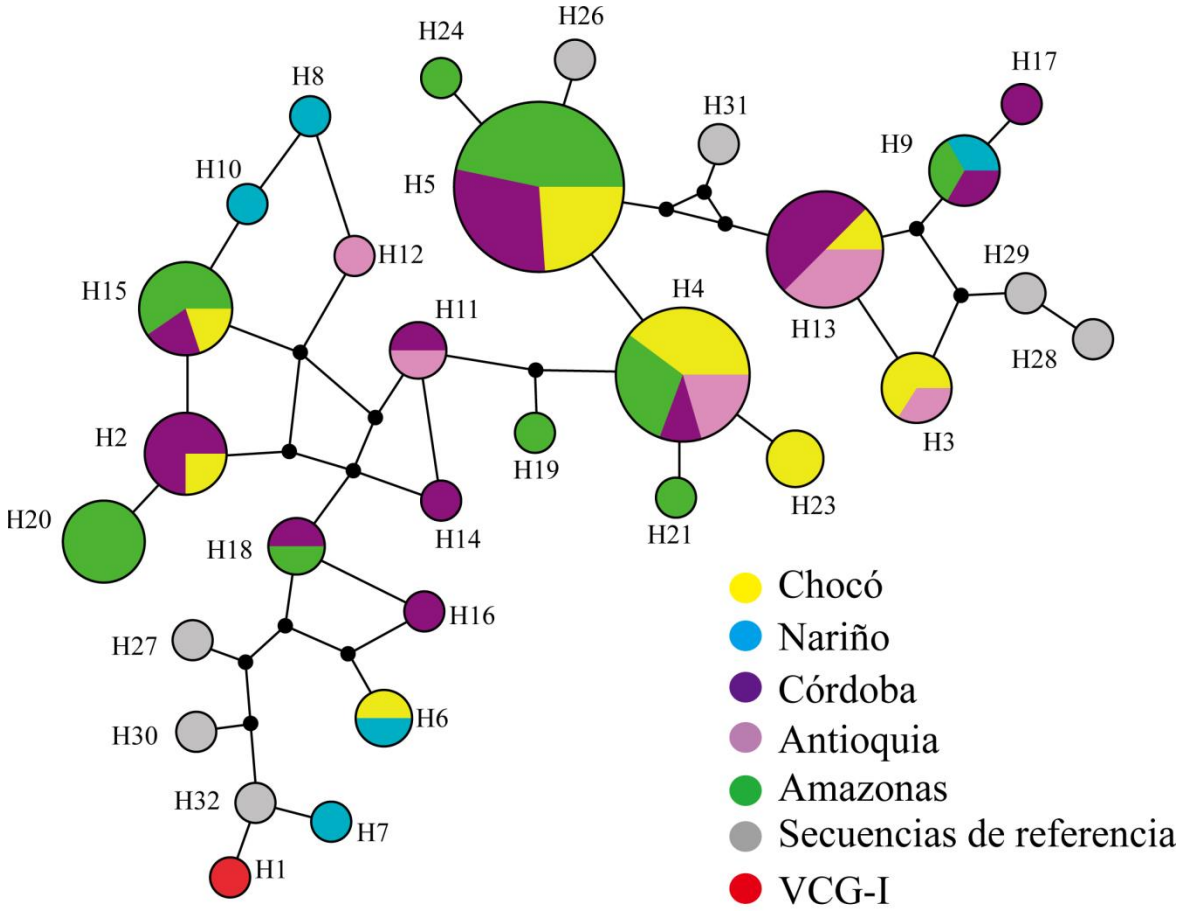
La gráfica muestra los haplotipos de *pvron4* identificados a partir de los aislados de las tres regiones de Colombia. Los haplotipos 22 y 28 fueron incluidos dentro del haplotipo 15 al utilizar el algoritmo "star contraction" [83], para simplificar la interpretación de la red. Cada nodo es un haplotipo y su tamaño indica la frecuencia del mismo. Las líneas que conectan los haplotipos entre sí, representan los distintos caminos mutacionales y los puntos negros (*median vectors*) son las secuencias ancestrales que explican la relación y origen evolutivo.

7.7 Predicción de dominios putativos y del potencial antigénico de *pvron4*

El análisis de la secuencia de *Pv*RON4 de la cepa Sal-I mostró la presencia de un dominio putativo de la superfamilia de proteínas serin/esterasa (pfam05057) entre los aminoácidos 311-425 (nucleótidos 931 a 1.275, números basados en la secuencia de referencia Sal-I, número de acceso PlasmoDB: PVX_091434) con un valor e de 1,10e-03. Este dominio está ubicado después de la región repetitiva y es altamente conservado (Figura 7-12).

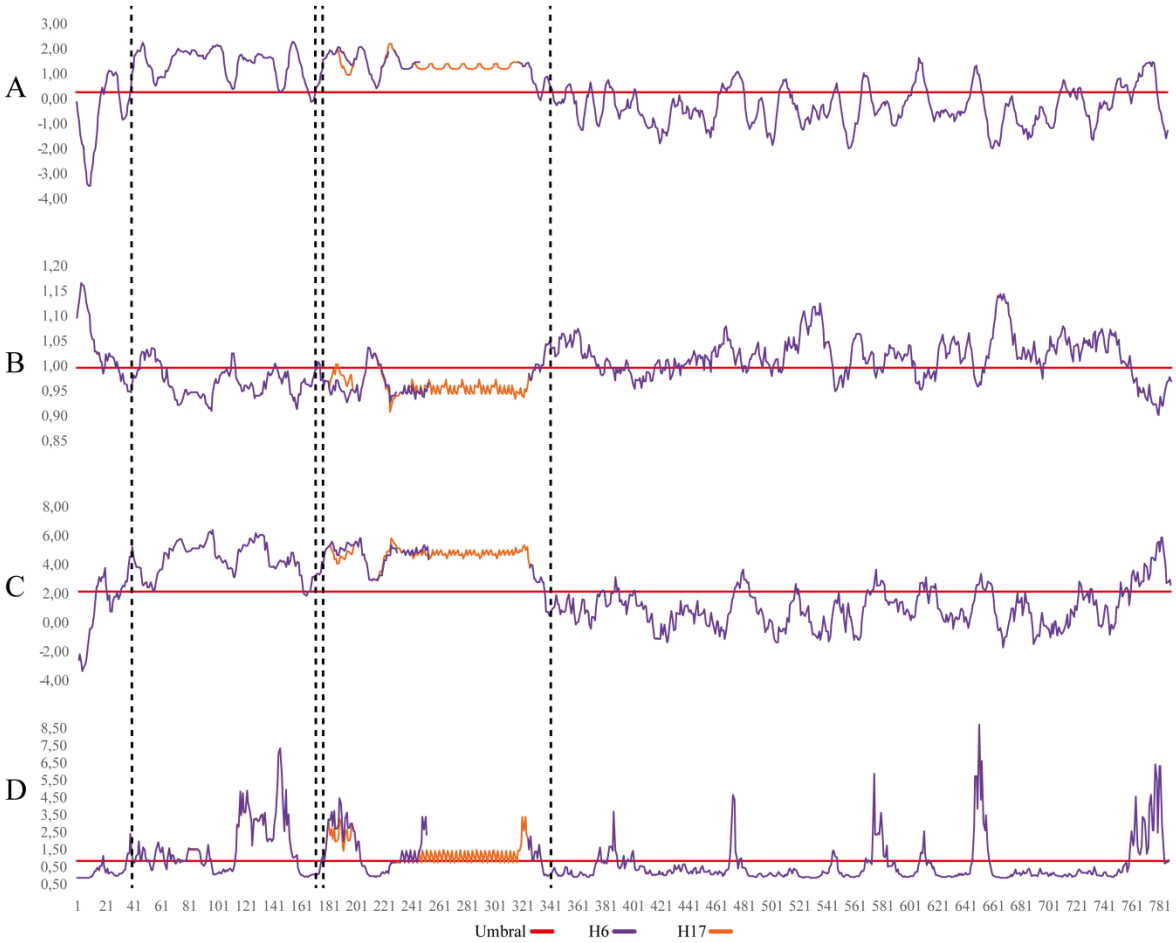
Por otra parte, de acuerdo con la predicción de la antigenicidad de Kolaskar & Tongaonkar y de epitopes B lineales, hay una región potencialmente antigénica entre las posiciones 41-171 y 178 hasta la 256 para el haplotipo 6 y hasta la posición 340 para el haplotipo 17 (Figura 7-18). Esto es congruente con las predicciones de hidrofilicidad y accesibilidad al solvente, de manera que la región N-terminal de *Pv*RON4 parece ser un potencial blanco inmune. Así mismo, el número de repeticiones en tándem expande la región antigénica, aumentando la posibilidad de ser reconocidas por el sistema inmune. En contraste, la región central (después de la región repetitiva) y la C-terminal, parecen ser menos antigénicas, así como menos expuestas al solvente (Figura 7-18).

Figura 7-17. *Median –joining network* para las poblaciones parasitarias por departamentos.



Relación evolutiva entre los 32 haplotipos identificados a partir de los aislados colectados de cinco departamentos de Colombia, junto con la secuencia de referencia. Los haplotipos 22 y 28 fueron incluidos dentro del haplotipo 15 al utilizar el algoritmo "star contraction" [89], para simplificar la interpretación de la red. Cada nodo es un haplotipo y su tamaño indica la frecuencia del mismo. Las líneas que conectan los haplotipos entre sí, representan los distintos caminos mutacionales y los puntos negros (*median vectors*) son las secuencias ancestrales que explican la relación y origen evolutivo. El color corresponde a cada una de las subpoblaciones de origen.

Figura 7-18. Regiones potencialmente antigénicas de *Pv*RON4.



Predicción de **A.** Epítopes B lineales (umbral=0,35); **B.** Antigenicidad de Kolaskar y Tongaonkar (umbral=1,0); **C.** Hidrofilicidad de Parker (umbral=2,31) y **D.** Accesibilidad al solvente (umbral=1,0). Las líneas punteadas indican las regiones que tiene la mayor probabilidad de ser reconocidas por el sistema inmune de acuerdo al puntaje en cada predicción. Los valores invertidos para la antigenicidad se deben al cálculo de la propensión antigénica, independientemente de la posible ocurrencia de aminoácidos en epítopes y superficie de la proteína.

8. Discusión

Varias de las proteínas contenidas dentro de los organelos apicales del parásito parecen ser cruciales para la invasión a la célula hospedera y, por ende, representan blancos vacunales prometedores. Dentro de las proteínas localizadas en estos organelos apicales, se distinguen las proteínas del complejo RON, las cuales se encuentran participando en la formación del *TJ* [13, 14, 36, 42]. Este *TJ* juega un rol decisivo para el ingreso del parásito a la célula hospedera y el cierre de la vacuola parasitófora [13]. La proteína RON4, localizada en el complejo de invasión, está presente en los miembros del phylum *Apicomplexa* [16, 19, 20], sugiriendo que ésta forma parte de una ruta conservada de invasión. Por tal motivo, esta proteína ha sido descrita como potencial candidato a vacuna.

Son varias las características de un candidato a vacuna [12, 90]. Una de las principales es que debe presentar una baja diversidad genética para así evitar las respuestas inmunes alelo-específicas que disminuyen la eficacia de las vacunas. Análisis de cepas de laboratorio de *P. falciparum* de diferentes orígenes geográficos, mostraron que el locus *pfron4* es altamente conservado, presentando solamente una sustitución de aminoácidos [19]. En *P. vivax*, este locus parece tener el mismo comportamiento. El análisis de 5 secuencias de referencia, reveló una baja diversidad genética [21]. Con el fin de confirmar que *pvron4* es un gen altamente conservado en *P. vivax*, este estudio analizó 73 asilados clínicos de la población colombiana. A pesar de aumentar el número de secuencias analizadas, la diversidad de *pvron4* se mantuvo baja. Nuestros resultados mostraron que *pvron4* presenta una diversidad genética más baja de lo reportado previamente [21]. Solo 8 SNPs fueron identificados en 80 secuencias disponibles, contrastando con los 14 previamente identificados [21]. Este alto número de SNPs (14 en estudios previos), es consecuencia de un alineamiento erróneo de la región repetitiva en los estudios anteriores.

pvron4 mostró un comportamiento similar a otras proteínas de organelos apicales [91, 92]. *pvrap1* y *pvrap2* tienen un π entre 0,0009 y 0,001, respectivamente [92], mientras *pvron4* mostró un valor aún más bajo al de *pvrap1,* sugiriendo que éste tiene una baja diversidad genética, siendo el locus más conservado reportado hasta la fecha para *P. vivax*. Aunque *pvron4* fue altamente conservado en secuencia, éste mostró un alto polimorfismo en tamaño. Estudios previos identificaron dos tipos de repeticiones hacia el extremo N-terminal de la proteína codificada [18]. Estas repeticiones fueron reportadas como copias imperfectas de los aminoácidos GGE[H/S]GE[H/S] y [G/A]EH. Sin embargo, nuestros resultados muestran que la región repetitiva de *pvron4* está compuesta por dos tipos de repeticiones con una identidad del 100%. La primera codifica para tres aminoácidos GES (1 a 3 copias) y la segunda para los aminoácidos GEHGEHAEHGE (1 a 7 copias). Estas repeticiones generan un alto número de haplotipos (alelos) diferentes en la población de *P. vivax*.

Estudios previos han sugerido que las repeticiones en tándem podrían tener un papel importante como mecanismo de evasión de las respuestas inmunes del hospedero [37, 54, 93, 94]. En este estudio, 21 haplotipos fueron identificados en *Pv*RON4 cuando los InDels fueron analizados. Adicionalmente, la región N-terminal de *Pv*RON4 parece ser la porción de la proteína más expuesta según los resultados de disponibilidad al solvente e hidrofilicidad. Esta región (después del péptido señal y la región repetitiva), parece ser un potencial blanco antigénico, debido a que es allí donde se predijo el mayor número de potenciales epítopes B lineales. Las secuencias repetitivas identificadas, amplían la región expuesta al solvente y el potencial antigénico de la proteína. Consecuentemente, la región N-terminal de *Pv*RON4 podría ser la región expuesta al sistema inmune del hospedero y la repetición podría estar actuando como una cortina de humo inmunológica. Futuros estudios de antigenicidad e inmunogenicidad son necesarios para confirmar esta hipótesis. Adicionalmente, ya que la región repetitiva fue altamente conservada en secuencia, ésta podría tener un papel estructural o funcional importante, como se ha sugerido recientemente para CSP [95, 96].

Mientras la región N-terminal parece estar expuesta al sistema inmune, la región central y la C-terminal podrían ser funcionalmente importantes. Los test de neutralidad (e.g. Tajima, Fu & Li) no mostraron valores estadísticamente significativos y por lo tanto no se descarta la neutralidad. Si *pvron4* está bajo neutralidad, éste debería mostrar un alto polimorfismo, a menos que exista una restricción funcional o estructural [97]. Dado que este locus fue altamente conservado en secuencia, una restricción funcional/estructural es probable. Sin embargo, una selección a nivel traduccional podría también ser la responsable de la alta conservación en secuencia de *pvron4*. Los

análisis del uso preferencial de codones no mostraron un sesgo en el uso de codones (ENC = 53 a 55). De hecho, este valor fue similar al reportado para el genoma completo (ENC = 52,18) [98]. Esto sugiere que la alta conservación del locus *pvron4* no se debe a una selección a nivel traduccional y podría ser resultado de una fuerte selección purificante.

En *pvron4*, la tasa d_S fue significativamente mayor a d_N según la prueba de Fisher (*Fisher exact test*), lo que sugiere que este locus está bajo selección purificante. Sin embargo, evaluar cómo actúa la selección natural en antígenos altamente conservados no es fácil [91, 99, 100]. Con el fin de evaluar el efecto de la selección sobre antígenos parasitarios, estudios previos han comparado las secuencias de *P. vivax* con especies filogenéticamente relacionadas [91, 99-102]. El análisis de la ventana deslizante de la tasa ω mostró valores menores a 1 hacia la región central de la proteína, así como hacia el extremo C-terminal. Adicionalmente, la tasa K_S fue estadísticamente mayor a K_N y varios sitios bajo selección purificante entre especies fueron detectados en estas regiones, sugiriendo que la selección purificante juega un papel importante durante la evolución del locus en el género. Teniendo en cuenta que las regiones funcionales tienden a evolucionar más lentamente y suelen ser conservadas entre especies [103], estos resultados sugieren que la región central y C-terminal de *Pv*RON4 podrían ser funcionalmente importantes. La presencia de cisteínas conservadas en la porción C-terminal (que en la mayoría de las veces están asociadas con la interacción proteína-proteína) podría estar mediando la interacción entre RON4 y AMA-1 y/u otras RONs [17, 104], mientras la presencia de un dominio putativo serin/esterasa en la región central de la proteína, podría estar involucrado en la entrada de RON4 a la célula hospedera.

Estudios en *P. falciparum* y *T. gondi* han mostrado que la región C-terminal de RON4 parece jugar un rol importante para la proceso de invasión [17, 20]. RON4 se encuentra dentro del glóbulo rojo sirviendo de anclaje para el complejo AMA1/RON2 [13, 20, 47]. Por lo tanto, durante las etapas iniciales de este proceso, RON4 debe ser secretada y entrar al glóbulo rojo por un mecanismo aún desconocido. La presencia de un dominio serin/esterasa en la región central de *Pv*RON4, podría dar pistas del mecanismo de acción. Esta región es una de las más estructuradas de la proteína, conteniendo varios sitios bajo selección purificante que reiteran un papel funcional/estructural. Este dominio podría entonces estar participando en el ingreso de RON4 a la célula huésped.

Mientras la región N-terminal de *Pv*RON4 parece estar asociado con la evasión de la respuesta inmune, la región central (conteniendo el dominio serin/esterasa), podría estar asociado con la ruptura de los enlaces éster en los fosfolípidos que constituyen la membrana de la célula

hospedera. Esta ruptura permitiría la entrada de RON4 al citoplasma del glóbulo rojo o del hepatocito. Una vez dentro, la región C-terminal de RON4 serviría de anclaje para la proteína RON2, que a su vez permitiría la interacción entre el parásito y la célula hospedera, mediada por AMA1. Aunque esto es sólo una hipótesis, la presencia de este dominio, así como la observación de varios sitos bajo selección purificante en la región central y C-terminal, y la alta conservación de esta proteína entre especies emparentadas, la sustentaría. Futuros estudios funcionales son necesarios para su confirmación.

A pesar de que *ron4* fue altamente conservada entre especies y que la selección purificante parece ser importante durante la evolución de este locus en *Plasmodium*, algunos sitios bajo selección positiva fueron identificados. Estos coinciden con las regiones donde se observó un $\omega > 1$. Estudios previos han mostrado que algunos antígenos (independiente de su diversidad genética) presentan regiones/codones bajo selección positiva episódica que podrían haber permitido la adaptación a diferentes hospederos [102, 105, 106]. La topología del árbol obtenida para el gen *ron4* es similar a la obtenida al analizar el ADN mitocondrial [107]. Las relaciones filogenéticas de las especies que infectan roedores y homínidos son obtenidas en la filogenia de *ron4*. Sin embargo, estas relaciones no se ven representadas para las especies que infectan primates del viejo mundo. Estas especies tienen una compleja historia evolutiva que incluye aspectos biogeográficos, adaptación a nuevos hospederos macacos, e incluso el cambio de monos a humanos [105, 108]. La selección episódica observada podría ser entonces consecuencia de una diversificación rápida (en la región N-terminal y una pequeña porción en la región C-terminal) de este grupo de parásitos que permitió la adaptación de RON4 de una población ancestral a nuevos hospederos disponibles, como previamente ha sido sugerido [102, 105, 106, 108].

En Colombia, se encontró un relativo alto número de haplotipos en el locus *pvron4* que resultan de la combinación entre SNPs y de repeticiones en tándem. El análisis *median joining* mostró que las regiones colombianas comparten la mayoría de los haplotipos y parecen ser genéticamente similares. Sin embargo, al analizar el AMOVA se encontró que existe un 6% de variación estimada entre las subpoblaciones que conforman las distintas regiones. El valor F_{ST} mostró que las subpoblaciones de Nariño, Antioquia y Amazonas pueden estar genéticamente aisladas, debido a la presencia de haplotipos únicos. Estos es concordante con previos estudios en Colombia con otros antígenos parasitarios [100], así como con estudios de ADN mitocondrial en América [109], que sugieren que la población parasitaria en América está estructurada y con limitado flujo génico. Sin embargo, ya que algunas subpoblaciones analizadas aquí tenían un tamaño de muestra limitado, se

hace necesario ampliar el número de secuencias para confirmar estos resultados y determinar si existe un efecto demográfico que permita el aislamiento de estas poblaciones.

9. Conclusiones y recomendaciones

9.1 Conclusiones

El diseño de una vacuna completamente efectiva contra los parásitos que causan malaria, requiere de antígenos con una limitada diversidad genética que evite las respuestas inmunes alelo-especificas. El locus *pvron4* mostró tener una baja diversidad genética a nivel de secuencia, pero con una gran cantidad de haplotipos debido a repeticiones en tándem ubicadas en la región 5' del gen, que podrían estar involucradas en la evasión de la respuesta inmune. Por otra parte, la región central y la 3', son altamente conservadas, incluso entre especies. Estas regiones están bajo selección purificante, sugiriendo que son funcional o estructuralmente importantes en la proteina. Adicionalmente, la región central de *Pv*RON4 presenta un dominio putativo serin/esterasa. Nosotros hipotetizamos que este dominio permitiría la entrada de RON4 a la célula hospedera, mientras la región C-terminal serviría de anclaje al complejo AMA1/RONs. Teniendo en cuenta estos resultados, la región central o C-terminal de *Pv*RON4 son candidatos prometedores a incluir en el diseño de una vacuna basada en subunidades contra *P. vivax*.

9.2 Recomendaciones

Con base en la baja variabilidad de la proteína *Pv*RON4 y la participación de sus ortólogos en mecanismos de invasión conservados, es necesario realizar ensayos funcionales que permitan identificar la interacción entre esta proteína y las demás RONs que participan en el complejo de unión a AMA1 en *Plasmodium vivax*. Adicionalmente, se debe evaluar la actividad esterasa predicha computacionalmente, mediante pruebas bioquímicas. Estos estudios permitirían elucidar la hipótesis planteada acerca del rol potencial de RON4 en la entrada del parasito a la célula hospedera.

Así mismo, estudios de antigenicidad e inmunogenicidad son requeridos para evaluar la respuesta inducida por péptidos de la proteína *Pv*RON4 al usarlos como inmunógenos. Específicamente, es necesario evaluar la región central y C-terminal, las cuales parecen ser blancos prometedores, ya que representan la región más conservada y funcionalmente importante, de acuerdo a los resultados obtenidos en este estudio.

Por otro lado, aunque poco se ha elucidado acerca de la función de las repeticiones en tándem en proteínas de agentes patógenos, este estudio sustenta su participación en la expansión de regiones antigénicas como posibles cortinas de humo. Sin embargo, aún queda por explicar las razones por las que estas repeticiones (independiente del número de veces que se presentan) son altamente conservadas, y si estas podrían tener alguna función estructural, como ha sido sugerido para la proteína de circumsporozoito (CSP), mediante enfoques de biología molecular con parásitos *knockout* para las repeticiones en tándem del gen *csp*.

La conservación interespecifica de RON4 evidencia que esta proteína es un componente esencial en el género *Plasmodium*. Estudios previos en *P. falciparum* indican que los *knockout* de *pfron4* son inviables, lo que confirma su importancia para el parásito. Esto debería comprobarse en *P. vivax*, ya que, de ser así, RON4 representa no sólo un candidato más a vacuna, sino una diana terapéutica para intentar controlar la malaria mediante enfoques diferentes a la vacunación.

10. Bibliografía

1. Esmaeili Rastaghi, A.R., et al., *Genetic diversity and effect of natural selection at apical membrane antigen-1 (AMA-1) among Iranian Plasmodium vivax isolates.* Folia Parasitol (Praha), 2014. **61**(5): p. 385-93.
2. Guerra, C.A., R.W. Snow, and S.I. Hay, *Mapping the global extent of malaria in 2005.* Trends Parasitol, 2006. **22**(8): p. 353-8.
3. UNICEF. *Achieving the Malaria MDG Target. Reversing the Incidence of Malaria 2000–2015*. 2015; Available from: http://www.unicef.org/publications/files/Achieving_the_Malaria_MDG_Target.pdf.
4. WHO, U. *Achieving the malaria MDG target: reversing the incidence of malaria 2000–2015*. 2015 08/2015 [cited 2015 25/10/2015].
5. Tjitra, E., et al., *Multidrug-resistant Plasmodium vivax associated with severe and fatal malaria: a prospective study in Papua, Indonesia.* PLoS Med, 2008. **5**(6): p. e128.
6. Price, R.N., N.M. Douglas, and N.M. Anstey, *New developments in Plasmodium vivax malaria: severe disease and the rise of chloroquine resistance.* Curr Opin Infect Dis, 2009. **22**(5): p. 430-5.
7. Brutus, L., et al., *Plasmodium vivax malaria during pregnancy, Bolivia.* Emerg Infect Dis, 2013. **19**(10): p. 1605-11.
8. Gething, P.W., et al., *A long neglected world malaria map: Plasmodium vivax endemicity in 2010.* PLoS Negl Trop Dis, 2012. **6**(9): p. e1814.
9. Pérez, H.A., *La malaria por Plasmodium Vivax (Grassi y Feletti, 1890) en los trópicos y los retos de la cura radical.* Interciencia, 2004. **29**: p. 490-495.
10. Mueller, I., et al., *Key gaps in the knowledge of Plasmodium vivax, a neglected human malaria parasite.* Lancet Infect Dis, 2009. **9**(9): p. 555-66.
11. Cotter, C., et al., *The changing epidemiology of malaria elimination: new strategies for new challenges.* Lancet, 2013. **382**(9895): p. 900-11.
12. Barry, A.E. and A. Arnott, *Strategies for designing and monitoring malaria vaccines targeting diverse antigens.* Front Immunol, 2014. **5**: p. 359.
13. Weiss, G.E., et al., *Revealing the sequence and resulting cellular morphology of receptor-ligand interactions during Plasmodium falciparum invasion of erythrocytes.* PLoS Pathog, 2015. **11**(2): p. e1004670.

14. Boucher, L.E. and J. Bosch, *The apicomplexan glideosome and adhesins - Structures and function.* J Struct Biol, 2015. **190**(2): p. 93-114.
15. Giovannini, D., et al., *Independent roles of apical membrane antigen 1 and rhoptry neck proteins during host cell invasion by apicomplexa.* Cell Host Microbe, 2011. **10**(6): p. 591-602.
16. Lebrun, M., et al., *The rhoptry neck protein RON4 re-localizes at the moving junction during Toxoplasma gondii invasion.* Cell Microbiol, 2005. **7**(12): p. 1823-33.
17. Takemae, H., et al., *Characterization of the interaction between Toxoplasma gondii rhoptry neck protein 4 and host cellular beta-tubulin.* Sci Rep, 2013. **3**: p. 3199.
18. Arevalo-Pinzon, G., et al., *Annotation and characterization of the Plasmodium vivax rhoptry neck protein 4 (PvRON4).* Malar J, 2013. **12**: p. 356.
19. Morahan, B.J., et al., *Plasmodium falciparum: genetic and immunogenic characterisation of the rhoptry neck protein PfRON4.* Exp Parasitol, 2009. **122**(4): p. 280-8.
20. Alexander, D.L., et al., *Plasmodium falciparum AMA1 binds a rhoptry neck protein homologous to TgRON4, a component of the moving junction in Toxoplasma gondii.* Eukaryot Cell, 2006. **5**(7): p. 1169-73.
21. Garzon-Ospina, D., J. Forero-Rodriguez, and M.A. Patarroyo, *Inferring natural selection signals in Plasmodium vivax-encoded proteins having a potential role in merozoite invasion.* Infect Genet Evol, 2015. **33**: p. 182-8.
22. Salavatifar, M., et al., *High-Level Expression, Purification and Characterization of A Recombinant Plasmodium vivax Apical Membrane Antigen 1: Implication for vivax Malaria Vaccine Development.* Cell J, 2015. **17**(3): p. 520-31.
23. Patarroyo, M.A., D. Calderon, and D.A. Moreno-Perez, *Vaccines against Plasmodium vivax: a research challenge.* Expert Rev Vaccines, 2012. **11**(10): p. 1249-60.
24. HJ, H.P., *One Million Deaths by Parasites*, in *PLOS Medical Journals' Community Blog*2015.
25. INS, *GESTIÓN PARA LA VIGILANCIA ENTOMOLÓGICA Y CONTROL DE LA TRANSMISIÓN DE MALARIA*, 2015.
26. WHO. *World malaria report 2012*. 2012; Available from: http://www.who.int/malaria/publications/world_malaria_report_2012/wmr2012_no_profiles.pdf.
27. Rich, S.M. and F.J. Ayala, *Progress in malaria research: the case for phylogenetics.* Adv Parasitol, 2003. **54**: p. 255-80.
28. Smith, J.D. and A.G. Craig, *The surface of the Plasmodium falciparum-infected erythrocyte.* Curr Issues Mol Biol, 2005. **7**(1): p. 81-93.
29. WHO. *World Malaria Report 2015* 2015; Available from: http://www.who.int/malaria/publications/world-malaria-report-2015/wmr2015-without-profiles.pdf?ua=1.

30. WHO. *World Malaria Report 2014*. 2014; Available from: http://www.who.int/malaria/publications/country-profiles/profile_col_en.pdf?ua=1.
31. SIVIGILA, *Boletin epidemiológico semanal*, 2015.
32. Camargo-Ayala, P.A., et al., *High Plasmodium malariae Prevalence in an Endemic Area of the Colombian Amazon Region.* PLoS One, 2016. **11**(7): p. e0159968.
33. *Protocolo de Vigilancia en Salud Pública*, 2014, Instituto Nacional de Salud. p. 25.
34. Winter, D.J., et al., *Whole Genome Sequencing of Field Isolates Reveals Extensive Genetic Diversity in Plasmodium vivax from Colombia.* PLoS Negl Trop Dis, 2015. **9**(12): p. e0004252.
35. Schneider, K.A. and A.A. Escalante, *Fitness components and natural selection: why are there different patterns on the emergence of drug resistance in Plasmodium falciparum and Plasmodium vivax?* Malar J, 2013. **12**: p. 15.
36. Cowman, A.F. and B.S. Crabb, *Invasion of red blood cells by malaria parasites.* Cell, 2006. **124**(4): p. 755-66.
37. Ferreira, M.U., M. da Silva Nunes, and G. Wunderlich, *Antigenic diversity and immune evasion by malaria parasites.* Clin Diagn Lab Immunol, 2004. **11**(6): p. 987-95.
38. Boeger, W.A., et al., *Neotropical Monogenoidea 59. Polyonchoineans from Characidium spp. (Characiformes: Crenuchidae) from southern Brazil.* Folia Parasitol (Praha), 2014. **61**(2): p. 120-32.
39. Farooq, F. and E.S. Bergmann-Leitner, *Immune Escape Mechanisms are Plasmodium's Secret Weapons Foiling the Success of Potent and Persistently Efficacious Malaria Vaccines.* Clin Immunol, 2015. **161**(2): p. 136-43.
40. Hoffman, S.L., et al., *The march toward malaria vaccines.* Vaccine, 2015.
41. Lin, C.S., et al., *The merozoite surface protein 1 complex is a platform for binding to human erythrocytes by Plasmodium falciparum.* J Biol Chem, 2014. **289**(37): p. 25655-69.
42. Tonkin, M.L., et al., *Host cell invasion by apicomplexan parasites: insights from the co-structure of AMA1 with a RON2 peptide.* Science, 2011. **333**(6041): p. 463-7.
43. Healer, J., et al., *Independent translocation of two micronemal proteins in developing Plasmodium falciparum merozoites.* Infect Immun, 2002. **70**(10): p. 5751-8.
44. Silvie, O., et al., *A role for apical membrane antigen 1 during invasion of hepatocytes by Plasmodium falciparum sporozoites.* J Biol Chem, 2004. **279**(10): p. 9490-6.
45. Hodder, A.N., et al., *The disulfide bond structure of Plasmodium apical membrane antigen-1.* J Biol Chem, 1996. **271**(46): p. 29446-52.
46. Besteiro, S., et al., *Export of a Toxoplasma gondii rhoptry neck protein complex at the host cell membrane to form the moving junction during invasion.* PLoS Pathog, 2009. **5**(2): p. e1000309.

47. Cao, J., et al., *Rhoptry neck protein RON2 forms a complex with microneme protein AMA1 in Plasmodium falciparum merozoites.* Parasitol Int, 2009. **58**(1): p. 29-35.
48. Lamarque, M., et al., *The RON2-AMA1 interaction is a critical step in moving junction-dependent invasion by apicomplexan parasites.* PLoS Pathog, 2011. **7**(2): p. e1001276.
49. Riglar, D.T., et al., *Super-resolution dissection of coordinated events during malaria parasite invasion of the human erythrocyte.* Cell Host Microbe, 2011. **9**(1): p. 9-20.
50. Alexander, D.L., et al., *Identification of the moving junction complex of Toxoplasma gondii: a collaboration between distinct secretory organelles.* PLoS Pathog, 2005. **1**(2): p. e17.
51. McBride, J.S. and H.G. Heidrich, *Fragments of the polymorphic Mr 185,000 glycoprotein from the surface of isolated Plasmodium falciparum merozoites form an antigenic complex.* Mol Biochem Parasitol, 1987. **23**(1): p. 71-84.
52. Fenton, B., A. Walker, and D. Walliker, *Protein variation in clones of Plasmodium falciparum detected by two dimensional electrophoresis.* Mol Biochem Parasitol, 1985. **16**(2): p. 173-83.
53. Grau, G.E. and C. Behr, *T cells and malaria: is Th1 cell activation a prerequisite for pathology?* Res Immunol, 1994. **145**(6): p. 441-54.
54. Hisaeda, H., K. Yasutomo, and K. Himeno, *Malaria: immune evasion by parasites.* Int J Biochem Cell Biol, 2005. **37**(4): p. 700-6.
55. Jorda, J. and A.V. Kajava, *T-REKS: identification of Tandem REpeats in sequences with a K-meanS based algorithm.* Bioinformatics, 2009. **25**(20): p. 2632-8.
56. Edgar, R.C., *MUSCLE: multiple sequence alignment with high accuracy and high throughput.* Nucleic Acids Res, 2004. **32**(5): p. 1792-7.
57. Nicholas, K.B., and Nicholas, H.B.J.,, *GeneDoc: A tool for editing and annotating multiple sequence alignments.* 1997.
58. Suyama, M., D. Torrents, and P. Bork, *PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments.* Nucleic Acids Res, 2006. **34**(Web Server issue): p. W609-12.
59. Librado, P. and J. Rozas, *DnaSP v5: a software for comprehensive analysis of DNA polymorphism data.* Bioinformatics, 2009. **25**(11): p. 1451-2.
60. Tajima, F., *Statistical method for testing the neutral mutation hypothesis by DNA polymorphism.* Genetics, 1989. **123**(3): p. 585-95.
61. Fu, Y.X. and W.H. Li, *Statistical tests of neutrality of mutations.* Genetics, 1993. **133**(3): p. 693-709.
62. Fu, Y.X., *Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection.* Genetics, 1997. **147**(2): p. 915-25.
63. Fay, J.C. and C.I. Wu, *Hitchhiking under positive Darwinian selection.* Genetics, 2000. **155**(3): p. 1405-13.

64. Zhang, J., H.F. Rosenberg, and M. Nei, *Positive Darwinian selection after gene duplication in primate ribonuclease genes.* Proc Natl Acad Sci U S A, 1998. **95**(7): p. 3708-13.
65. Tamura, K., et al., *MEGA6: Molecular Evolutionary Genetics Analysis version 6.0.* Mol Biol Evol, 2013. **30**(12): p. 2725-9.
66. Jukes, T.H., *Evolutionary pattern of specificity regions in light chains of immunoglobulins.* Biochem Genet, 1969. **3**(2): p. 109-17.
67. McDonald, J.H. and M. Kreitman, *Adaptive protein evolution at the Adh locus in Drosophila.* Nature, 1991. **351**(6328): p. 652-4.
68. Egea, R., S. Casillas, and A. Barbadilla, *Standard and generalized McDonald-Kreitman test: a website to detect selection by comparing different classes of DNA sites.* Nucleic Acids Res, 2008. **36**(Web Server issue): p. W157-62.
69. Kosakovsky Pond, S.L. and S.D. Frost, *Not so different after all: a comparison of methods for detecting amino acid sites under selection.* Mol Biol Evol, 2005. **22**(5): p. 1208-22.
70. Murrell, B., et al., *Detecting individual sites subject to episodic diversifying selection.* PLoS Genet, 2012. **8**(7): p. e1002764.
71. Murrell, B., et al., *FUBAR: a fast, unconstrained bayesian approximation for inferring selection.* Mol Biol Evol, 2013. **30**(5): p. 1196-205.
72. Delport, W., et al., *Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology.* Bioinformatics, 2010. **26**(19): p. 2455-7.
73. Kosakovsky Pond, S.L., et al., *A random effects branch-site model for detecting episodic diversifying selection.* Mol Biol Evol, 2011. **28**(11): p. 3033-43.
74. Pond, S.L., S.D. Frost, and S.V. Muse, *HyPhy: hypothesis testing using phylogenies.* Bioinformatics, 2005. **21**(5): p. 676-9.
75. Novembre, J.A., *Accounting for background nucleotide composition when measuring codon usage bias.* Mol Biol Evol, 2002. **19**(8): p. 1390-4.
76. Shields, D.C., et al., *"Silent" sites in Drosophila genes are not neutral: evidence of selection among synonymous codons.* Mol Biol Evol, 1988. **5**(6): p. 704-16.
77. Morton, B.R., *Chloroplast DNA codon use: evidence for selection at the psb A locus based on tRNA availability.* J Mol Evol, 1993. **37**(3): p. 273-80.
78. Kelly, J.K., *A test of neutrality based on interlocus associations.* Genetics, 1997. **146**(3): p. 1197-206.
79. Rozas, J., et al., *DNA variation at the rp49 gene region of Drosophila simulans: evolutionary inferences from an unusual haplotype structure.* Genetics, 2001. **158**(3): p. 1147-55.
80. Hudson, R.R. and N.L. Kaplan, *Statistical properties of the number of recombination events in the history of a sample of DNA sequences.* Genetics, 1985. **111**(1): p. 147-64.
81. Kosakovsky Pond, S.L., et al., *Automated Phylogenetic Detection of Recombination Using a Genetic Algorithm.* Mol Biol Evol, 2006. **23**(10): p. 1891-1901.

82. Martin, D. and E. Rybicki, *RDP: detection of recombination amongst aligned sequences.* Bioinformatics, 2000. **16**(6): p. 562-3.
83. Excoffier, L., G. Laval, and S. Schneider, *Arlequin (version 3.0): an integrated software package for population genetics data analysis.* Evol Bioinform Online, 2005. **1**: p. 47-50.
84. Bandelt, H.J., P. Forster, and A. Rohl, *Median-joining networks for inferring intraspecific phylogenies.* Mol Biol Evol, 1999. **16**(1): p. 37-48.
85. Kolaskar, A.S. and P.C. Tongaonkar, *A semi-empirical method for prediction of antigenic determinants on protein antigens.* FEBS Lett, 1990. **276**(1-2): p. 172-4.
86. Parker, J.M., D. Guo, and R.S. Hodges, *New hydrophilicity scale derived from high-performance liquid chromatography peptide retention data: correlation of predicted surface residues with antigenicity and X-ray-derived accessible sites.* Biochemistry, 1986. **25**(19): p. 5425-32.
87. Emini, E.A., et al., *Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide.* J Virol, 1985. **55**(3): p. 836-9.
88. Larsen, J.E., O. Lund, and M. Nielsen, *Improved method for predicting linear B-cell epitopes.* Immunome Res, 2006. **2**: p. 2.
89. Forster, P., et al., *Phylogenetic star contraction applied to Asian and Papuan mtDNA evolution.* Mol Biol Evol, 2001. **18**(10): p. 1864-81.
90. Richie, T.L. and A. Saul, *Progress and challenges for malaria vaccines.* Nature, 2002. **415**(6872): p. 694-701.
91. Pacheco, M.A., et al., *Evidence for negative selection on the gene encoding rhoptry-associated protein 1 (RAP-1) in Plasmodium spp.* Infect Genet Evol, 2010. **10**(5): p. 655-61.
92. Garzon-Ospina, D., L. Romero-Murillo, and M.A. Patarroyo, *Limited genetic polymorphism of the Plasmodium vivax low molecular weight rhoptry protein complex in the Colombian population.* Infect Genet Evol, 2010. **10**(2): p. 261-7.
93. Verra, F. and A.L. Hughes, *Biased amino acid composition in repeat regions of Plasmodium antigens.* Mol Biol Evol, 1999. **16**(5): p. 627-33.
94. Ramasamy, R., *Molecular basis for evasion of host immunity and pathogenesis in malaria.* Biochim Biophys Acta, 1998. **1406**(1): p. 10-27.
95. Ferguson, D.J., et al., *The repeat region of the circumsporozoite protein is critical for sporozoite formation and maturation in Plasmodium.* PLoS One, 2014. **9**(12): p. e113923.
96. Aldrich, C., et al., *Roles of the amino terminal region and repeat region of the Plasmodium berghei circumsporozoite protein in parasite infectivity.* PLoS One, 2012. **7**(2): p. e32524.
97. Kimura, M., *The neutral theory of molecular evolution*1983, Cambridge Cambridgeshire ; New York: Cambridge University Press. xv, 367 p.
98. Cornejo, O.E., D. Fisher, and A.A. Escalante, *Genome-wide patterns of genetic polymorphism and signatures of selection in Plasmodium vivax.* Genome Biol Evol, 2015. **7**(1): p. 106-19.

99. Forero-Rodriguez, J., D. Garzon-Ospina, and M.A. Patarroyo, *Low genetic diversity in the locus encoding the Plasmodium vivax P41 protein in Colombia's parasite population.* Malar J, 2014. **13**: p. 388.
100. Forero-Rodriguez, J., D. Garzon-Ospina, and M.A. Patarroyo, *Low genetic diversity and functional constraint in loci encoding Plasmodium vivax P12 and P38 proteins in the Colombian population.* Malar J, 2014. **13**: p. 58.
101. Pacheco, M.A., et al., *Evidence of purifying selection on merozoite surface protein 8 (MSP8) and 10 (MSP10) in Plasmodium spp.* Infect Genet Evol, 2012. **12**(5): p. 978-86.
102. Garzon-Ospina, D., J. Forero-Rodriguez, and M.A. Patarroyo, *Heterogeneous genetic diversity pattern in Plasmodium vivax genes encoding merozoite surface proteins (MSP) -7E, -7F and -7L.* Malar J, 2014. **13**: p. 495.
103. Graur, D., et al., *On the immortality of television sets: "function" in the human genome according to the evolution-free gospel of ENCODE.* Genome Biol Evol, 2013. **5**(3): p. 578-90.
104. Narum, D.L., et al., *Identification and characterization of the Plasmodium yoelii PyP140/RON4 protein, an orthologue of Toxoplasma gondii RON4, whose cysteine-rich domain does not protect against lethal parasite challenge infection.* Infect Immun, 2008. **76**(11): p. 4876-82.
105. Muehlenbein, M.P., et al., *Accelerated diversification of nonhuman primate malarias in Southeast Asia: adaptive radiation or geographic speciation?* Mol Biol Evol, 2015. **32**(2): p. 422-39.
106. Sawai, H., et al., *Lineage-specific positive selection at the merozoite surface protein 1 (msp1) locus of Plasmodium vivax and related simian malaria parasites.* BMC Evol Biol, 2010. **10**: p. 52.
107. Pacheco, M.A., et al., *Malarial parasite diversity in chimpanzees: the value of comparative approaches to ascertain the evolution of Plasmodium falciparum antigens.* Malar J, 2013. **12**: p. 328.
108. Mu, J., et al., *Host switch leads to emergence of Plasmodium vivax malaria in humans.* Mol Biol Evol, 2005. **22**(8): p. 1686-93.
109. Taylor, J.E., et al., *The evolutionary history of Plasmodium vivax as inferred from mitochondrial genomes: parasite genetic diversity in the Americas.* Mol Biol Evol, 2013. **30**(9): p. 2050-64.

11. Anexos

11.1 Anexo 1. Variación en el número de repeticiones en tándem en el locus *pvron4*.

```
                        *        20         *        40         *        60         *        80         *       100         *
H 6  : GCTGTCACCTGCGCATGTGGGTGGCGAGA---------GTGGCGGCGAGCACGGAGAGCACGGTGAACACGCTGAACATGGGGAGCACGGAGAGCACGGTGA------------ :  94
H 17 : GCTGTCACCTGCGCATGTGGGTGGCGAGAGTGGCGAGAGTGGCGGCGAGCACGGAGAGCACGGTGAACACGCTGAACATGGGGAGCACGGAGAGCACGGTGAACACGCTGAACA : 114

                  120         *       140         *       160         *       180         *       200         *       220
H 6  : ------------------------------------------------------------------------------------------------------------------ :   -
H 17 : TGGGGAGCACGGAGAGCACGGTGAACACGCTGAACATGGGGAGCACGGAGAGCACGGTGAACACGCTGAACATGGGGAGCACGGAGAGCACGGTGAACACGCTGAACATGGGGA : 228

                 *       240         *       260         *       280         *       300         *       320         *       340
H 6  : ---------------------------------------------------------------------------------------ACATGGCGAGCACGGTGAACAT :   115
H 17 : GCACGGAGAGCACGGTGAACACGCTGAACATGGGGAGCACGGAGAGCACGGTGAACACGCTGAACATGGGGAGCACGGAGAGCACGGTGAACATGGCGAGCACGGTGAACAT :   340
```

Secuencia de ADN de los haplotipos 6 (el alelo con el menor número de repeticiones) y 17 (el alelo con el mayor número de repeticiones) del locus *pvron4*.

11.2 Anexo 2. Alineamiento de los 32 haplotipos identificados para *pvron4*.

```
                                  *        20         *        40         *        60         *        80         *       100
H1_VCG                  : ATGTCTCGTAAAAGGGTTTTTCTTCTGTGCATTTTTCTGGCCATTTCCGCCGTTGTTGACGAAGCCGAGTCGTTCAGGAGTAAGCAAAAAGAAAGGAAAAAAAA :  104
H2                      : ATGTCTCGTAAAAGGGTTTTTCTTCTGTGCATTTTTCTGGCCATTTCCGCCGTTGTTGACGAAGCCGAGTCGTTCAGGAGTAAGCAAAAAGAAAGGAAAAAAAA :  104
H3                      : ATGTCTCGTAAAAGGGTTTTTCTTCTGTGCATTTTTCTGGCCATTTCCGCCGTTGTTGACGAAGCCGAGTCGTTCAGGAGTAAGCAAAAAGAAAGGAAAAAAAA :  104
H4                      : ATGTCTCGTAAAAGGGTTTTTCTTCTGTGCATTTTTCTGGCCATTTCCGCCGTTGTTGACGAAGCCGAGTCGTTCAGGAGTAAGCAAAAAGAAAGGAAAAAAAA :  104
H5                      : ATGTCTCGTAAAAGGGTTTTTCTTCTGTGCATTTTTCTGGCCATTTCCGCCGTTGTTGACGAAGCCGAGTCGTTCAGGAGTAAGCAAAAAGAAAGGAAAAAAAA :  104
H6                      : ATGTCTCGTAAAAGGGTTTTTCTTCTGTGCATTTTTCTGGCCATTTCCGCCGTTGTTGACGAAGCCGAGTCGTTCAGGAGTAAGCAAAAAGAAAGGAAAAAAAA :  104
H7                      : ATGTCTCGTAAAAGGGTTTTTCTTCTGTGCATTTTTCTGGCCATTTCCGCCGTTGTTGACGAAGCCGAGTCGTTCAGGAGTAAGCAAAAAGAAAGGAAAAAAAA :  104
H8                      : ATGTCTCGTAAAAGGGTTTTTCTTCTGTGCATTTTTCTGGCCATTTCCGCCGTTGTTGACGAAGCCGAGTCGTTCAGGAGTAAGCAAAAAGAAAGGAAAAAAAA :  104
H9                      : ATGTCTCGTAAAAGGGTTTTTCTTCTGTGCATTTTTCTGGCCATTTCCGCCGTTGTTGACGAAGCCGAGTCGTTCAGGAGTAAGCAAAAAGAAAGGAAAAAAAA :  104
H10                     : ATGTCTCGTAAAAGGGTTTTTCTTCTGTGCATTTTTCTGGCCATTTCCGCCGTTGTTGACGAAGCCGAGTCGTTCAGGAGTAAGCAAAAAGAAAGGAAAAAAAA :  104
H11                     : ATGTCTCGTAAAAGGGTTTTTCTTCTGTGCATTTTTCTGGCCATTTCCGCCGTTGTTGACGAAGCCGAGTCGTTCAGGAGTAAGCAAAAAGAAAGGAAAAAAAA :  104
H12                     : ATGTCTCGTAAAAGGGTTTTTCTTCTGTGCATTTTTCTGGCCATTTCCGCCGTTGTTGACGAAGCCGAGTCGTTCAGGAGTAAGCAAAAAGAAAGGAAAAAAAA :  104
H13                     : ATGTCTCGTAAAAGGGTTTTTCTTCTGTGCATTTTTCTGGCCATTTCCGCCGTTGTTGACGAAGCCGAGTCGTTCAGGAGTAAGCAAAAAGAAAGGAAAAAAAA :  104
H14                     : ATGTCTCGTAAAAGGGTTTTTCTTCTGTGCATTTTTCTGGCCATTTCCGCCGTTGTTGACGAAGCCGAGTCGTTCAGGAGTAAGCAAAAAGAAAGGAAAAAAAA :  104
H15                     : ATGTCTCGTAAAAGGGTTTTTCTTCTGTGCATTTTTCTGGCCATTTCCGCCGTTGTTGACGAAGCCGAGTCGTTCAGGAGTAAGCAAAAAGAAAGGAAAAAAAA :  104
H16                     : ATGTCTCGTAAAAGGGTTTTTCTTCTGTGCATTTTTCTGGCCATTTCCGCCGTTGTTGACGAAGCCGAGTCGTTCAGGAGTAAGCAAAAAGAAAGGAAAAAAAA :  104
H17                     : ATGTCTCGTAAAAGGGTTTTTCTTCTGTGCATTTTTCTGGCCATTTCCGCCGTTGTTGACGAAGCCGAGTCGTTCAGGAGTAAGCAAAAAGAAAGGAAAAAAAA :  104
H18                     : ATGTCTCGTAAAAGGGTTTTTCTTCTGTGCATTTTTCTGGCCATTTCCGCCGTTGTTGACGAAGCCGAGTCGTTCAGGAGTAAGCAAAAAGAAAGGAAAAAAAA :  104
H19                     : ATGTCTCGTAAAAGGGTTTTTCTTCTGTGCATTTTTCTGGCCATTTCCGCCGTTGTTGACGAAGCCGAGTCGTTCAGGAGTAAGCAAAAAGAAAGGAAAAAAAA :  104
H20                     : ATGTCTCGTAAAAGGGTTTTTCTTCTGTGCATTTTTCTGGCCATTTCCGCCGTTGTTGACGAAGCCGAGTCGTTCAGGAGTAAGCAAAAAGAAAGGAAAAAAAA :  104
H21                     : ATGTCTCGTAAAAGGGTTTTTCTTCTGTGCATTTTTCTGGCCATTTCCGCCGTTGTTGACGAAGCCGAGTCGTTCAGGAGTAAGCAAAAAGAAAGGAAAAAAAA :  104
H22                     : ATGTCTCGTAAAAGGGTTTTTCTTCTGTGCATTTTTCTGGCCATTTCCGCCGTTGTTGACGAAGCCGAGTCGTTCAGGAGTAAGCAAAAAGAAAGGAAAAAAAA :  104
H23                     : ATGTCTCGTAAAAGGGTTTTTCTTCTGTGCATTTTTCTGGCCATTTCCGCCGTTGTTGACGAAGCCGAGTCGTTCAGGAGTAAGCAAAAAGAAAGGAAAAAAAA :  104
H24                     : ATGTCTCGTAAAAGGGTTTTTCTTCTGTGCATTTTTCTGGCCATTTCCGCCGTTGTTGACGAAGCCGAGTCGTTCAGGAGTAAGCAAAAAGAAAGGAAAAAAAA :  104
H25                     : ATGTCTCGTAAAAGGGTTTTTCTTCTGTGCATTTTTCTGGCCATTTCCGCCGTTGTTGACGAAGCCGAGTCGTTCAGGAGTAAGCAAAAAGAAAGGAAAAAAAA :  104
H26_Mauritania-I_ref    : ATGTCTCGTAAAAGGGTTTTTCTTCTGTGCATTTTTCTGGCCATTTCCGCCGTTGTTGACGAAGCCGAGTCGTTCAGGAGTAAGCAAAAAGAAAGGAAAAAAAA :  104
H27_India-VII_refseq    : ATGTCTCGTAAAAGGGTTTTTCTTCTGTGCATTTTTCTGGCCATTTCCGCCGTTGTTGACGAAGCCGAGTCGTTCAGGAGTAAGCAAAAAGAAAGGAAAAAAAA :  104
H28_Ctg_refsequence     : ATGTCTCGTAAAAGGGTTTTTCTTCTGTGCATTTTTCTGGCCATTTCCGCCGTTGTTGACGAAGCCGAGTCGTTCAGGAGTAAGCAAAAAGAAAGGAAAAAAAA :  104
H29_Salvador-I_refse    : ATGTCTCGTAAAAGGGTTTTTCTTCTGTGCATTTTTCTGGCCATTTCCGCCGTTGTTGACGAAGCCGAGTCGTTCAGGAGTAAGCAAAAAGAAAGGAAAAAAAA :  104
H30_NKorea_refsequen    : ATGTCTCGTAAAAGGGTTTTTCTTCTGTGCATTTTTCTGGCCATTTCCGCCGTTGTTGACGAAGCCGAGTCGTTCAGGAGTAAGCAAAAAGAAAGGAAAAAAAA :  104
H31_P.vivax01_refseq    : ATGTCTCGTAAAAGGGTTTTTCTTCTGTGCATTTTTCTGGCCATTTCCGCCGTTGTTGACGAAGCCGAGTCGTTCAGGAGTAAGCAAAAAGAAAGGAAAAAAAA :  104
H32_Brazil-I_refsequ    : ATGTCTCGTAAAAGGGTTTTTCTTCTGTGCATTTTTCTGGCCATTTCCGCCGTTGTTGACGAAGCCGAGTCGTTCAGGAGTAAGCAAAAAGAAAGGAAAAAAAA :  104
                          ATGTCTCGTAAAAGGGTTTTTCTTCTGTGCATTTTTCTGGCCATTTCCGCCGTTGTTGACGAAGCCGAGTCGTTCAGGAGTAAGCAAAAAGAAAGGAAAAAAAA
```

```
                                  *       120         *       140         *       160         *       180         *       200
H1_VCG                  : GGAGAGATAGGTAGAGAGGGGCATCACACGCAACTGCTATACGCATGCAACTGCGATACGCACGTGACTGCGGTACACACCTGTGCAATGGCGCGTGCCGTATG : 208
H2                      : GGAGAGATAGGTAGAGAGGGGCATCACACGCAACTGCTATACGCATGCAACTGCGATACGCACGTGACTGCGGTACACACCTGTGCAATGGCGCGTGCCGTATG : 208
H3                      : GGAGAGATAGGTAGAGAGGGGCATCACACGCAACTGCTATACGCATGCAACTGCGATACGCACGTGACTGCGGTACACACCTGTGCAATGGCGCGTGCCGTATG : 208
H4                      : GGAGAGATAGGTAGAGAGGGGCATCACACGCAACTGCTATACGCATGCAACTGCGATACGCACGTGACTGCGGTACACACCTGTGCAATGGCGCGTGCCGTATG : 208
H5                      : GGAGAGATAGGTAGAGAGGGGCATCACACGCAACTGCTATACGCATGCAACTGCGATACGCACGTGACTGCGGTACACACCTGTGCAATGGCGCGTGCCGTATG : 208
H6                      : GGAGAGATAGGTAGAGAGGGGCATCACACGCAACTGCTATACGCATGCAACTGCGATACGCACGTGACTGCGGTACACACCTGTGCAATGGCGCGTGCCGTATG : 208
H7                      : GGAGAGATAGGTAGAGAGGGGCATCACACGCAACTGCTATACGCATGCAACTGCGATACGCACGTGACTGCGGTACACACCTGTGCAATGGCGCGTGCCGTATG : 208
H8                      : GGAGAGATAGGTAGAGAGGGGCATCACACGCAACTGCTATACGCATGCAACTGCGATACGCACGTGACTGCGGTACACACCTGTGCAATGGCGCGTGCCGTATG : 208
H9                      : GGAGAGATAGGTAGAGAGGGGCATCACACGCAACTGCTATACGCATGCAACTGCG----------------GTACACACCTGTGCAATGGCGCGTGCCGTATG : 191
H10                     : GGAGAGATAGGTAGAGAGGGGCATCACACGCAACTGCTATACGCATGCAACTGCGATACGCACGTGACTGCGGTACACACCTGTGCAATGGCGCGTGCCGTATG : 208
H11                     : GGAGAGATAGGTAGAGAGGGGCATCACACGCAACTGCTATACGCATGCAACTGCGATACGCACGTGACTGCGGTACACACCTGTGCAATGGCGCGTGCCGTATG : 208
H12                     : GGAGAGATAGGTAGAGAGGGGCATCACACGCAACTGCTATACGCATGCAACTGCGATACGCACGTGACTGCGGTACACACCTGTGCAATGGCGCGTGCCGTATG : 208
H13                     : GGAGAGATAGGTAGAGAGGGGCATCACACGCAACTGCTATACGCATGCAACTGCGATACGCACGTGACTGCGGTACACACCTGTGCAATGGCGCGTGCCGTATG : 208
H14                     : GGAGAGATAGGTAGAGAGGGGCATCACACGCAACTGCTATACGCATGCAACTGCGATACGCACGTGACTGCGGTACACACCTGTGCAATGGCGCGTGCCGTATG : 208
H15                     : GGAGAGATAGGTAGAGAGGGGCATCACACGCAACTGCTATACGCATGCAACTGCGATACGCACGTGACTGCGGTACACACCTGTGCAATGGCGCGTGCCGTATG : 208
H16                     : GGAGAGATAGGTAGAGAGGGGCATCACACGCAACTGCTATACGCATGCAACTGCGATACGCACGTGACTGCGGTACACACCTGTGCAATGGCGCGTGCCGTATG : 208
H17                     : GGAGAGATAGGTAGAGAGGGGCATCACACGCAACTGCTATACGCATGCAACTGCGATACGCACGTGACTGCGGTACACACCTGTGCAATGGCGCGTGCCGTATG : 208
H18                     : GGAGAGATAGGTAGAGAGGGGCATCACACGCAACTGCTATACGCATGCAACTGCGATACGCACGTGACTGCGGTACACACCTGTGCAATGGCGCGTGCCGTATG : 208
H19                     : GGAGAGATAGGTAGAGAGGGGCATCACACGCAACTGCTATACGCATGCAACTGCGATACGCACGTGACTGCGGTACACACCTGTGCAATGGCGCGTGCCGTATG : 208
H20                     : GGAGAGATAGGTAGAGAGGGGCATCACACGCAACTGCTATACGCATGCAACTGCGATACGCACGTGACTGCGGTACACACCTGTGCAATGGCGCGTGCCGTATG : 208
H21                     : GGAGAGATAGGTAGAGAGGGGCATCACACGCAACTGCTATACGCATGCAACTGCGATACGCACGTGACTGCGGTACACACCTGTGCAATGGCGCGTGCCGTATG : 208
H22                     : GGAGAGATAGGTAGAGAGGGGCATCACACGCAACTGCTATACGCATGCAACTGCGATACGCACGTGACTGCGGTACACACCTGTGCAATGGCGCGTGCCGTATG : 208
H23                     : GGAGAGATAGGTAGAGAGGGGCATCACACGCAACTGCTATACGCATGCAACTGCGATACGCACGTGACTGCGGTACACACCTGTGCAATGGCGCGTGCCGTATG : 208
H24                     : GGAGAGATAGGTAGAGAGGGGCATCACACGCAACTGCTATACGCATGCAACTGCGATACGCACGTGACTGCGGTACACACCTGTGCAATGGCGCGTGCCGTATG : 208
H25                     : GGAGAGATAGGTAGAGAGGGGCATCACACGCAACTGCTATACGCATGCAACTGCGATACGCACGTGACTGCGGTACACACCTGTGCAATGGCGCGTGCCGTATG : 208
H26_Mauritania-I_ref    : GGAGAGATAGGTAGAGAGGGGCATCACACGCAACTGCTATACGCATGCAACTGCGATACGCACGTGACTGCGGTACACACCTGTGCAATGGCGCGTGCCGTATG : 208
H27_India-VII_refseq    : GGAGAGATAGGTAGAGAGGGGCATCACACGCAACTGCTATACGCATGCAACTGCGATACGCACGTGACTGCGGTACACACCTGTGCAATGGCGCGTGCCGTATG : 208
H28_Ctg_refsequence     : GGAGAGATAGGTAGAGAGGGGCATCACACGCAACTGCTATACGCATGCAACTGCGATACGCACGTGACTGCGGTACACACCTGTGCAATGGCGCGTGCCGTATG : 208
H29_Salvador-I_refse    : GGAGAGATAGGTAGAGAGGGGCATCACACGCAACTGCTATACGCATGCAACTGCGATACGCACGTGACTGCGGTACACACCTGTGCAATGGCGCGTGCCGTATG : 208
H30_NKorea_refsequen    : GGAGAGATAGGTAGAGAGGGGCATCACACGCAACTGCTATACGCATGCAACTGCGATACGCACGTGACTGCGGTACACACCTGTGCAATGGCGCGTGCCGTATG : 208
H31_P.vivax01_refseq    : GGAGAGATAGGTAGAGAGGGGCATCACACGCAACTGCTATACGCATGCAACTGCGATACGCACGTGACTGCGGTACACACCTGTGCAATGGCGCGTGCCGTATG : 208
H32_Brazil-I_refsequ    : GGAGAGATAGGTAGAGAGGGGCATCACACGCAACTGCTATACGCATGCAACTGCGATACGCACGTGACTGCGGTACACACCTGTGCAATGGCGCGTGCCGTATG : 208
                          GGAGAGATAGGTAGAGAGGGGCATCACACGCAACTGCTATACGCATGCAACTGCGatacgcacgtgactgcgGTACACACCTGTGCAATGGCGCGTGCCGTATG

                          *       220         *       240         *       260         *       280         *       300         *
H1_VCG                  : CATGCTGTCCACGGCGCTCCACCCCCAAAGCAACGAAATAAAAAATAAAAAGTAAATAAATAAATAAATAAATTCCCCCTTTCAGAGCCGGACGCGGCAGCGAA : 312
H2                      : CATGCTGTCCACGGCGCTCCACCCCCAAAGCAACGAAATAAAAAATAAAAAGTAAATAAATAAATAAATAAATTCCCCCTTTCAGAGCCGGACGCGGCAGCGAA : 312
H3                      : CATGCTGTCCACGGCGCTCCACCCCCAAAGCAACGAAATAAAAAATAAAAAGTAAATAAATAAATAAATAAATTCCCCCTTTCAGAGCCAGACGCGGCAGCGAA : 312
H4                      : CATGCTGTCCACGGCGCTCCACCCCCAAAGCAACGAAATAAAAAATAAAAAGTAAATAAATAAATAAATAAATTCCCCCTTTCAGAGCCAGACGCGGCAGCGAA : 312
H5                      : CATGCTGTCCACGGCGCTCCACCCCCAAAGCAACGAAATAAAAAATAAAAAGTAAATAAATAAATAAATAAATTCCCCCTTTCAGAGCCGGACGCGGCAGCGAA : 312
H6                      : CATGCTGTCCACGGCGCTCCACCCCCAAAGCAACGAAATAAAAAATAAAAAGTAAATAAATAAATAAATAAATTCCCCCTTTCAGAGCCGGACGCGGCAGCGAA : 312
H7                      : CATGCTGTCCACGGCGCTCCACCCCCAAAGCAACGAAATAAAAAATAAAAAGTAAATAAATAAATAAATAAATTCCCCCTTTCAGAGCCAGACGCGGCAGCGAA : 312
H8                      : CATGCTGTCCACGGCGCTCCACCCCCAAAGCAACGAAATAAAAAATAAAAAGTAAATAAATAAATAAATAAATTCCCCCTTTCAGAGCCAGACGCGGCAGCGAA : 312
H9                      : CATGCTGTCCACGGCGCTCCACCCCCAAAGCAACGAAATAAAAAATAAAAAGTAAATAAATAAATAAATAAATTCCCCCTTTCAGAGCCGGACGCGGCAGCGAA : 295
```

```
H10                  : CATGCTGTCCACGGCGCTCCACCCCCAAAGCAACGAAATAAAAAATAAAAAGTAAATAAATAAATAAATAAATTCCCCCTTTCAGAGCCAGACGCGGCAGCGAA :  312
H11                  : CATGCTGTCCACGGCGCTCCACCCCCAAAGCAACGAAATAAAAAATAAAAAGTAAATAAATAAATAAATAAATTCCCCCTTTCAGAGCCAGACGCGGCAGCGAA :  312
H12                  : CATGCTGTCCACGGCGCTCCACCCCCAAAGCAACGAAATAAAAAATAAAAAGTAAATAAATAAATAAATAAATTCCCCCTTTCAGAGCCAGACGCGGCAGCGAA :  312
H13                  : CATGCTGTCCACGGCGCTCCACCCCCAAAGCAACGAAATAAAAAATAAAAAGTAAATAAATAAATAAATAAATTCCCCCTTTCAGAGCCGGACGCGGCAGCGAA :  312
H14                  : CATGCTGTCCACGGCGCTCCACCCCCAAAGCAACGAAATAAAAAATAAAAAGTAAATAAATAAATAAATAAATTCCCCCTTTCAGAGCCGGACGCGGCAGCGAA :  312
H15                  : CATGCTGTCCACGGCGCTCCACCCCCAAAGCAACGAAATAAAAAATAAAAAGTAAATAAATAAATAAATAAATTCCCCCTTTCAGAGCCAGACGCGGCAGCGAA :  312
H16                  : CATGCTGTCCACGGCGCTCCACCCCCAAAGCAACGAAATAAAAAATAAAAAGTAAATAAATAAATAAATAAATTCCCCCTTTCAGAGCCGGACGCGGCAGCGAA :  312
H17                  : CATGCTGTCCACGGCGCTCCACCCCCAAAGCAACGAAATAAAAAATAAAAAGTAAATAAATAAATAAATAAATTCCCCCTTTCAGAGCCGGACGCGGCAGCGAA :  312
H18                  : CATGCTGTCCACGGCGCTCCACCCCCAAAGCAACGAAATAAAAAATAAAAAGTAAATAAATAAATAAATAAATTCCCCCTTTCAGAGCCGGACGCGGCAGCGAA :  312
H19                  : CATGCTGTCCACGGCGCTCCACCCCCAAAGCAACGAAATAAAAAATAAAAAGTAAATAAATAAATAAATAAATTCCCCCTTTCAGAGCCAGACGCGGCAGCGAA :  312
H20                  : CATGCTGTCCACGGCGCTCCACCCCCAAAGCAACGAAATAAAAAATAAAAAGTAAATAAATAAATAAATAAATTCCCCCTTTCAGAGCCGGACGCGGCAGCGAA :  312
H21                  : CATGCTGTCCACGGCGCTCCACCCCCAAAGCAACGAAATAAAAAATAAAAAGTAAATAAATAAATAAATAAATTCCCCCTTTCAGAGCCAGACGCGGCAGCGAA :  312
H22                  : CATGCTGTCCACGGCGCTCCACCCCCAAAGCAACGAAATAAAAAATAAAAAGTAAATAAATAAATAAATAAATTCCCCCTTTCAGAGCCAGACGCGGCAGCGAA :  312
H23                  : CATGCTGTCCACGGCGCTCCACCCCCAAAGCAACGAAATAAAAAATAAAAAGTAAATAAATAAATAAATAAATTCCCCCTTTCAGAGCCAGACGCGGCAGCGAA :  312
H24                  : CATGCTGTCCACGGCGCTCCACCCCCAAAGCAACGAAATAAAAAATAAAAAGTAAATAAATAAATAAATAAATTCCCCCTTTCAGAGCCGGACGCGGCAGCGAA :  312
H25                  : CATGCTGTCCACGGCGCTCCACCCCCAAAGCAACGAAATAAAAAATAAAAAGTAAATAAATAAATAAATAAATTCCCCCTTTCAGAGCCAGACGCGGCAGCGAA :  312
H26_Mauritania-I_ref : CATGCTGTCCACGGCGCTCCACCCCCAAAGCAACGAAATAAAAAATAAAAAGTAAATAAATAAATAAATAAATTCCCCCTTTCAGAGCCGGACGCGGCAGCGAA :  312
H27_India-VII_refseq : CATGCTGTCCACGGCGCTCCACCCCCAAAGCAACGAAATAAAAAATAAAAAGTAAATAAATAAATAAATAAATTCCCCCTTTCAGAGCCGGACGCGGCAGCGAA :  312
H28_Ctg_refsequence  : CATGCTGTCCACGGCGCTCCACCCCCAAAGCAACGAAATAAAAAATAAAAAGTAAATAAATAAATAAATAAATTCCCCCTTTCAGAGCCAGACGCGGCAGCGAA :  312
H29_Salvador-I_refse : CATGCTGTCCACGGCGCTCCACCCCCAAAGCAACGAAATAAAAAATAAAAAGTAAATAAATAAATAAATAAATTCCCCCTTTCAGAGCCAGACGCGGCAGCGAA :  312
H30_NKorea_refsequen : CATGCTGTCCACGGCGCTCCACCCCCAAAGCAACGAAATAAAAAATAAAAAGTAAATAAATAAATAAATAAATTCCCCCTTTCAGAGCCGGACGCGGCAGCGAA :  312
H31_P.vivax01_refseq : CATGCTGTCCACGGCGCTCCACCCCCAAAGCAACGAAATAAAAAATAAAAAGTAAATAAATAAATAAATAAATTCCCCCTTTCAGAGCCGGACGCGGCAGCGAA :  312
H32_Brazil-I_refsequ : CATGCTGTCCACGGCGCTCCACCCCCAAAGCAACGAAATAAAAAATAAAAAGTAAATAAATAAATAAATAAATTCCCCCTTTCAGAGCCGGACGCGGCAGCGAA :  312
                       CATGCTGTCCACGGCGCTCCACCCCCAAAGCAACGAAATAAAAAATAAAAAGTAAATAAATAAATAAATAAATTCCCCCTTTCAGAGCC GACGCGGCAGCGAA

                                 320         *       340         *       360         *       380         *       400         *
H1_VCG               : CGTCTCCTTCCTGGAGTTCACAGCAGATGGCAACGCTGAAGAGGGGCAGGCGCAGCCAAGTGAACCCATCGTTAACACCCAGCCGCTGCAGGAAAATGTAAGTG :  416
H2                   : CGTCTCCTTCCTGGAGTTCACAGCAGATGGCAACGCTGAAGAGGGGCAGGCGCAGCCAAGTGAACCCATCGTTAACACCCAGCCGCTGCAGGAAAATGTAAGTG :  416
H3                   : CGTCTCCTTCCTGGAGTTCACAGCAGATGGCAACGCTGAAGAGGGGCAGGCGCAGCCAAGTGAACCCATCGTTAACACCCAGCCGCTGCAGGAAAATGTAAGTG :  416
H4                   : CGTCTCCTTCCTGGAGTTCACAGCAGATGGCAACGCTGAAGAGGGGCAGGCGCAGCCAAGTGAACCCATCGTTAACACCCAGCCGCTGCAGGAAAATGTAAGTG :  416
H5                   : CGTCTCCTTCCTGGAGTTCACAGCAGATGGCAACGCTGAAGAGGGGCAGGCGCAGCCAAGTGAACCCATCGTTAACACCCAGCCGCTGCAGGAAAATGTAAGTG :  416
H6                   : CGTCTCCTTCCTGGAGTTCACAGCAGATGGCAACGCTGAAGAGGGGCAGGCGCAGCCAAGTGAACCCATCGTTAACACCCAGCCGCTGCAGGAAAATGTAAGTG :  416
H7                   : CGTCTCCTTCCTGGAGTTCACAGCAGATGGCAACGCTGAAGAGGGGCAGGCGCAGCCAAGTGAACCCATCGTTAACACCCAGCCGCTGCAGGAAAATGTAAGTG :  416
H8                   : CGTCTCCTTCCTGGAGTTCACAGCAGATGGCAACGCTGAAGAGGGGCAGGCGCAGCCAAGTGAACCCATCGTTAACACCCAGCCGCTGCAGGAAAATGTAAGTG :  416
H9                   : CGTCTCCTTCCTGGAGTTCACAGCAGATGGCAACGCTGAAGAGGGGCAGGCGCAGCCAAGTGAACCCATCGTTAACACCCAGCCGCTGCAGGAAAATGTAAGTG :  399
H10                  : CGTCTCCTTCCTGGAGTTCACAGCAGATGGCAACGCTGAAGAGGGGCAGGCGCAGCCAAGTGAACCCATCGTTAACACCCAGCCGCTGCAGGAAAATGTAAGTG :  416
H11                  : CGTCTCCTTCCTGGAGTTCACAGCAGATGGCAACGCTGAAGAGGGGCAGGCGCAGCCAAGTGAACCCATCGTTAACACCCAGCCGCTGCAGGAAAATGTAAGTG :  416
H12                  : CGTCTCCTTCCTGGAGTTCACAGCAGATGGCAACGCTGAAGAGGGGCAGGCGCAGCCAAGTGAACCCATCGTTAACACCCAGCCGCTGCAGGAAAATGTAAGTG :  416
H13                  : CGTCTCCTTCCTGGAGTTCACAGCAGATGGCAACGCTGAAGAGGGGCAGGCGCAGCCAAGTGAACCCATCGTTAACACCCAGCCGCTGCAGGAAAATGTAAGTG :  416
H14                  : CGTCTCCTTCCTGGAGTTCACAGCAGATGGCAACGCTGAAGAGGGGCAGGCGCAGCCAAGTGAACCCATCGTTAACACCCAGCCGCTGCAGGAAAATGTAAGTG :  416
H15                  : CGTCTCCTTCCTGGAGTTCACAGCAGATGGCAACGCTGAAGAGGGGCAGGCGCAGCCAAGTGAACCCATCGTTAACACCCAGCCGCTGCAGGAAAATGTAAGTG :  416
H16                  : CGTCTCCTTCCTGGAGTTCACAGCAGATGGCAACGCTGAAGAGGGGCAGGCGCAGCCAAGTGAACCCATCGTTAACACCCAGCCGCTGCAGGAAAATGTAAGTG :  416
H17                  : CGTCTCCTTCCTGGAGTTCACAGCAGATGGCAACGCTGAAGAGGGGCAGGCGCAGCCAAGTGAACCCATCGTTAACACCCAGCCGCTGCAGGAAAATGTAAGTG :  416
H18                  : CGTCTCCTTCCTGGAGTTCACAGCAGATGGCAACGCTGAAGAGGGGCAGGCGCAGCCAAGTGAACCCATCGTTAACACCCAGCCGCTGCAGGAAAATGTAAGTG :  416
H19                  : CGTCTCCTTCCTGGAGTTCACAGCAGATGGCAACGCTGAAGAGGGGCAGGCGCAGCCAAGTGAACCCATCGTTAACACCCAGCCGCTGCAGGAAAATGTAAGTG :  416
```

```
H20                  : CGTCTCCTTCCTGGAGTTCACAGCAGATGGCAACGCTGAAGAGGGGCAGGCGCAGCCAAGTGAACCCATCGTTAACACCCAGCCGCTGCAGGAAAATGTAAGTG :  416
H21                  : CGTCTCCTTCCTGGAGTTCACAGCAGATGGCAACGCTGAAGAGGGGCAGGCGCAGCCAAGTGAACCCATCGTTAACACCCAGCCGCTGCAGGAAAATGTAAGTG :  416
H22                  : CGTCTCCTTCCTGGAGTTCACAGCAGATGGCAACGCTGAAGAGGGGCAGGCGCAGCCAAGTGAACCCATCGTTAACACCCAGCCGCTGCAGGAAAATGTAAGTG :  416
H23                  : CGTCTCCTTCCTGGAGTTCACAGCAGATGGCAACGCTGAAGAGGGGCAGGCGCAGCCAAGTGAACCCATCGTTAACACCCAGCCGCTGCAGGAAAATGTAAGTG :  416
H24                  : CGTCTCCTTCCTGGAGTTCACAGCAGATGGCAACGCTGAAGAGGGGCAGGCGCAGCCAAGTGAACCCATCGTTAACACCCAGCCGCTGCAGGAAAATGTAAGTG :  416
H25                  : CGTCTCCTTCCTGGAGTTCACAGCAGATGGCAACGCTGAAGAGGGGCAGGCGCAGCCAAGTGAACCCATCGTTAACACCCAGCCGCTGCAGGAAAATGTAAGTG :  416
H26_Mauritania-I_ref : CGTCTCCTTCCTGGAGTTCACAGCAGATGGCAACGCTGAAGAGGGGCAGGCGCAGCCAAGTGAACCCATCGTTAACACCCAGCCGCTGCAGGAAAATGTAAGTG :  416
H27_India-VII_refseq : CGTCTCCTTCCTGGAGTTCACAGCAGATGGCAACGCTGAAGAGGGGCAGGCGCAGCCAAGTGAACCCATCGTTAACACCCAGCCGCTGCAGGAAAATGTAAGTG :  416
H28_Ctg_refsequence  : CGTCTCCTTCCTGGAGTTCACAGCAGATGGCAACGCTGAAGAGGGGCAGGCGCAGCCAAGTGAACCCATCGTTAACACCCAGCCGCTGCAGGAAAATGTAAGTG :  416
H29_Salvador-I_refse : CGTCTCCTTCCTGGAGTTCACAGCAGATGGCAACGCTGAAGAGGGGCAGGCGCAGCCAAGTGAACCCATCGTTAACACCCAGCCGCTGCAGGAAAATGTAAGTG :  416
H30_NKorea_refsequen : CGTCTCCTTCCTGGAGTTCACAGCAGATGGCAACGCTGAAGAGGGGCAGGCGCAGCCAAGTGAACCCATCGTTAACACCCAGCCGCTGCAGGAAAATGTAAGTG :  416
H31_P.vivax01_refseq : CGTCTCCTTCCTGGAGTTCACAGCAGATGGCAACGCTGAAGAGGGGCAGGCGCAGCCAAGTGAACCCATCGTTAACACCCAGCCGCTGCAGGAAAATGTAAGTG :  416
H32_Brazil-I_refsequ : CGTCTCCTTCCTGGAGTTCACAGCAGATGGCAACGCTGAAGAGGGGCAGGCGCAGCCAAGTGAACCCATCGTTAACACCCAGCCGCTGCAGGAAAATGTAAGTG :  416
                       CGTCTCCTTCCTGGAGTTCACAGCAGATGGCAACGCTGAAGAGGGGCAGGCGCAGCCAAGTGAACCCATCGTTAACACCCAGCCGCTGCAGGAAAATGTAAGTG

                        420         *       440         *       460         *       480         *       500         *       520
H1_VCG               : AAGCCGCTCCCCACTCGGAGGAGTCCACTGGGGGAAGCGCGGAGAAGGGAAGTGGCGGCCACGAGGAAACCCCCCACGCAGGAGAACATGACAAAGGTAGTGGC :  520
H2                   : AAGCCGCTCCCCACTCGGAGGAGTCCACTGGGGGAAGCGCGGAGAAGGGAAGTGGCGGCCACGAGGAAACCCCCCACGCAGGAGAACATGACAAAGGTAGTGGC :  520
H3                   : AAGCCGCTCCCCACTCGGAGGAGTCCACTGGGGGAAGCGCGGAGAAGGGAAGTGGCGGCCACGAGGAAACCCCCCACGCAGGAGAACATGACAAAGGTAGTGGC :  520
H4                   : AAGCCGCTCCCCACTCGGAGGAGTCCACTGGGGGAAGCGCGGAGAAGGGAAGTGGCGGCCACGAGGAAACCCCCCACGCAGGAGAACATGACAAAGGTAGTGGC :  520
H5                   : AAGCCGCTCCCCACTCGGAGGAGTCCACTGGGGGAAGCGCGGAGAAGGGAAGTGGCGGCCACGAGGAAACCCCCCACGCAGGAGAACATGACAAAGGTAGTGGC :  520
H6                   : AAGCCGCTCCCCACTCGGAGGAGTCCACTGGGGGAAGCGCGGAGAAGGGAAGTGGCGGCCACGAGGAAACCCCCCACGCAGGAGAACATGACAAAGGTAGTGGC :  520
H7                   : AAGCCGCTCCCCACTCGGAGGAGTCCACTGGGGGAAGCGCGGAGAAGGGAAGTGGCGGCCACGAGGAAACCCCCCACGCAGGAGAACATGACAAAGGTAGTGGC :  520
H8                   : AAGCCGCTCCCCACTCGGAGGAGTCCACTGGGGGAAGCGCGGAGAAGGGAAGTGGCGGCCACGAGGAAACCCCCCACGCAGGAGAACATGACAAAGGTAGTGGC :  520
H9                   : AAGCCGCTCCCCACTCGGAGGAGTCCACTGGGGGAAGCGCGGAGAAGGGAAGTGGCGGCCACGAGGAAACCCCCCACGCAGGAGAACATGACAAAGGTAGTGGC :  503
H10                  : AAGCCGCTCCCCACTCGGAGGAGTCCACTGGGGGAAGCGCGGAGAAGGGAAGTGGCGGCCACGAGGAAACCCCCCACGCAGGAGAACATGACAAAGGTAGTGGC :  520
H11                  : AAGCCGCTCCCCACTCGGAGGAGTCCACTGGGGGAAGCGCGGAGAAGGGAAGTGGCGGCCACGAGGAAACCCCCCACGCAGGAGAACATGACAAAGGTAGTGGC :  520
H12                  : AAGCCGCTCCCCACTCGGAGGAGTCCACTGGGGGAAGCGCGGAGAAGGGAAGTGGCGGCCACGAGGAAACCCCCCACGCAGGAGAACATGACAAAGGTAGTGGC :  520
H13                  : AAGCCGCTCCCCACTCGGAGGAGTCCACTGGGGGAAGCGCGGAGAAGGGAAGTGGCGGCCACGAGGAAACCCCCCACGCAGGAGAACATGACAAAGGTAGTGGC :  520
H14                  : AAGCCGCTCCCCACTCGGAGGAGTCCACTGGGGGAAGCGCGGAGAAGGGAAGTGGCGGCCACGAGGAAACCCCCCACGCAGGAGAACATGACAAAGGTAGTGGC :  520
H15                  : AAGCCGCTCCCCACTCGGAGGAGTCCACTGGGGGAAGCGCGGAGAAGGGAAGTGGCGGCCACGAGGAAACCCCCCACGCAGGAGAACATGACAAAGGTAGTGGC :  520
H16                  : AAGCCGCTCCCCACTCGGAGGAGTCCACTGGGGGAAGCGCGGAGAAGGGAAGTGGCGGCCACGAGGAAACCCCCCACGCAGGAGAACATGACAAAGGTAGTGGC :  520
H17                  : AAGCCGCTCCCCACTCGGAGGAGTCCACTGGGGGAAGCGCGGAGAAGGGAAGTGGCGGCCACGAGGAAACCCCCCACGCAGGAGAACATGACAAAGGTAGTGGC :  520
H18                  : AAGCCGCTCCCCACTCGGAGGAGTCCACTGGGGGAAGCGCGGAGAAGGGAAGTGGCGGCCACGAGGAAACCCCCCACGCAGGAGAACATGACAAAGGTAGTGGC :  520
H19                  : AAGCCGCTCCCCACTCGGAGGAGTCCACTGGGGGAAGCGCGGAGAAGGGAAGTGGCGGCCACGAGGAAACCCCCCACGCAGGAGAACATGACAAAGGTAGTGGC :  520
H20                  : AAGCCGCTCCCCACTCGGAGGAGTCCACTGGGGGAAGCGCGGAGAAGGGAAGTGGCGGCCACGAGGAAACCCCCCACGCAGGAGAACATGACAAAGGTAGTGGC :  520
H21                  : AAGCCGCTCCCCACTCGGAGGAGTCCACTGGGGGAAGCGCGGAGAAGGGAAGTGGCGGCCACGAGGAAACCCCCCACGCAGGAGAACATGACAAAGGTAGTGGC :  520
H22                  : AAGCCGCTCCCCACTCGGAGGAGTCCACTGGGGGAAGCGCGGAGAAGGGAAGTGGCGGCCACGAGGAAACCCCCCACGCAGGAGAACATGACAAAGGTAGTGGC :  520
H23                  : AAGCCGCTCCCCACTCGGAGGAGTCCACTGGGGGAAGCGCGGAGAAGGGAAGTGGCGGCCACGAGGAAACCCCCCACGCAGGAGAACATGACAAAGGTAGTGGC :  520
H24                  : AAGCCGCTCCCCACTCGGAGGAGTCCACTGGGGGAAGCGCGGAGAAGGGAAGTGGCGGCCACGAGGAAACCCCCCACGCAGGAGAACATGACAAAGGTAGTGGC :  520
H25                  : AAGCCGCTCCCCACTCGGAGGAGTCCACTGGGGGAAGCGCGGAGAAGGGAAGTGGCGGCCACGAGGAAACCCCCCACGCAGGAGAACATGACAAAGGTAGTGGC :  520
H26_Mauritania-I_ref : AAGCCGCTCCCCACTCGGAGGAGTCCACTGGGGGAAGCGCGGAGAAGGGAAGTGGCGGCCACGAGGAAACCCCCCACGCAGGAGAACATGACAAAGGTAGTGGC :  520
H27_India-VII_refseq : AAGCCGCTCCCCACTCGGAGGAGTCCACTGGGGGAAGCGCGGAGAAGGGAAGTGGCGGCCACGAGGAAACCCCCCACGCAGGAGAACGTGACAAAGGTAGTGGC :  520
H28_Ctg_refsequence  : AAGCCGCTCCCCACTCGGAGGAGTCCACTGGGGGAAGCGCGGAGAAGGGAAGTGGCGGCCACGAGGAAACCCCCCACGCAGGAGAACATGACAAAGGTAGTGGC :  520
H29_Salvador-I_refse : AAGCCGCTCCCCACTCGGAGGAGTCCACTGGGGGAAGCGCGGAGAAGGGAAGTGGCGGCCACGAGGAAACCCCCCACGCAGGAGAACATGACAAAGGTAGTGGC :  520
```

```
H30_NKorea_refsequen : AAGCCGCTCCCCACTCGGAGGAGTCCACTGGGGGAAGCGCGGAGAAGGGAAGTGGCGGCCACGAGGAAACCCCCCACGCAGGAGAACATGACAAAGGTAGTGGC :  520
H31_P.vivax01_refseq : AAGCCGCTCCCCACTCGGAGGAGTCCACTGGGGGAAGCGCGGAGAAGGGAAGTGGCGGCCACGAGGAAACCCCCCACGCAGGAGAACATGACAAAGGTAGTGGC :  520
H32_Brazil-I_refsequ : AAGCCGCTCCCCACTCGGAGGAGTCCACTGGGGGAAGCGCGGAGAAGGGAAGTGGCGGCCACGAGGAAACCCCCCACGCAGGAGAACATGACAAAGGTAGTGGC :  520
                       AAGCCGCTCCCCACTCGGAGGAGTCCACTGGGGGAAGCGCGGAGAAGGGAAGTGGCGGCCACGAGGAAACCCCCCACGCAGGAGAACaTGACAAAGGTAGTGGC

                                 *       540         *       560         *       580         *       600         *       620
H1_VCG               : GGAAGCAACGGGAGCAGCGAAAGCAACGGAAGCGTCTTGGCTGATGGCCACGCCGAAGAAGCGCCGAACGAGCAGGTGAGTAGCGAGAAGAACGACGGACACAG :  624
H2                   : GGAAGCAACGGGAGCAGCGAAAGCAACGGAAGCGTCTTGGCTGATGGCCACGCCGAAGAAGCGCCGAACGAGCAGGTGAGTAGCGAGAAGAACGACGGACACAG :  624
H3                   : GGAAGCAACGGGAGCAGCGAAAGCAACGGAAGCGTCTTGGCTGATGGCCACGCCGAAGAAGCGCCGAACGAGCAGGTGAGTAGCGAGAAGAACGACGGACACAG :  624
H4                   : GGAAGCAACGGGAGCAGCGAAAGCAACGGAAGCGTCTTGGCTGATGGCCACGCCGAAGAAGCGCCGAACGAGCAGGTGAGTAGCGAGAAGAACGACGGACACAG :  624
H5                   : GGAAGCAACGGGAGCAGCGAAAGCAACGGAAGCGTCTTGGCTGATGGCCACGCCGAAGAAGCGCCGAACGAGCAGGTGAGTAGCGAGAAGAACGACGGACACAG :  624
H6                   : GGAAGCAACGGGAGCAGCGAAAGCAACGGAAGCGTCTTGGCTGATGGCCACGCCGAAGAAGCGCCGAACGAGCAGGTGAGTAGCGAGAAGAACGACGGACACAG :  624
H7                   : GGAAGCAACGGGAGCAGCGAAAGCAACGGAAGCGTCTTGGCTGATGGCCACGCCGAAGAAGCGCCGAACGAGCAGGTGAGTAGCGAGAAGAACGACGGACACAG :  624
H8                   : GGAAGCAACGGGAGCAGCGAAAGCAACGGAAGCGTCTTGGCTGATGGCCACGCCGAAGAAGCGCCGAACGAGCAGGTGAGTAGCGAGAAGAACGACGGACACAG :  624
H9                   : GGAAGCAACGGGAGCAGCGAAAGCAACGGAAGCGTCTTGGCTGATGGCCACGCCGAAGAAGCGCCGAACGAGCAGGTGAGTAGCGAGAAGAACGACGGACACAG :  607
H10                  : GGAAGCAACGGGAGCAGCGAAAGCAACGGAAGCGTCTTGGCTGATGGCCACGCCGAAGAAGCGCCGAACGAGCAGGTGAGTAGCGAGAAGAACGACGGACACAG :  624
H11                  : GGAAGCAACGGGAGCAGCGAAAGCAACGGAAGCGTCTTGGCTGATGGCCACGCCGAAGAAGCGCCGAACGAGCAGGTGAGTAGCGAGAAGAACGACGGACACAG :  624
H12                  : GGAAGCAACGGGAGCAGCGAAAGCAACGGAAGCGTCTTGGCTGATGGCCACGCCGAAGAAGCGCCGAACGAGCAGGTGAGTAGCGAGAAGAACGACGGACACAG :  624
H13                  : GGAAGCAACGGGAGCAGCGAAAGCAACGGAAGCGTCTTGGCTGATGGCCACGCCGAAGAAGCGCCGAACGAGCAGGTGAGTAGCGAGAAGAACGACGGACACAG :  624
H14                  : GGAAGCAACGGGAGCAGCGAAAGCAACGGAAGCGTCTTGGCTGATGGCCACGCCGAAGAAGCGCCGAACGAGCAGGTGAGTAGCGAGAAGAACGACGGACACAG :  624
H15                  : GGAAGCAACGGGAGCAGCGAAAGCAACGGAAGCGTCTTGGCTGATGGCCACGCCGAAGAAGCGCCGAACGAGCAGGTGAGTAGCGAGAAGAACGACGGACACAG :  624
H16                  : GGAAGCAACGGGAGCAGCGAAAGCAACGGAAGCGTCTTGGCTGATGGCCACGCCGAAGAAGCGCCGAACGAGCAGGTGAGTAGCGAGAAGAACGACGGACACAG :  624
H17                  : GGAAGCAACGGGAGCAGCGAAAGCAACGGAAGCGTCTTGGCTGATGGCCACGCCGAAGAAGCGCCGAACGAGCAGGTGAGTAGCGAGAAGAACGACGGACACAG :  624
H18                  : GGAAGCAACGGGAGCAGCGAAAGCAACGGAAGCGTCTTGGCTGATGGCCACGCCGAAGAAGCGCCGAACGAGCAGGTGAGTAGCGAGAAGAACGACGGACACAG :  624
H19                  : GGAAGCAACGGGAGCAGCGAAAGCAACGGAAGCGTCTTGGCTGATGGCCACGCCGAAGAAGCGCCGAACGAGCAGGTGAGTAGCGAGAAGAACGACGGACACAG :  624
H20                  : GGAAGCAACGGGAGCAGCGAAAGCAACGGAAGCGTCTTGGCTGATGGCCACGCCGAAGAAGCGCCGAACGAGCAGGTGAGTAGCGAGAAGAACGACGGACACAG :  624
H21                  : GGAAGCAACGGGAGCAGCGAAAGCAACGGAAGCGTCTTGGCTGATGGCCACGCCGAAGAAGCGCCGAACGAGCAGGTGAGTAGCGAGAAGAACGACGGACACAG :  624
H22                  : GGAAGCAACGGGAGCAGCGAAAGCAACGGAAGCGTCTTGGCTGATGGCCACGCCGAAGAAGCGCCGAACGAGCAGGTGAGTAGCGAGAAGAACGACGGACACAG :  624
H23                  : GGAAGCAACGGGAGCAGCGAAAGCAACGGAAGCGTCTTGGCTGATGGCCACGCCGAAGAAGCGCCGAACGAGCAGGTGAGTAGCGAGAAGAACGACGGACACAG :  624
H24                  : GGAAGCAACGGGAGCAGCGAAAGCAACGGAAGCGTCTTGGCTGATGGCCACGCCGAAGAAGCGCCGAACGAGCAGGTGAGTAGCGAGAAGAACGACGGACACAG :  624
H25                  : GGAAGCAACGGGAGCAGCGAAAGCAACGGAAGCGTCTTGGCTGATGGCCACGCCGAAGAAGCGCCGAACGAGCAGGTGAGTAGCGAGAAGAACGACGGACACAG :  624
H26_Mauritania-I_ref : GG------------------AAGCAACGGAAGCGTCTTGGCTGATGGCCACGCCGAAGAAGCGCCGAACGAGCAGGTGAGTAGCGAGAAGAACGACGGACACAG :  606
H27_India-VII_refseq : GGAAGCAACGGGAGCAGCGAAAGCAACGGAAGCGTCTTGGCTGATGGCCACGCCGAAGAAGCGCCGAACGAGCAGGTGAGTAGCGAGAAGAACGACGGACACAG :  624
H28_Ctg_refsequence  : GGAAGCAACGGGAGCAGCGAAAGCAACGGAAGCGTCTTGGCTGATGGCCACGCCGAAGAAGCGCCGAACGAGCAGGTGAGTAGCGAGAAGAACGACGGACACAG :  624
H29_Salvador-I_refse : GGAAGCAACGGGAGCAGCGAAAGCAACGGAAGCGTCTTGGCTGATGGCCACGCCGAAGAAGCGCCGAACGAGCAGGTGAGTAGCGAGAAGAACGACGGACACAG :  624
H30_NKorea_refsequen : GGAAGCAACGGGAGCAGCGAAAGCAACGGAAGCGTCTTGGCTGATGGCCACGCCGAAGAAGCGCCGAACGAGCAGGTGAGTAGCGAGAAGAACGACGGACACAG :  624
H31_P.vivax01_refseq : GGAAGCAACGGGAGCAGCGAAAGCAACGGGAGCGTCTTGGCTGATGGCCACGCCGAAGAAGCGCCGAACGAGCAGGTGAGTAGCGAGAAGAACGACGGACACAG :  624
H32_Brazil-I_refsequ : GGAAGCAACGGGAGCAGCGAAAGCAACGGAAGCGTCTTGGCTGATGGCCACGCCGAAGAAGCGCCGAACGAGCAGGTGAGTAGCGAGAAGAACGACGGACACAG :  624
                       GGaagcaacgggagcagcgaAAGCAACGGaAGCGTCTTGGCTGATGGCCACGCCGAAGAAGCGCCGAACGAGCAGGTGAGTAGCGAGAAGAACGACGGACACAG

                            *       640         *       660         *       680         *       700         *       720
H1_VCG               : TGAAAGCAGCACCACGGGCTCACACTCAGATAAGGAGGAGCTGGACTTGTATACGTCCTCCCCCTCCGATAAAGAAGAAAACAGCGCCCCTGCCATGGGTGGTA :  728
H2                   : TGAAAGCAGCACCACGGGCTCACACTCAGATAAGGAGGAGCTGGACTTGTATACGTCCTCCCCCTCCGATAAAGAAGAAAACAGCGCCCCTGCCATGGGTGGTA :  728
H3                   : TGAAAGCAGCACCACGGGCTCACACTCAGATAAGGAGGAGCTGGACTTGTATACGTCCTCCCCCTCCGATAAAGAAGAAAACAGCGCCCCTGCCATGGGTGGTA :  728
H4                   : TGAAAGCAGCACCACGGGCTCACACTCAGATAAGGAGGAGCTGGACTTGTATACGTCCTCCCCCTCCGATAAAGAAGAAAACAGCGCCCCTGCCATGGGTGGTA :  728
```

```
H5                    : TGAAAGCAGCACCACGGGCTCACACTCAGATAAGGAGGAGCTGGACTTGTATACGTCCTCCCCCTCCGATAAAGAAGAAAACAGCGCCCCTGCCATGGGTGGTA : 728
H6                    : TGAAAGCAGCACCACGGGCTCACACTCAGATAAGGAGGAGCTGGACTTGTATACGTCCTCCCCCTCCGATAAAGAAGAAAACAGCGCCCCTGCCATGGGTGGTA : 728
H7                    : TGAAAGCAGCACCACGGGCTCACACTCAGATAAGGAGGAGCTGGACTTGTATACGTCCTCCCCCTCCGATAAAGAAGAAAACAGCGCCCCTGCCATGGGTGGTA : 728
H8                    : TGAAAGCAGCACCACGGGCTCACACTCAGATAAGGAGGAGCTGGACTTGTATACGTCCTCCCCCTCCGATAAAGAAGAAAACAGCGCCCCTGCCATGGGTGGTA : 728
H9                    : TGAAAGCAGCACCACGGGCTCACACTCAGATAAGGAGGAGCTGGACTTGTATACGTCCTCCCCCTCCGATAAAGAAGAAAACAGCGCCCCTGCCATGGGTGGTA : 711
H10                   : TGAAAGCAGCACCACGGGCTCACACTCAGATAAGGAGGAGCTGGACTTGTATACGTCCTCCCCCTCCGATAAAGAAGAAAACAGCGCCCCTGCCATGGGTGGTA : 728
H11                   : TGAAAGCAGCACCACGGGCTCACACTCAGATAAGGAGGAGCTGGACTTGTATACGTCCTCCCCCTCCGATAAAGAAGAAAACAGCGCCCCTGCCATGGGTGGTA : 728
H12                   : TGAAAGCAGCACCACGGGCTCACACTCAGATAAGGAGGAGCTGGACTTGTATACGTCCTCCCCCTCCGATAAAGAAGAAAACAGCGCCCCTGCCATGGGTGGTA : 728
H13                   : TGAAAGCAGCACCACGGGCTCACACTCAGATAAGGAGGAGCTGGACTTGTATACGTCCTCCCCCTCCGATAAAGAAGAAAACAGCGCCCCTGCCATGGGTGGTA : 728
H14                   : TGAAAGCAGCACCACGGGCTCACACTCAGATAAGGAGGAGCTGGACTTGTATACGTCCTCCCCCTCCGATAAAGAAGAAAACAGCGCCCCTGCCATGGGTGGTA : 728
H15                   : TGAAAGCAGCACCACGGGCTCACACTCAGATAAGGAGGAGCTGGACTTGTATACGTCCTCCCCCTCCGATAAAGAAGAAAACAGCGCCCCTGCCATGGGTGGTA : 728
H16                   : TGAAAGCAGCACCACGGGCTCACACTCAGATAAGGAGGAGCTGGACTTGTATACGTCCTCCCCCTCCGATAAAGAAGAAAACAGCGCCCCTGCCATGGGTGGTA : 728
H17                   : TGAAAGCAGCACCACGGGCTCACACTCAGATAAGGAGGAGCTGGACTTGTATACGTCCTCCCCCTCCGATAAAGAAGAAAACAGCGCCCCTGCCATGGGTGGTA : 728
H18                   : TGAAAGCAGCACCACGGGCTCACACTCAGATAAGGAGGAGCTGGACTTGTATACGTCCTCCCCCTCCGATAAAGAAGAAAACAGCGCCCCTGCCATGGGTGGTA : 728
H19                   : TGAAAGCAGCACCACGGGCTCACACTCAGATAAGGAGGAGCTGGACTTGTATACGTCCTCCCCCTCCGATAAAGAAGAAAACAGCGCCCCTGCCATGGGTGGTA : 728
H20                   : TGAAAGCAGCACCACGGGCTCACACTCAGATAAGGAGGAGCTGGACTTGTATACGTCCTCCCCCTCCGATAAAGAAGAAAACAGCGCCCCTGCCATGGGTGGTA : 728
H21                   : TGAAAGCAGCACCACGGGCTCACACTCAGATAAGGAGGAGCTGGACTTGTATACGTCCTCCCCCTCCGATAAAGAAGAAAACAGCGCCCCTGCCATGGGTGGTA : 728
H22                   : TGAAAGCAGCACCACGGGCTCACACTCAGATAAGGAGGAGCTGGACTTGTATACGTCCTCCCCCTCCGATAAAGAAGAAAACAGCGCCCCTGCCATGGGTGGTA : 728
H23                   : TGAAAGCAGCACCACGGGCTCACACTCAGATAAGGAGGAGCTGGACTTGTATACGTCCTCCCCCTCCGATAAAGAAGAAAACAGCGCCCCTGCCATGGGTGGTA : 728
H24                   : TGAAAGCAGCACCACGGGCTCACACTCAGATAAGGAGGAGCTGGACTTGTATACGTCCTCCCCCTCCGATAAAGAAGAAAACAGCGCCCCTGCCATGGGTGGTA : 728
H25                   : TGAAAGCAGCACCACGGGCTCACACTCAGATAAGGAAGAGCTGGACTTGTATACGTCCTCCCCCTCCGATAAAGAAGAAAACAGCGCCCCTGCCATGGGTGGTA : 728
H26_Mauritania-I_ref  : TGAAAGCAGCACCACGGGCTCACACTCAGATAAGGAGGAGCTGGACTTGTATACGTCCTCCCCCTCCGATAAAGAAGAAAACAGCGCCCCTGCCATGGGTGGTA : 710
H27_India-VII_refseq  : TGAAAGCAGCACCACGGGCTCACACTCAGATAAGGAGGAGCTGGACTTGTATACGTCCTCCCCCTCCGATAAAGAAGAAAACAGCGCCCCTGCCATGGGTGGTA : 728
H28_Ctg_refsequence   : TGAAAGCAGCACCACGGGCTCACACTCAGATAAGGAGGAGCTGGACTTGTATACGTCCTCCCCCTCCGATAAAGAAGAAAACAGCGCCCCTGCCATGGGTGGTA : 728
H29_Salvador-I_refse  : TGAAAGCAGCACCACGGGCTCACACTCAGATAAGGAGGAGCTGGACTTGTATACGTCCTCCCCCTCCGATAAAGAAGAAAACAGCGCCCCTGCCATGGGTGGTA : 728
H30_NKorea_refsequen  : TGAAAGCAGCACCACGGGCTCACACTCAGATAAGGAGGAGCTGGACTTGTATACGTCCTCCCCCTCCGATAAAGAAGAAAACAGCGCCCCTGCCATGGGTGGTA : 728
H31_P.vivax01_refseq  : TGAAAGCAGCACCACGGGCTCACACTCAGATAAGGAGGAGCTGGACTTGTATACGTCCTCCCCCTCCGATAAAGAAGAAAACAGCGCCCCTGCCATGGGTGGTA : 728
H32_Brazil-I_refsequ  : TGAAAGCAGCACCACGGGCTCACACTCAGATAAGGAGGAGCTGGACTTGTATACGTCCTCCCCCTCCGATAAAGAAGAAAACAGCGCCCCTGCCATGGGTGGTA : 728
                        TGAAAGCAGCACCACGGGCTCACACTCAGATAAGGAgGAGCTGGACTTGTATACGTCCTCCCCCTCCGATAAAGAAGAAAACAGCGCCCCTGCCATGGGTGGTA

                         *       740         *       760         *       780         *       800         *       820         *
H1_VCG                : TGCACAACACGGTAGGCCTGAGGGGGGGAAGCGCCTCTGAATCGGCGGCCCCAAGTGAAGAGAAAGCATCCGAAGAGAATGGCCCCCATGAAGTGGCAGATAAA : 832
H2                    : TGCACAACACGGTAGGCCTGAGGGGGGGAAGCGCCTCTGAATCGGCGGCCCCAAGTGAAGAGAAAGCATCCGAAGAGAATGTCCCCCATGAAGTGGCAGATAAA : 832
H3                    : TGCACAACACGGTAGGCCTGAGGGGGGGAAGCGCCTCTGAATCGGCGGCCCCAAGTGAAGAGAAAGCATCCGAAGAGAATGTCCCCCATGAAGTGGCAGATAAA : 832
H4                    : TGCACAACACGGTAGGCCTGAGGGGGGGAAGCGCCTCTGAATCGGCGGCCCCAAGTGAAGAGAAAGCATCCGAAGAGAATGTCCCCCATGAAGTGGCAGATAAA : 832
H5                    : TGCACAACACGGTAGGCCTGAGGGGGGGAAGCGCCTCTGAATCGGCGGCCCCAAGTGAAGAGAAAGCATCCGAAGAGAATGTCCCCCATGAAGTGGCAGATAAA : 832
H6                    : TGCACAACACGGTAGGCCTGAGGGGGGGAAGCGCCTCTGAATCGGCGGCCCCAAGTGAAGAGAAAGCATCCGAAGAGAATGGCCCCCATGAAGTGGCAGATAAA : 832
H7                    : TGCACAACACGGTAGGCCTGAGGGGGGGAAGCGCCTCTGAATCGGCGGCCCCAAGTGAAGAGAAAGCATCCGAAGAGAATGTCCCCCATGAAGTGGCAGATAAA : 832
H8                    : TGCACAACACGGTAGGCCTGAGGGGGGGAAGCGCCTCTGAATCGGCGGCCCCAAGTGAAGAGAAAGCATCCGAAGAGAATGTCCCCCATGAAGTGGCAGATAAA : 832
H9                    : TGCACAACACGGTAGGCCTGAGGGGGGGAAGCGCCTCTGAATCGGCGGCCCCAAGTGAAGAGAAAGCATCCGAAGAGAATGTCCCCCATGAAGTGGCAGATAAA : 815
H10                   : TGCACAACACGGTAGGCCTGAGGGGGGGAAGCGCCTCTGAATCGGCGGCCCCAAGTGAAGAGAAAGCATCCGAAGAGAATGTCCCCCATGAAGTGGCAGATAAA : 832
H11                   : TGCACAACACGGTAGGCCTGAGGGGGGGAAGCGCCTCTGAATCGGCGGCCCCAAGTGAAGAGAAAGCATCCGAAGAGAATGTCCCCCATGAAGTGGCAGATAAA : 832
H12                   : TGCACAACACGGTAGGCCTGAGGGGGGGAAGCGCCTCTGAATCGGCGGCCCCAAGTGAAGAGAAAGCATCCGAAGAGAATGTCCCCCATGAAGTGGCAGATAAA : 832
H13                   : TGCACAACACGGTAGGCCTGAGGGGGGGAAGCGCCTCTGAATCGGCGGCCCCAAGTGAAGAGAAAGCATCCGAAGAGAATGTCCCCCATGAAGTGGCAGATAAA : 832
H14                   : TGCACAACACGGTAGGCCTGAGGGGGGGAAGCGCCTCTGAATCGGCGGCCCCAAGTGAAGAGAAAGCATCCGAAGAGAATGTCCCCCATGAAGTGGCAGATAAA : 832
```

```
H15                  : TGCACAACACGGTAGGCCTGAGGGGGGGAAGCGCCTCTGAATCGGCGGCCCCAAGTGAAGAGAAAGCATCCGAAGAGAATGTCCCCCATGAAGTGGCAGATAAA :  832
H16                  : TGCACAACACGGTAGGCCTGAGGGGGGGAAGCGCCTCTGAATCGGCGGCCCCAAGTGAAGAGAAAGCATCCGAAGAGAATGGCCCCCATGAAGTGGCAGATAAA :  832
H17                  : TGCACAACACGGTAGGCCTGAGGGGGGGAAGCGCCTCTGAATCGGCGGCCCCAAGTGAAGAGAAAGCATCCGAAGAGAATGTCCCCCATGAAGTGGCAGATAAA :  832
H18                  : TGCACAACACGGTAGGCCTGAGGGGGGGAAGCGCCTCTGAATCGGCGGCCCCAAGTGAAGAGAAAGCATCCGAAGAGAATGTCCCCCATGAAGTGGCAGATAAA :  832
H19                  : TGCACAACACGGTAGGCCTGAGGGGGGGAAGCGCCTCTGAATCGGCGGCCCCAAGTGAAGAGAAAGCATCCGAAGAGAATGTCCCCCATGAAGTGGCAGATAAA :  832
H20                  : TGCACAACACGGTAGGCCTGAGGGGGGGAAGCGCCTCTGAATCGGCGGCCCCAAGTGAAGAGAAAGCATCCGAAGAGAATGTCCCCCATGAAGTGGCAGATAAA :  832
H21                  : TGCACAACACGGTAGGCCTGAGGGGGGGAAGCGCCTCTGAATCGGCGGCCCCAAGTGAAGAGAAAGCATCCGAAGAGAATGTCCCCCATGAAGTGGCAGATAAA :  832
H22                  : TGCACAACACGGTAGGCCTGAGGGGGGGAAGCGCCTCTGAATCGGCGGCCCCAAGTGAAGAGAAAGCATCCGAAGAGAATGTCCCCCATGAAGTGGCAGATAAA :  832
H23                  : TGCACAACACGGTAGGCCTGAGGGGGGGAAGCGCCTCTGAATCGGCGGCCCCAAGTGAAGAGAAAGCATCCGAAGAGAATGTCCCCCATGAAGTGGCAGATAAA :  832
H24                  : TGCACAACACGGTAGGCCTGAGGGGGGGAAGCGCCTCTGAATCGGCGGCCCCAAGTGAAGAGAAAGCATCCGAAGAGAATGTCCCCCATGAAGTGGCAGATAAA :  832
H25                  : TGCACAACACGGTAGGCCTGAGGGGGGGAAGCGCCTCTGAATCGGCGGCCCCAAGTGAAGAGAAAGCATCCGAAGAGAATGTCCCCCATGAAGTGGCAGATAAA :  832
H26_Mauritania-I_ref : TGCACAACACGGTAGGCCTGAGGGGGGGAAGCGCCTCTGAATCGGCGGCCCCAAGTGAAGAGAAAGCATCCGAAGAGAATGTCCCCCATGAAGTGGCAGATAAA :  814
H27_India-VII_refseq : TGCACAACACGGTAGGCCTGAGGGGGGGAAGCGCCTCTGAATCGGCGGCCCCAAGTGAAGAGAAAGCATCCGAAGAGAATGTCCCCCATGAAGTGGCAGATAAA :  832
H28_Ctg_refsequence  : TGCACAACACGGTAGGCCTGAGGGGGGGAAGCGCCTCTGAATCGGCGGCCCCAAGTGAAGAGAAAGCATCCGAAGAGAATGTCCCCCATGAAGTGGCAGATAAA :  832
H29_Salvador-I_refse : TGCACAACACGGTAGGCCTGAGGGGGGGAAGCGCCTCTGAATCGGCGGCCCCAAGTGAAGAGAAAGCATCCGAAGAGAATGTCCCCCATGAAGTGGCAGATAAA :  832
H30_NKorea_refsequen : TGCACAACACGGTAGGCCTGAGGGGGGGAAGCGCCTCTGAATCGGCGGCCCCAAGTGAAGAGAAAGCATCCGAAGAGAATGTCCCCCATGAAGTGGCAGATAAA :  832
H31_P.vivax01_refseq : TGCACAACACGGTAGGCCTGAGGGGGGGAAGCGCCTCTGAATCGGCGGCCCCAAGTGAAGAGAAAGCATCCGAAGAGAATGTCCCCCATGAAGTGGCAGATAAA :  832
H32_Brazil-I_refsequ : TGCACAACACGGTAGGCCTGAGGGGGGGAAGCGCCTCTGAATCGGCGGCCCCAAGTGAAGAGAAAGCATCCGAAGAGAATGTCCCCCATGAAGTGGCAGATAAA :  832
                       TGCACAACACGGTAGGCCTGAGGGGGGGAAGCGCCTCTGAATCGGCGGCCCCAAGTGAAGAGAAAGCATCCGAAGAGAATGtCCCCCATGAAGTGGCAGATAAA

                            840         *       860         *       880         *       900         *       920         *
H1_VCG               : AATGGGGAGGACACCTCACAGGAGGGTTCCCACGCGGGAGCGCTGTCACCTGCGCATGTGGGTGGCGA---------GAGTGGCGAGAGTGGCGGCGAGCACGG :  927
H2                   : AATGGGGAGGACACCTCACAGGAGGGTTCCCACGCGGGAGCGCTGTCACCTGCGCATGTGGGTGGCGA---------GAGTGGCGAGAGTGGCGGCGAGCACGG :  927
H3                   : AATGGGGAGGACACCTCACAGGAGGGTTCCCACGCGGGAGCGCTGTCACCTGCGCATGTGGGTGGCGA---------GAGTGGCGAGAGTGGCGGCGAGCACGG :  927
H4                   : AATGGGGAGGACACCTCACAGGAGGGTTCCCACGCGGGAGCGCTGTCACCTGCGCATGTGGGTGGCGA---------GAGTGGCGAGAGTGGCGGCGAGCACGG :  927
H5                   : AATGGGGAGGACACCTCACAGGAGGGTTCCCACGCGGGAGCGCTGTCACCTGCGCATGTGGGTGGCGA---------GAGTGGCGAGAGTGGCGGCGAGCACGG :  927
H6                   : AATGGGGAGGACACCTCACAGGAGGGTTCCCACGCGGGAGCGCTGTCACCTGCGCATGTGGGTGGCGA-------------------GAGTGGCGGCGAGCACGG :  918
H7                   : AATGGGGAGGACACCTCACAGGAGGGTTCCCACGCGGGAGCGCTGTCACCTGCGCATGTGGGTGGCGAGAGTGGCGAGAGTGGCGAGAGTGGCGGCGAGCACGG :  936
H8                   : AATGGGGAGGACACCTCACAGGAGGGTTCCCACGCGGGAGCGCTGTCACCTGCGCATGTGGGTGGCGAGAGTGGCGAGAGTGGCGAGAGTGGCGGCGAGCACGG :  936
H9                   : AATGGGGAGGACACCTCACAGGAGGGTTCCCACGCGGGAGCGCTGTCACCTGCGCATGTGGGTGGCGA---------GAGTGGCGAGAGTGGCGGCGAGCACGG :  910
H10                  : AATGGGGAGGACACCTCACAGGAGGGTTCCCACGCGGGAGCGCTGTCACCTGCGCATGTGGGTGGCGAGAGTGGCGAGAGTGGCGAGAGTGGCGGCGAGCACGG :  936
H11                  : AATGGGGAGGACACCTCACAGGAGGGTTCCCACGCGGGAGCGCTGTCACCTGCGCATGTGGGTGGCGA---------GAGTGGCGAGAGTGGCGGCGAGCACGG :  927
H12                  : AATGGGGAGGACACCTCACAGGAGGGTTCCCACGCGGGAGCGCTGTCACCTGCGCATGTGGGTGGCGA---------GAGTGGCGAGAGTGGCGGCGAGCACGG :  927
H13                  : AATGGGGAGGACACCTCACAGGAGGGTTCCCACGCGGGAGCGCTGTCACCTGCGCATGTGGGTGGCGA---------GAGTGGCGAGAGTGGCGGCGAGCACGG :  927
H14                  : AATGGGGAGGACACCTCACAGGAGGGTTCCCACGCGGGAGCGCTGTCACCTGCGCATGTGGGTGGCGA---------GAGTGGCGAGAGTGGCGGCGAGCACGG :  927
H15                  : AATGGGGAGGACACCTCACAGGAGGGTTCCCACGCGGGAGCGCTGTCACCTGCGCATGTGGGTGGCGA---------GAGTGGCGAGAGTGGCGGCGAGCACGG :  927
H16                  : AATGGGGAGGACACCTCACAGGAGGGTTCCCACGCGGGAGCGCTGTCACCTGCGCATGTGGGTGGCGA---------GAGTGGCGAGAGTGGCGGCGAGCACGG :  927
H17                  : AATGGGGAGGACACCTCACAGGAGGGTTCCCACGCGGGAGCGCTGTCACCTGCGCATGTGGGTGGCGA---------GAGTGGCGAGAGTGGCGGCGAGCACGG :  927
H18                  : AATGGGGAGGACACCTCACAGGAGGGTTCCCACGCGGGAGCGCTGTCACCTGCGCATGTGGGTGGCGA---------GAGTGGCGAGAGTGGCGGCGAGCACGG :  927
H19                  : AATGGGGAGGACACCTCACAGGAGGGTTCCCACGCGGGAGCGCTGTCACCTGCGCATGTGGGTGGCGA---------GAGTGGCGAGAGTGGCGGCGAGCACGG :  927
H20                  : AATGGGGAGGACACCTCACAGGAGGGTTCCCACGCGGGAGCGCTGTCACCTGCGCATGTGGGTGGCGA---------GAGTGGCGAGAGTGGCGGCGAGCACGG :  927
H21                  : AATGGGGAGGACACCTCACAGGAGGGTTCCCACGCGGGAGCGCTGTCACCTGCGCATGTGGGTGGCGA---------GAGTGGCGAGAGTGGCGGCGAGCACGG :  927
H22                  : AATGGGGAGGACACCTCACAGGAGGGTTCCCACGCGGGAGCGCTGTCACCTGCGCATGTGGGTGGCGA---------GAGTGGCGAGAGTGGCGGCGAGCACGG :  927
H23                  : AATGGGGAGGACACCTCACAGGAGGGTTCCCACGCGGGAGCGCTGTCACCTGCGCATGTGGGTGGCGA---------GAGTGGCGAGAGTGGCGGCGAGCACGG :  927
H24                  : AATGGGGAGGACACCTCACAGGAGGGTTCCCACGCGGGAGCGCTGTCACCTGCGCATGTGGGTGGCGAGAGTGGCGAGAGTGGCGAGAGTGGCGGCGAGCACGG :  936
```

```
H25                  : AATGGGGAGGACACCTCACAGGAGGGTTCCCACGCGGGAGCGCTGTCACCTGCGCATGTGGGTGGCGA---------GAGTGGCGAGAGTGGCGGCGAGCACGG :  927
H26_Mauritania-I_ref : AATGGGGAGGACACCTCACAGGAGGGTTCCCACGCGGGAGCGCTGTCACCTGCGCATGTGGGTGGCGA-------------------GAGTGGCGGCGAGCACGG :  900
H27_India-VII_refseq : AATGGGGAGGACACCTCACAGGAGGGTTCCCACGCGGGAGCGCTGTCACCTGCGCATGTGGGTGGCGA-------------------GAGTGGCGGCGAGCACGG :  918
H28_Ctg_refsequence  : AATGGGGAGGACACCTCACAGGAGGGTTCCCACGCGGGAGCGCTGTCACCTGCGCATGTGGGTGGCGA---------GAGTGGCGAGAGTGGCGGCGAGCACGG :  927
H29_Salvador-I_refse : AATGGGGAGGACACCTCACAGGAGGGTTCCCACGCGGGAGCGCTGTCACCTGCGCATGTGGGTGGCGA---------GAGTGGCGAGAGTGGCGGCGAGCACGG :  927
H30_NKorea_refsequen : AATGGGGAGGACACCTCACAGGAGGGTTCCCACGCGGGAGCGCTGTCACCTGCGCATGTGGGTGGCGA-------------------GAGTGGCGGCGAGCACGG :  918
H31_P.vivax01_refseq : AATGGGGAGGACACCTCACAGGAGGGTTCCCACGCGGGAGCGCTGTCACCTGCGCATGTGGGTGGCGA-------------------GAGTGGCGGCGAGCACGG :  918
H32_Brazil-I_refsequ : AATGGGGAGGACACCTCACAGGAGGGTTCCCACGCGGGAGCGCTGTCACCTGCGCATGTGGGTGGCGA---------GAGTGGCGAGAGTGGCGGCGAGCACGG :  927
                       AATGGGGAGGACACCTCACAGGAGGGTTCCCACGCGGGAGCGCTGTCACCTGCGCATGTGGGTGGCGA         gagtggcgaGAGTGGCGGCGAGCACGG

                        940         *       960         *       980         *      1000         *      1020         *      1040
H1_VCG               : AGAGCACGGTGAACACGCTGAACATGGGGAGCACGGAGAGCACGGTGAACATGGCGAGCACGGTGAACA------------------------------------ :  996
H2                   : AGAGCACGGTGAACACGCTGAACATGGGGAGCACGGA------------------GAGCACGGTGAACA----------------CGCTGAACATGGGGAGCAC :  997
H3                   : AGAGCACGGTGAACACGCTGAACATGGGGAGCACGGA------------------GAGCACGGTGAACA----------------CGCTGAACATGGGGAGCAC :  997
H4                   : AGAGCACGGTGAACACGCTGAACATGGGGAGCACGGA------------------GAGCACGGTGAACA----------------CGCTGAACATGGGGAGCAC :  997
H5                   : AGAGCACGGTGAACACGCTGAACATGGGGAGCACGGA------------------GAGCACGGTGAACA----------------CGCTGAACATGGGGAGCAC :  997
H6                   : AGAGCACGGTGAACACGCTGAACATGGGGAGCACGGA------------------GAGCACGGTGAACATGGCGAGCACGGTGAA------------------- :  985
H7                   : AGAGCACGGTGAACACGCTGAACATGGGGAGCACGGAGAGCACGGTGAACATGGCGAGCACGGTGAACA------------------------------------ : 1005
H8                   : AGAGCACGGTGAACACGCTGAACATGGGGAGCACGGA------------------GAGCACGGTGAACA----------------CGCTGAACATGGGGAGCAC : 1006
H9                   : AGAGCACGGTGAACACGCTGAACATGGGGAGCACGGA------------------GAGCACGGTGAACA----------------CGCTGAACATGGGGAGCAC :  980
H10                  : AGAGCACGGTGAACACGCTGAACATGGGGAGCACGGA------------------GAGCACGGTGAACA----------------CGCTGAACATGGGGAGCAC : 1006
H11                  : AGAGCACGGTGAACACGCTGAACATGGGGAGCACGGA------------------GAGCACGGTGAACA----------------CGCTGAACATGGGGAGCAC :  997
H12                  : AGAGCACGGTGAACACGCTGAACATGGGGAGCACGGA------------------GAGCACGGTGAACA----------------CGCTGAACATGGGGAGCAC :  997
H13                  : AGAGCACGGTGAACACGCTGAACATGGGGAGCACGGA------------------GAGCACGGTGAACA----------------CGCTGAACATGGGGAGCAC :  997
H14                  : AGAGCACGGTGAACACGCTGAACATGGGGAGCACGGA------------------GAGCACGGTGAACA----------------CGCTGAACATGGGGAGCAC :  997
H15                  : AGAGCACGGTGAACACGCTGAACATGGGGAGCACGGA------------------GAGCACGGTGAACA----------------CGCTGAACATGGGGAGCAC :  997
H16                  : AGAGCACGGTGAACACGCTGAACATGGGGAGCACGGA------------------GAGCACGGTGAACATGGCGAGCACGGTGAA------------------- :  994
H17                  : AGAGCACGGTGAACACGCTGAACATGGGGAGCACGGA------------------GAGCACGGTGAACA----------------CGCTGAACATGGGGAGCAC :  997
H18                  : AGAGCACGGTGAACACGCTGAACATGGGGAGCACGGA------------------GAGCACGGTGAACATGGCGAGCACGGTGAA------------------- :  994
H19                  : AGAGCACGGTGAACACGCTGAACATGGGGAGCACGGA------------------GAGCACGGTGAACA----------------CGCTGAACATGGGGAGCAC :  997
H20                  : AGAGCACGGTGAACACGCTGAACATGGGGAGCACGGA------------------GAGCACGGTGAACA----------------CGCTGAACATGGGGAGCAC :  997
H21                  : AGAGCACGGTGAACACGCTGAACATGGGGAGCACGGA------------------GAGCACGGTGAACA----------------CGCTGAACATGGGGAGCAC :  997
H22                  : AGAGCACGGTGAACACGCTGAACATGGGGAGCACGGA------------------GAGCACGGTGAACA----------------CGCTGAACATGGGGAGCAC :  997
H23                  : AGAGCACGGTGAACACGCTGAACATGGGGAGCACGGA------------------GAGCACGGTGAACA----------------CGCTGAACATGGGGAGCAC :  997
H24                  : AGAGCACGGTGAACACGCTGAACATGGGGAGCACGGA------------------GAGCACGGTGAACA----------------CGCTGAACATGGGGAGCAC : 1006
H25                  : AGAGCACGGTGAACACGCTGAACATGGGGAGCACGGA------------------GAGCACGGTGAACA----------------CGCTGAACATGGGGAGCAC :  997
H26_Mauritania-I_ref : AGAGCACGGTGAACACGCTGAACATGGGGAGCACGGA------------------GAGCACGGTGAACA----------------CGCTGAACATGGGGAGCAC :  970
H27_India-VII_refseq : AGAGCACGGTGAACACGCTGAACATGGGGAGCACGGAGAGCACGGTGAACATGGCGAGCACGGTGAACATGGCGAGCACGGTGAA------------------- : 1003
H28_Ctg_refsequence  : AGAGCACGGTGAACACGCTGAACATGGGGAGCACGGAGAGCACGG------------------------------------------------------------ :  972
H29_Salvador-I_refse : AGAGCACGGTGAACACGCTGAACATGGGGAGCACGGA------------------------------------------------------------------ :  964
H30_NKorea_refsequen : AGAGCACGGTGAACACGCTGAACATGGGGAGCACGGAGAGCACGGT---------------------------------------------------------- :  964
H31_P.vivax01_refseq : ---------TGAACACGCTGAACATGGGGAGCACGGAGAGCACGGTGAACA--------------------------------CGCTGAACATGGGGAGCAC :  979
H32_Brazil-I_refsequ : AGAGCACGGTGAACACGCTGAACATGGGGAGCACGGAGAGCACGGTGAACATGGCGAGCACGGTGAACA------------------------------------ :  996
                       agagcacggTGAACACGCTGAACATGGGGAGCACGGA                  gagcacggtgaaca
```

```
                                  *      1060         *      1080         *      1100         *      1120         *      1140
H1_VCG                : ------------------------------------------------------------------------------------------------------ :    -
H2                    : GGAGAGCACGGTGAA----------------------------------------------------------------------------------------- : 1012
H3                    : GGAGAGCACGGTGAA----------------------------------------------------------------------------------------- : 1012
H4                    : GGAGAGCACGGTGAA----------------------------------------------------------------------------------------- : 1012
H5                    : GGAGAGCACGGTGAA----------------------------------------------------------------------------------------- : 1012
H6                    : ------------------------------------------------------------------------------------------------------ :    -
H7                    : ------------------------------------------------------------------------------------------------------ :    -
H8                    : GGAGAGCACGGTGAACAT-------------------------------------------------------------------------------------- : 1024
H9                    : GGAGAGCACGGTGAA----------------------------------------------------------------------------------------- :  995
H10                   : GGAGAGCACGGTGAA----------------------------------------------------------------------------------------- : 1021
H11                   : GGAGAGCACGGTGAA----------------------------------------------------------------------------------------- : 1012
H12                   : GGAGAGCACGGTGAACAT-------------------------------------------------------------------------------------- : 1015
H13                   : GGAGAGCACGGTGAA----------------------------------------------------------------------------------------- : 1012
H14                   : GGAGAGCACGGTGAA----------------------------------------------------------------------------------------- : 1012
H15                   : GGAGAGCACGGTGAA----------------------------------------------------------------------------------------- : 1012
H16                   : ------------------------------------------------------------------------------------------------------ :    -
H17                   : GGAGAGCACGGTGAA----------------------------------------------------------------------------------CACGCTGA : 1020
H18                   : ------------------------------------------------------------------------------------------------------ :    -
H19                   : GGAGAGCACGGTGAA----------------------------------------------------------------------------------------- : 1012
H20                   : GGAGAGCACGGTGAA----------------------------------------------------------------------------------------- : 1012
H21                   : GGAGAGCACGGTGAA----------------------------------------------------------------------------------------- : 1012
H22                   : GGAGAGCACGGTGAA----------------------------------------------------------------------------------------- : 1012
H23                   : GGAGAGCACGGTGAA----------------------------------------------------------------------------------------- : 1012
H24                   : GGAGAGCACGGTGAA----------------------------------------------------------------------------------------- : 1021
H25                   : GGAGAGCACGGTGAA----------------------------------------------------------------------------------------- : 1012
H26_Mauritania-I_ref  : GGAGAGCACGGTGAACATGGTGAGCATGGTGAGCACGGTGAGCATGGTGAGCATGGTGAGCACGGTGAACATGGTGAGCATGGGGAGCACGGTGAG-------- : 1066
H27_India-VII_refseq  : ------------------------------------------------------------------------------------------------------ :    -
H28_Ctg_refsequence   : ------------------------------------------------------------------------------------------------------ :    -
H29_Salvador-I_refse  : ------------------------------------------------------------------------------------------------------ :    -
H30_NKorea_refsequen  : ------------------------------------------------------------------------------------------------------ :    -
H31_P.vivax01_refseq  : GGAGAGCACGGTGAA----------------------------------------------------------------------------------------- :  994
H32_Brazil-I_refsequ  : ------------------------------------------------------------------------------------------------------ :    -

                                      *      1160         *      1180         *      1200         *      1220         *      1240
H1_VCG                : ------------------------------------------------------------------------------------------------------ :    -
H2                    : ------------------------------------------------------------------------------------------------------ :    -
H3                    : ----------------------------------------------------------------------------------------------------CACG : 1016
H4                    : ------------------------------------------------------------------------------------------------------ :    -
H5                    : ------------------------------------------------------------------------------------------------------ :    -
H6                    : ------------------------------------------------------------------------------------------------------ :    -
H7                    : ------------------------------------------------------------------------------------------------------ :    -
H8                    : ------------------------------------------------------------------------------------------------------ :    -
H9                    : ------------------------------------------------------------------CACGCTGAACATGGGGAGCACGGAGAGCACGGTGAACACG : 1035
```

```
H10                  : ------------------------------------------------------------------------------------------------------------ :    -
H11                  : ------------------------------------------------------------------------------------------------------------ :    -
H12                  : ------------------------------------------------------------------------------------------------------------ :    -
H13                  : ------------------------------------------------------------------------------------------------------------CACG : 1016
H14                  : ------------------------------------------------------------------------------------------------------------ :    -
H15                  : ------------------------------------------------------------------------------------------------------------ :    -
H16                  : ------------------------------------------------------------------------------------------------------------ :    -
H17                  : ACATGGGGAGCACGGAGAGCACGGTGAACACGCTGAACATGGGGAGCACGGAGAGCACGGTGAACACGCTGAACATGGGGAGCACGGAGAGCACGGTGAACACG : 1124
H18                  : ------------------------------------------------------------------------------------------------------------ :    -
H19                  : ------------------------------------------------------------------------------------------------------------ :    -
H20                  : ------------------------------------------------------------------------------------------------------------ :    -
H21                  : ------------------------------------------------------------------------------------------------------------ :    -
H22                  : ------------------------------------------------------------------------------------------------------------ :    -
H23                  : ------------------------------------------------------------------------------------------------------------ :    -
H24                  : ------------------------------------------------------------------------------------------------------------ :    -
H25                  : ------------------------------------------------------------------------------------------------------------ :    -
H26_Mauritania-I_ref : ------------------------------------------------------------------------------------------------------------ :    -
H27_India-VII_refseq : ------------------------------------------------------------------------------------------------------------ :    -
H28_Ctg_refsequence  : ------------------------------------------------------------------------------------------AGAGCACGGTGAACACG :  989
H29_Salvador-I_refse : ------------------------------------------------------------------------------GAGCACGGAGAGCACGGTGAACACG :  989
H30_NKorea_refsequen : ------------------------------------------------------------------------------------------------------------ :    -
H31_P.vivax01_refseq : ------------------------------------------------------------------------------------------------------------ :    -
H32_Brazil-I_refsequ : ------------------------------------------------------------------------------------------------------------ :    -

                       *      1260         *      1280         *      1300         *      1320         *      1340         *
H1_VCG               : ------------------------------------------------------------------------------------------TGCTGAACAGGAGGAG : 1012
H2                   : ------------------------------------------------------------------CATGGCGAGCACGGTGAACATGCTGAACAGGAGGAG : 1048
H3                   : CTGAACATGGGGAGCACGGAGAGCACGGTGAACACGCTGAACATGGGGAGCACGGAGAGCACGGTGAACATGGCGAGCACGGTGAACATGCTGAACAGGAGGAG : 1120
H4                   : ---------------------------------CACGCTGAACATGGGGAGCACGGAGAGCACGGTGAACATGGCGAGCACGGTGAACATGCTGAACAGGAGGAG : 1084
H5                   : ---------------------------------CACGCTGAACATGGGGAGCACGGAGAGCACGGTGAACATGGCGAGCACGGTGAACATGCTGAACAGGAGGAG : 1084
H6                   : ------------------------------------------------------------------------------------------CATGCTGAACAGGAGGAG : 1003
H7                   : ------------------------------------------------------------------------------------------TGCTGAACAGGAGGAG : 1021
H8                   : ------------------------------------------------------------------------GGCGAGCACGGTGAACATGCTGAACAGGAGGAG : 1057
H9                   : CTGAACATGGGGAGCACGGAGAGCACGGTGAACACGCTGAACATGGGGAGCACGGAGAGCACGGTGAACATGGCGAGCACGGTGAACATGCTGAACAGGAGGAG : 1139
H10                  : ------------------------------------------------------------------CATGGCGAGCACGGTGAACATGCTGAACAGGAGGAG : 1057
H11                  : ---------------------------------CACGCTGAACATGGGGAGCACGGAGA------------------GCACGGTGAACATGCTGAACAGGAGGAG : 1066
H12                  : ------------------------------------------------------------------------GGCGAGCACGGTGAACATGCTGAACAGGAGGAG : 1048
H13                  : CTGAACATGGGGAGCACGGAGAGCACGGTGAACACGCTGAACATGGGGAGCACGGAGAGCACGGTGAACATGGCGAGCACGGTGAACATGCTGAACAGGAGGAG : 1120
H14                  : ---------------------------------CACGCTGAACATGGGGAGCACGGAGA------------------GCACGGTGAACATGCTGAACAGGAGGAG : 1066
H15                  : ------------------------------------------------------------------CATGGCGAGCACGGTGAACATGCTGAACAGGAGGAG : 1048
H16                  : ------------------------------------------------------------------------------------------CATGCTGAACAGGAGGAG : 1012
H17                  : CTGAACATGGGGAGCACGGAGAGCACGGTGAACACGCTGAACATGGGGAGCACGGAGAGCACGGTGAACATGGCGAGCACGGTGAACATGCTGAACAGGAGGAG : 1228
H18                  : ------------------------------------------------------------------------------------------CATGCTGAACAGGAGGAG : 1012
H19                  : ---------------------------------CACGCTGAACATGGGGAGCACGGAGAGCACGGTGAA------------------CATGCTGAACAGGAGGAG : 1066
```

```
H20                  : ------------------------------------------------------------------CATGGCGAGCACGGTGAACATGCTGAACAGGAGGAG : 1048
H21                  : --------------------------------CACGCTGAACATGGGGAGCACGGAGAGCACGGTGAACATGGCGAGCACGGTGAACATGCTGAACAGGAGGAG : 1084
H22                  : ------------------------------------------------------------------CATGGCGAGCACGGTGAACATGCTGAACAGGAGGAG : 1048
H23                  : --------------------------------CACGCTGAACATGGGGAGCACGGAGAGCACGGTGAACATGGCGAGCACGGTGAACATGCTGAACAGGAGGAG : 1084
H24                  : --------------------------------CACGCTGAACATGGGGAGCACGGAGAGCACGGTGAACATGGCGAGCACGGTGAACATGCTGAACAGGAGGAG : 1093
H25                  : ------------------------------------------------------------------CATGGCGAGCACGGTGAACATGCTGAACAGGAGGAG : 1048
H26_Mauritania-I_ref : --------------------------------CACGCTGAACATGGGGAGCACGGAGAGCACGGTGAACATGGCGAGCACGGTGAACATGCTGAACAGGAGGAG : 1138
H27_India-VII_refseq : ------------------------------------------------------------------------------------CATGCTGAACAGGAGGAG : 1021
H28_Ctg_refsequence  : CTGAACATGGGGAGCACGGAGAGCACGGTGAACACGCTGAACATGGGGAGCACGGAGAGCACGGTGAACATGGCGAGCACGGTGAACATGCTGAACAGGAGGAG : 1093
H29_Salvador-I_refse : CTGAACATGGGGAGCACGGAGAGCACGGTGAACACGCTGAACATGGGGAGCACGGAGAGCACGGTGAACATGGCGAGCACGGTGAACATGCTGAACAGGAGGAG : 1093
H30_NKorea_refsequen : ------------------------------------------------------------------GAACATGGCGAGCACGGTGAACATGCTGAACAGGAGGAG : 1003
H31_P.vivax01_refseq : --------------CATGGCGAGCACGGTGAACACGCTGAACATGGGGAGCACGGAGAGCACGGTGAACATGGCGAGCACGGTGAACATGCTGAACAGGAGGAG : 1084
H32_Brazil-I_refsequ : ------------------------------------------------------------------------------------------TGCTGAACAGGAGGAG : 1012
                                                                                                                caTGCTGAACAGGAGGAG

                                 1360         *      1380         *      1400         *      1420         *      1440         *
H1_VCG               : CACCAAATCTATGGGACGATTCACCCCGAGGCGGAAGCCTCCATTTTGCCCGCGCTGAAGGAAAAGGGCATCCACGTGTCCAGCCCAGAAAGGGTAATCCTCGA : 1116
H2                   : CACCAAATCTATGGGACGATTCACCCCGAGGCGGAAGCCTCCATTTTGCCCGCGCTGAAGGAAAAGGGCATCCACGTGTCCAGCCCAGAAAGGGTAATCCTCGA : 1152
H3                   : CACCAAATCTATGGGACGATTCACCCCGAGGCGGAAGCCTCCATTTTGCCCGCGCTGAAGGAAAAGGGCATCCACGTGTCCAGCCCAGAAAGGGTAATCCTCGA : 1224
H4                   : CACCAAATCTATGGGACGATTCACCCCGAGGCGGAAGCCTCCATTTTGCCCGCGCTGAAGGAAAAGGGCATCCACGTGTCCAGCCCAGAAAGGGTAATCCTCGA : 1188
H5                   : CACCAAATCTATGGGACGATTCACCCCGAGGCGGAAGCCTCCATTTTGCCCGCGCTGAAGGAAAAGGGCATCCACGTGTCCAGCCCAGAAAGGGTAATCCTCGA : 1188
H6                   : CACCAAATCTATGGGACGATTCACCCCGAGGCGGAAGCCTCCATTTTGCCCGCGCTGAAGGAAAAGGGCATCCACGTGTCCAGCCCAGAAAGGGTAATCCTCGA : 1107
H7                   : CACCAAATCTATGGGACGATTCACCCCGAGGCGGAAGCCTCCATTTTGCCCGCGCTGAAGGAAAAGGGCATCCACGTGTCCAGCCCAGAAAGGGTAATCCTCGA : 1125
H8                   : CACCAAATCTATGGGACGATTCACCCCGAGGCGGAAGCCTCCATTTTGCCCGCGCTGAAGGAAAAGGGCATCCACGTGTCCAGCCCAGAAAGGGTAATCCTCGA : 1161
H9                   : CACCAAATCTATGGGACGATTCACCCCGAGGCGGAAGCCTCCATTTTGCCCGCGCTGAAGGAAAAGGGCATCCACGTGTCCAGCCCAGAAAGGGTAATCCTCGA : 1243
H10                  : CACCAAATCTATGGGACGATTCACCCCGAGGCGGAAGCCTCCATTTTGCCCGCGCTGAAGGAAAAGGGCATCCACGTGTCCAGCCCAGAAAGGGTAATCCTCGA : 1161
H11                  : CACCAAATCTATGGGACGATTCACCCCGAGGCGGAAGCCTCCATTTTGCCCGCGCTGAAGGAAAAGGGCATCCACGTGTCCAGCCCAGAAAGGGTAATCCTCGA : 1170
H12                  : CACCAAATCTATGGGACGATTCACCCCGAGGCGGAAGCCTCCATTTTGCCCGCGCTGAAGGAAAAGGGCATCCACGTGTCCAGCCCAGAAAGGGTAATCCTCGA : 1152
H13                  : CACCAAATCTATGGGACGATTCACCCCGAGGCGGAAGCCTCCATTTTGCCCGCGCTGAAGGAAAAGGGCATCCACGTGTCCAGCCCAGAAAGGGTAATCCTCGA : 1224
H14                  : CACCAAATCTATGGGACGATTCACCCCGAGGCGGAAGCCTCCATTTTGCCCGCGCTGAAGGAAAAGGGCATCCACGTGTCCAGCCCAGAAAGGGTAATCCTCGA : 1170
H15                  : CACCAAATCTATGGGACGATTCACCCCGAGGCGGAAGCCTCCATTTTGCCCGCGCTGAAGGAAAAGGGCATCCACGTGTCCAGCCCAGAAAGGGTAATCCTCGA : 1152
H16                  : CACCAAATCTATGGGACGATTCACCCCGAGGCGGAAGCCTCCATTTTGCCCGCGCTGAAGGAAAAGGGCATCCACGTGTCCAGCCCAGAAAGGGTAATCCTCGA : 1116
H17                  : CACCAAATCTATGGGACGATTCACCCCGAGGCGGAAGCCTCCATTTTGCCCGCGCTGAAGGAAAAGGGCATCCACGTGTCCAGCCCAGAAAGGGTAATCCTCGA : 1332
H18                  : CACCAAATCTATGGGACGATTCACCCCGAGGCGGAAGCCTCCATTTTGCCCGCGCTGAAGGAAAAGGGCATCCACGTGTCCAGCCCAGAAAGGGTAATCCTCGA : 1116
H19                  : CACCAAATCTATGGGACGATTCACCCCGAGGCGGAAGCCTCCATTTTGCCCGCGCTGAAGGAAAAGGGCATCCACGTGTCCAGCCCAGAAAGGGTAATCCTCGA : 1170
H20                  : CACCAAATCTATGGGACGATTCACCCCGAGGCGGAAGCCTCCATTTTGCCCGCGCTGAAGGAAAAGGGCATCCACGTGTCCAGCCCAGAAAGGGTAATCCTCGA : 1152
H21                  : CACCAAATCTATGGGACGATTCACCCCGAGGCGGAAGCCTCCATTTTGCCCGCGCTGAAGGAAAAGGGCATCCACGTGTCCAGCCCAGAAAGGGTAATCCTCGA : 1188
H22                  : CACCAAATCTATGGGACGATTCACCCCGAGGCGGAAGCCTCCATTTTGCCCGCGCTGAAGGAAAAGGGCATCCACGTGTCCAGCCCAGAAAGGGTAATCCTCGA : 1152
H23                  : CACCAAATCTATGGGACGATTCACCCCGAGGCGGAAGCCTCCATTTTGCCCGCGCTGAAGGAAAAGGGCATCCACGTGTCCAGCCCAGAAAGGGTAATCCTCGA : 1188
H24                  : CACCAAATCTATGGGACGATTCACCCCGAGGCGGAAGCCTCCATTTTGCCCGCGCTGAAGGAAAAGGGCATCCACGTGTCCAGCCCAGAAAGGGTAATCCTCGA : 1197
H25                  : CACCAAATCTATGGGACGATTCACCCCGAGGCGGAAGCCTCCATTTTGCCCGCGCTGAAGGAAAAGGGCATCCACGTGTCCAGCCCAGAAAGGGTAATCCTCGA : 1152
H26_Mauritania-I_ref : CACCAAATCTATGGGACGATTCACCCCGAGGCGGAAGCCTCCATTTTGCCCGCGCTGAAGGAAAAGGGCATCCACGTGTCCAGCCCAGAAAGGGTAATCCTCGA : 1242
H27_India-VII_refseq : CACCAAATCTATGGGACGATTCACCCCGAGGCGGAAGCCTCCATTTTGCCCGCGCTGAAGGAAAAGGGCATCCACGTGTCCAGCCCAGAAAGGGTAATCCTCGA : 1125
H28_Ctg_refsequence  : CACCAAATCTATGGGACGATTCACCCCGAGGCGGAAGCCTCCATTTTGCCCGCGCTGAAGGAAAAGGGCATCCACGTGTCCAGCCCAGAAAGGGTAATCCTCGA : 1197
H29_Salvador-I_refse : CACCAAATCTATGGGACGATTCACCCCGAGGCGGAAGCCTCCATTTTGCCCGCGCTGAAGGAAAAGGGCATCCACGTGTCCAGCCCAGAAAGGGTAATCCTCGA : 1197
```

```
H30_NKorea_refsequen : CACCAAATCTATGGGACGATTCACCCCGAGGCGGAAGCCTCCATTTTGCCCGCGCTGAAGGAAAAGGGCATCCACGTGTCCAGCCCAGAAAGGGTAATCCTCGA : 1107
H31_P.vivax01_refseq : CACCAAATCTATGGGACGATTCACCCCGAGGCGGAAGCCTCCATTTTGCCCGCGCTGAAGGAAAAGGGCATCCACGTGTCCAGCCCAGAAAGGGTAATCCTCGA : 1188
H32_Brazil-I_refsequ : CACCAAATCTATGGGACGATTCACCCCGAGGCGGAAGCCTCCATTTTGCCCGCGCTGAAGGAAAAGGGCATCCACGTGTCCAGCCCAGAAAGGGTAATCCTCGA : 1116
                       CACCAAATCTATGGGACGATTCACCCCGAGGCGGAAGCCTCCATTTTGCCCGCGCTGAAGGAAAAGGGCATCCACGTGTCCAGCCCAGAAAGGGTAATCCTCGA

                       1460         *      1480         *      1500         *      1520         *      1540         *      1560
H1_VCG               : GATCATCGAAAGCGCCAAGGAGGGAATAAAAGGATTGCTGAAGCTCCGGAAGAGCAAAGACACAGGCAAGCTGCTGGAGGAGGCACTAGAAAAGCTAAACATAA : 1220
H2                   : GATCATCGAAAGCGCCAAGGAGGGAATAAAAGGATTGCTGAAGCTCCGGAAGAGCAAAGACACAGGCAAGCTGCTGGAGGAGGCACTAGAAAAGCTAAACATAA : 1256
H3                   : GATCATCGAAAGCGCCAAGGAGGGAATAAAAGGATTGCTGAAGCTCCGGAAGAGCAAAGACACAGGCAAGCTGCTGGAGGAGGCACTAGAAAAGCTAAACATAA : 1328
H4                   : GATCATCGAAAGCGCCAAGGAGGGAATAAAAGGATTGCTGAAGCTCCGGAAGAGCAAAGACACAGGCAAGCTGCTGGAGGAGGCACTAGAAAAGCTAAACATAA : 1292
H5                   : GATCATCGAAAGCGCCAAGGAGGGAATAAAAGGATTGCTGAAGCTCCGGAAGAGCAAAGACACAGGCAAGCTGCTGGAGGAGGCACTAGAAAAGCTAAACATAA : 1292
H6                   : GATCATCGAAAGCGCCAAGGAGGGAATAAAAGGATTGCTGAAGCTCCGGAAGAGCAAAGACACAGGCAAGCTGCTGGAGGAGGCACTAGAAAAGCTAAACATAA : 1211
H7                   : GATCATCGAAAGCGCCAAGGAGGGAATAAAAGGATTGCTGAAGCTCCGGAAGAGCAAAGACACAGGCAAGCTGCTGGAGGAGGCACTAGAAAAGCTAAACATAA : 1229
H8                   : GATCATCGAAAGCGCCAAGGAGGGAATAAAAGGATTGCTGAAGCTCCGGAAGAGCAAAGACACAGGCAAGCTGCTGGAGGAGGCACTAGAAAAGCTAAACATAA : 1265
H9                   : GATCATCGAAAGCGCCAAGGAGGGAATAAAAGGATTGCTGAAGCTCCGGAAGAGCAAAGACACAGGCAAGCTGCTGGAGGAGGCACTAGAAAAGCTAAACATAA : 1347
H10                  : GATCATCGAAAGCGCCAAGGAGGGAATAAAAGGATTGCTGAAGCTCCGGAAGAGCAAAGACACAGGCAAGCTGCTGGAGGAGGCACTAGAAAAGCTAAACATAA : 1265
H11                  : GATCATCGAAAGCGCCAAGGAGGGAATAAAAGGATTGCTGAAGCTCCGGAAGAGCAAAGACACAGGCAAGCTGCTGGAGGAGGCACTAGAAAAGCTAAACATAA : 1274
H12                  : GATCATCGAAAGCGCCAAGGAGGGAATAAAAGGATTGCTGAAGCTCCGGAAGAGCAAAGACACAGGCAAGCTGCTGGAGGAGGCACTAGAAAAGCTAAACATAA : 1256
H13                  : GATCATCGAAAGCGCCAAGGAGGGAATAAAAGGATTGCTGAAGCTCCGGAAGAGCAAAGACACAGGCAAGCTGCTGGAGGAGGCACTAGAAAAGCTAAACATAA : 1328
H14                  : GATCATCGAAAGCGCCAAGGAGGGAATAAAAGGATTGCTGAAGCTCCGGAAGAGCAAAGACACAGGCAAGCTGCTGGAGGAGGCACTAGAAAAGCTAAACATAA : 1274
H15                  : GATCATCGAAAGCGCCAAGGAGGGAATAAAAGGATTGCTGAAGCTCCGGAAGAGCAAAGACACAGGCAAGCTGCTGGAGGAGGCACTAGAAAAGCTAAACATAA : 1256
H16                  : GATCATCGAAAGCGCCAAGGAGGGAATAAAAGGATTGCTGAAGCTCCGGAAGAGCAAAGACACAGGCAAGCTGCTGGAGGAGGCACTAGAAAAGCTAAACATAA : 1220
H17                  : GATCATCGAAAGCGCCAAGGAGGGAATAAAAGGATTGCTGAAGCTCCGGAAGAGCAAAGACACAGGCAAGCTGCTGGAGGAGGCACTAGAAAAGCTAAACATAA : 1436
H18                  : GATCATCGAAAGCGCCAAGGAGGGAATAAAAGGATTGCTGAAGCTCCGGAAGAGCAAAGACACAGGCAAGCTGCTGGAGGAGGCACTAGAAAAGCTAAACATAA : 1220
H19                  : GATCATCGAAAGCGCCAAGGAGGGAATAAAAGGATTGCTGAAGCTCCGGAAGAGCAAAGACACAGGCAAGCTGCTGGAGGAGGCACTAGAAAAGCTAAACATAA : 1274
H20                  : GATCATCGAAAGCGCCAAGGAGGGAATAAAAGGATTGCTGAAGCTCCGGAAGAGCAAAGACACAGGCAAGCTGCTGGAGGAGGCACTAGAAAAGCTAAACATAA : 1256
H21                  : GATCATCGAAAGCGCCAAGGAGGGAATAAAAGGATTGCTGAAGCTCCGGAAGAGCAAAGACACAGGCAAGCTGCTGGAGGAGGCACTAGAAAAGCTAAACATAA : 1292
H22                  : GATCATCGAAAGCGCCAAGGAGGGAATAAAAGGATTGCTGAAGCTCCGGAAGAGCAAAGACACAGGCAAGCTGCTGGAGGAGGCACTAGAAAAGCTAAACATAA : 1256
H23                  : GATCATCGAAAGCGCCAAGGAGGGAATAAAAGGATTGCTGAAGCTCCGGAAGAGCAAAGACACAGGCAAGCTGCTGGAGGAGGCACTAGAAAAGCTAAACATAA : 1292
H24                  : GATCATCGAAAGCGCCAAGGAGGGAATAAAAGGATTGCTGAAGCTCCGGAAGAGCAAAGACACAGGCAAGCTGCTGGAGGAGGCACTAGAAAAGCTAAACATAA : 1301
H25                  : GATCATCGAAAGCGCCAAGGAGGGAATAAAAGGATTGCTGAAGCTCCGGAAGAGCAAAGACACAGGCAAGCTGCTGGAGGAGGCACTAGAAAAGCTAAACATAA : 1256
H26_Mauritania-I_ref : GATCATCGAAAGCGCCAAGGAGGGAATAAAAGGATTGCTGAAGCTCCGGAAGAGCAAAGACACAGGCAAGCTGCTGGAGGAGGCACTAGAAAAGCTAAACATAA : 1346
H27_India-VII_refseq : GATCATCGAAAGCGCCAAGGAGGGAATAAAAGGATTGCTGAAGCTCCGGAAGAGCAAAGACACAGGCAAGCTGCTGGAGGAGGCACTAGAAAAGCTAAACATAA : 1229
H28_Ctg_refsequence  : GATCATCGAAAGCGCCAAGGAGGGAATAAAAGGATTGCTGAAGCTCCGGAAGAGCAAAGACACAGGCAAGCTGCTGGAGGAGGCACTAGAAAAGCTAAACATAA : 1301
H29_Salvador-I_refse : GATCATCGAAAGCGCCAAGGAGGGAATAAAAGGATTGCTGAAGCTCCGGAAGAGCAAAGACACAGGCAAGCTGCTGGAGGAGGCACTAGAAAAGCTAAACATAA : 1301
H30_NKorea_refsequen : GATCATCGAAAGCGCCAAGGAGGGAATAAAAGGATTGCTGAAGCTCCGGAAGAGCAAAGACACAGGCAAGCTGCTGGAGGAGGCACTAGAAAAGCTAAACATAA : 1211
H31_P.vivax01_refseq : GATCATCGAAAGCGCCAAGGAGGGAATAAAAGGATTGCTGAAGCTCCGGAAGAGCAAAGACACAGGCAAGCTGCTGGAGGAGGCACTAGAAAAGCTAAACATAA : 1292
H32_Brazil-I_refsequ : GATCATCGAAAGCGCCAAGGAGGGAATAAAAGGATTGCTGAAGCTCCGGAAGAGCAAAGACACAGGCAAGCTGCTGGAGGAGGCACTAGAAAAGCTAAACATAA : 1220
                       GATCATCGAAAGCGCCAAGGAGGGAATAAAAGGATTGCTGAAGCTCCGGAAGAGCAAAGACACAGGCAAGCTGCTGGAGGAGGCACTAGAAAAGCTAAACATAA

                                *      1580         *      1600         *      1620         *      1640         *      1660
H1_VCG               : ACCAAAGGGATCTGCATGCAAACAATAATTTCATTACCCTAGAAATGTATGACCGGATCCTCTCCGAAATGTTTAAAATATTAACAGAAATGTCATTCTATAAG : 1324
H2                   : ACCAAAGGGATCTGCATGCAAACAATAATTTCATTACCCTAGAAATGTATGACCGGATCCTCTCCGAAATGTTTAAAATATTAACAGAAATGTCATTCTATAAG : 1360
H3                   : ACCAAAGGGATCTGCATGCAAACAATAATTTCATTACCCTAGAAATGTATGACCGGATCCTCTCCGAAATGTTTAAAATATTAACAGAAATGTCATTCTATAAG : 1432
H4                   : ACCAAAGGGATCTGCATGCAAACAATAATTTCATTACCCTAGAAATGTATGACCGGATCCTCTCCGAAATGTTTAAAATATTAACAGAAATGTCATTCTATAAG : 1396
```

```
H5                    : ACCAAAGGGATCTGCATGCAAACAATAATTTCATTACCCTAGAAATGTATGACCGGATCCTCTCCGAAATGTTTAAAATATTAACAGAAATGTCATTCTATAAG : 1396
H6                    : ACCAAAGGGATCTGCATGCAAACAATAATTTCATTACCCTAGAAATGTATGACCGGATCCTCTCCGAAATGTTTAAAATATTAACAGAAATGTCATTCTATAAG : 1315
H7                    : ACCAAAGGGATCTGCATGCAAACAATAATTTCATTACCCTAGAAATGTATGACCGGATCCTCTCCGAAATGTTTAAAATATTAACAGAAATGTCATTCTATAAG : 1333
H8                    : ACCAAAGGGATCTGCATGCAAACAATAATTTCATTACCCTAGAAATGTATGACCGGATCCTCTCCGAAATGTTTAAAATATTAACAGAAATGTCATTCTATAAG : 1369
H9                    : ACCAAAGGGATCTGCATGCAAACAATAATTTCATTACCCTAGAAATGTATGACCGGATCCTCTCCGAAATGTTTAAAATATTAACAGAAATGTCATTCTATAAG : 1451
H10                   : ACCAAAGGGATCTGCATGCAAACAATAATTTCATTACCCTAGAAATGTATGACCGGATCCTCTCCGAAATGTTTAAAATATTAACAGAAATGTCATTCTATAAG : 1369
H11                   : ACCAAAGGGATCTGCATGCAAACAATAATTTCATTACCCTAGAAATGTATGACCGGATCCTCTCCGAAATGTTTAAAATATTAACAGAAATGTCATTCTATAAG : 1378
H12                   : ACCAAAGGGATCTGCATGCAAACAATAATTTCATTACCCTAGAAATGTATGACCGGATCCTCTCCGAAATGTTTAAAATATTAACAGAAATGTCATTCTATAAG : 1360
H13                   : ACCAAAGGGATCTGCATGCAAACAATAATTTCATTACCCTAGAAATGTATGACCGGATCCTCTCCGAAATGTTTAAAATATTAACAGAAATGTCATTCTATAAG : 1432
H14                   : ACCAAAGGGATCTGCATGCAAACAATAATTTCATTACCCTAGAAATGTATGACCGGATCCTCTCCGAAATGTTTAAAATATTAACAGAAATGTCATTCTATAAG : 1378
H15                   : ACCAAAGGGATCTGCATGCAAACAATAATTTCATTACCCTAGAAATGTATGACCGGATCCTCTCCGAAATGTTTAAAATATTAACAGAAATGTCATTCTATAAG : 1360
H16                   : ACCAAAGGGATCTGCATGCAAACAATAATTTCATTACCCTAGAAATGTATGACCGGATCCTCTCCGAAATGTTTAAAATATTAACAGAAATGTCATTCTATAAG : 1324
H17                   : ACCAAAGGGATCTGCATGCAAACAATAATTTCATTACCCTAGAAATGTATGACCGGATCCTCTCCGAAATGTTTAAAATATTAACAGAAATGTCATTCTATAAG : 1540
H18                   : ACCAAAGGGATCTGCATGCAAACAATAATTTCATTACCCTAGAAATGTATGACCGGATCCTCTCCGAAATGTTTAAAATATTAACAGAAATGTCATTCTATAAG : 1324
H19                   : ACCAAAGGGATCTGCATGCAAACAATAATTTCATTACCCTAGAAATGTATGACCGGATCCTCTCCGAAATGTTTAAAATATTAACAGAAATGTCATTCTATAAG : 1378
H20                   : ACCAAAGGGATCTGCATGCAAACAATAATTTCATTACCCTAGAAATGTATGACCGGATCCTCTCCGAAATGTTTAAAATATTAACAGAAATGTCATTCTATAAG : 1360
H21                   : ACCAAAGGGATCTGCATGCAAACAATAATTTCATTACCCTAGAAATGTATGACCGGATCCTCTCCGAAATGTTTAAAATATTAACAGAAATGTCATTCTATAAG : 1396
H22                   : ACCAAAGGGATCTGCATGCAAACAATAATTTCATTACCCTAGAAATGTATGACCGGATCCTCTCCGAAATGTTTAAAATATTAACAGAAATGTCATTCTATAAG : 1360
H23                   : ACCAAAGGGATCTGCATGCAAACAATAATTTCATTACCCTAGAAATGTATGACCGGATCCTCTCCGAAATGTTTAAAATATTAACAGAAATGTCATTCTATAAG : 1396
H24                   : ACCAAAGGGATCTGCATGCAAACAATAATTTCATTACCCTAGAAATGTATGACCGGATCCTCTCCGAAATGTTTAAAATATTAACAGAAATGTCATTCTATAAG : 1405
H25                   : ACCAAAGGGATCTGCATGCAAACAATAATTTCATTACCCTAGAAATGTATGACCGGATCCTCTCCGAAATGTTTAAAATATTAACAGAAATGTCATTCTATAAG : 1360
H26_Mauritania-I_ref  : ACCAAAGGGATCTGCATGCAAACAATAATTTCATTACCCTAGAAATGTATGACCGGATCCTCTCCGAAATGTTTAAAATATTAACAGAAATGTCATTCTATAAG : 1450
H27_India-VII_refseq  : ACCAAAGGGATCTGCATGCAAACAATAATTTCATTACCCTAGAAATGTATGACCGGATCCTCTCCGAAATGTTTAAAATATTAACAGAAATGTCATTCTATAAG : 1333
H28_Ctg_refsequence   : ACCAAAGGGATCTGCATGCAAACAATAATTTCATTACCCTAGAAATGTATGACCGGATCCTCTCCGAAATGTTTAAAATATTAACAGAAATGTCATTCTATAAG : 1405
H29_Salvador-I_refse  : ACCAAAGGGATCTGCATGCAAACAATAATTTCATTACCCTAGAAATGTATGACCGGATCCTCTCCGAAATGTTTAAAATATTAACAGAAATGTCATTCTATAAG : 1405
H30_NKorea_refsequen  : ACCAAAGGGATCTGCATGCAAACAATAATTTCATTACCCTAGAAATGTATGACCGGATCCTCTCCGAAATGTTTAAAATATTAACAGAAATGTCATTCTATAAG : 1315
H31_P.vivax01_refseq  : ACCAAAGGGATCTGCATGCAAACAATAATTTCATTACCCTAGAAATGTATGACCGGATCCTCTCCGAAATGTTTAAAATATTAACAGAAATGTCATTCTATAAG : 1396
H32_Brazil-I_refsequ  : ACCAAAGGGATCTGCATGCAAACAATAATTTCATTACCCTAGAAATGTATGACCGGATCCTCTCCGAAATGTTTAAAATATTAACAGAAATGTCATTCTATAAG : 1324
                        ACCAAAGGGATCTGCATGCAAACAATAATTTCATTACCCTAGAAATGTATGACCGGATCCTCTCCGAAATGTTTAAAATATTAACAGAAATGTCATTCTATAAG

                              *      1680         *      1700         *      1720         *      1740         *      1760
H1_VCG                : GATGGCCATTTTTACGAAACGTTAGGTTTGAATAAATCTATACTAAATCAATCGCTGAAAGAAATTAAAATAAAAATGTTGAGGACCATTGGAGTGCCATACAC : 1428
H2                    : GATGGCCATTTTTACGAAACGTTAGGTTTGAATAAATCTATACTAAATCAATCGCTGAAAGAAATTAAAATAAAAATGTTGAGGACCATTGGAGTGCCATACAC : 1464
H3                    : GATGGCCATTTTTACGAAACGTTAGGTTTGAATAAATCTATACTAAATCAATCGCTGAAAGAAATTAAAATAAAAATGTTGAGGACCATTGGAGTGCCATACAC : 1536
H4                    : GATGGCCATTTTTACGAAACGTTAGGTTTGAATAAATCTATACTAAATCAATCGCTGAAAGAAATTAAAATAAAAATGTTGAGGACCATTGGAGTGCCATACAC : 1500
H5                    : GATGGCCATTTTTACGAAACGTTAGGTTTGAATAAATCTATACTAAATCAATCGCTGAAAGAAATTAAAATAAAAATGTTGAGGACCATTGGAGTGCCATACAC : 1500
H6                    : GATGGCCATTTTTACGAAACGTTAGGTTTGAATAAATCTATACTAAATCAATCGCTGAAAGAAATTAAAATAAAAATGTTGAGGACCATTGGAGTGCCATACAC : 1419
H7                    : GATGGCCATTTTTACGAAACGTTAGGTTTGAATAAATCTATACTAAATCAATCGCTGAAAGAAATTAAAATAAAAATGTTGAGGACCATTGGAGTGCCATACAC : 1437
H8                    : GATGGCCATTTTTACGAAACGTTAGGTTTGAATAAATCTATACTAAATCAATCGCTGAAAGAAATTAAAATAAAAATGTTGAGGACCATTGGAGTGCCATACAC : 1473
H9                    : GATGGCCATTTTTACGAAACGTTAGGTTTGAATAAATCTATACTAAATCAATCGCTGAAAGAAATTAAAATAAAAATGTTGAGGACCATTGGAGTGCCATACAC : 1555
H10                   : GATGGCCATTTTTACGAAACGTTAGGTTTGAATAAATCTATACTAAATCAATCGCTGAAAGAAATTAAAATAAAAATGTTGAGGACCATTGGAGTGCCATACAC : 1473
H11                   : GATGGCCATTTTTACGAAACGTTAGGTTTGAATAAATCTATACTAAATCAATCGCTGAAAGAAATTAAAATAAAAATGTTGAGGACCATTGGAGTGCCATACAC : 1482
H12                   : GATGGCCATTTTTACGAAACGTTAGGTTTGAATAAATCTATACTAAATCAATCGCTGAAAGAAATTAAAATAAAAATGTTGAGGACCATTGGAGTGCCATACAC : 1464
H13                   : GATGGCCATTTTTACGAAACGTTAGGTTTGAATAAATCTATACTAAATCAATCGCTGAAAGAAATTAAAATAAAAATGTTGAGGACCATTGGAGTGCCATACAC : 1536
H14                   : GATGGCCATTTTTACGAAACGTTAGGTTTGAATAAATCTATACTAAATCAATCGCTGAAAGAAATTAAAATAAAAATGTTGAGGACCATTGGAGTGCCATACAC : 1482
```

```
H15                  : GATGGCCATTTTTACGAAACGTTAGGTTTGAATAAATCTATACTAAATCAATCGCTGAAAGAAATTAAAATAAAAATGTTGAGGACCATTGGAGTGCCATACAC : 1464
H16                  : GATGGCCATTTTTACGAAACGTTAGGTTTGAATAAATCTATACTAAATCAATCGCTGAAAGAAATTAAAATAAAAATGTTGAGGACCATTGGAGTGCCATACAC : 1428
H17                  : GATGGCCATTTTTACGAAACGTTAGGTTTGAATAAATCTATACTAAATCAATCGCTGAAAGAAATTAAAATAAAAATGTTGAGGACCATTGGAGTGCCATACAC : 1644
H18                  : GATGGCCATTTTTACGAAACGTTAGGTTTGAATAAATCTATACTAAATCAATCGCTGAAAGAAATTAAAATAAAAATGTTGAGGACCATTGGAGTGCCATACAC : 1428
H19                  : GATGGCCATTTTTACGAAACGTTAGGTTTGAATAAATCTATACTAAATCAATCGCTGAAAGAAATTAAAATAAAAATGTTGAGGACCATTGGAGTGCCATACAC : 1482
H20                  : GATGGCCATTTTTACGAAACGTTAGGTTTGAATAAATCTATACTAAATCAATCGCTGAAAGAAATTAAAATAAAAATGTTGAGGACCATTGGAGTGCCATACAC : 1464
H21                  : GATGGCCATTTTTACGAAACGTTAGGTTTGAATAAATCTATACTAAATCAATCGCTGAAAGAAATTAAAATAAAAATGTTGAGGACCATTGGAGTGCCATACAC : 1500
H22                  : GATGGCCATTTTTACGAAACGTTAGGTTTGAATAAATCTATACTAAATCAATCGCTGAAAGAAATTAAAATAAAAATGTTGAGGACCATTGGAGTGCCATACAC : 1464
H23                  : GATGGCCATTTTTACGAAACGTTAGGTTTGAATAAATCTATACTAAATCAATCGCTGAAAGAAATTAAAATAAAAATGTTGAGGACCATTGGAGTGCCATACAC : 1500
H24                  : GATGGCCATTTTTACGAAACGTTAGGTTTGAATAAATCTATACTAAATCAATCGCTGAAAGAAATTAAAATAAAAATGTTGAGGACCATTGGAGTGCCATACAC : 1509
H25                  : GATGGCCATTTTTACGAAACGTTAGGTTTGAATAAATCTATACTAAATCAATCGCTGAAAGAAATTAAAATAAAAATGTTGAGGACCATTGGAGTGCCATACAC : 1464
H26_Mauritania-I_ref : GATGGCCATTTTTACGAAACGTTAGGTTTGAATAAATCTATACTAAATCAATCGCTGAAAGAAATTAAAATAAAAATGTTGAGGACCATTGGAGTGCCATACAC : 1554
H27_India-VII_refseq : GATGGCCATTTTTACGAAACGTTAGGTTTGAATAAATCTATACTAAATCAATCGCTGAAAGAAATTAAAATAAAAATGTTGAGGACCATTGGAGTGCCATACAC : 1437
H28_Ctg_refsequence  : GATGGCCATTTTTACGAAACGTTAGGTTTGAATAAATCTATACTAAATCAATCGCTGAAAGAAATTAAAATAAAAATGTTGAGGACCATTGGAGTGCCATACAC : 1509
H29_Salvador-I_refse : GATGGCCATTTTTACGAAACGTTAGGTTTGAATAAATCTATACTAAATCAATCGCTGAAAGAAATTAAAATAAAAATGTTGAGGACCATTGGAGTGCCATACAC : 1509
H30_NKorea_refsequen : GATGGCCATTTTTACGAAACGTTAGGTTTGAATAAATCTATACTAAATCAATCGCTGAAAGAAATTAAAATAAAAATGTTGAGGACCATTGGAGTGCCATACAC : 1419
H31_P.vivax01_refseq : GATGGCCATTTTTACGAAACGTTAGGTTTGAATAAATCTATACTAAATCAATCGCTGAAAGAAATTAAAATAAAAATGTTGAGGACCATTGGAGTGCCATACAC : 1500
H32_Brazil-I_refsequ : GATGGCCATTTTTACGAAACGTTAGGTTTGAATAAATCTATACTAAATCAATCGCTGAAAGAAATTAAAATAAAAATGTTGAGGACCATTGGAGTGCCATACAC : 1428
                       GATGGCCATTTTTACGAAACGTTAGGTTTGAATAAATCTATACTAAATCAATCGCTGAAAGAAATTAAAATAAAAATGTTGAGGACCATTGGAGTGCCATACAC

                        *      1780         *      1800         *      1820         *      1840         *      1860         *
H1_VCG               : GAAGCTACCGCCAATCGTTAAGAACAAAGAGAAGGAAAGCACATGTGCTGCGAACAATCTCATCATTAGCATTACATCAAAAGAGCTAGCTCAACGAATGGCTA : 1532
H2                   : GAAGCTACCGCCAATCGTTAAGAACAAAGAGAAGGAAAGCACATGTGCTGCGAACAATCTCATCATTAGCATTACATCAAAAGAGCTAGCTCAACGAATGGCTA : 1568
H3                   : GAAGCTACCGCCAATCGTTAAGAACAAAGAGAAGGAAAGCACATGTGCTGCGAACAATCTCATCATTAGCATTACATCAAAAGAGCTAGCTCAACGAATGGCTA : 1640
H4                   : GAAGCTACCGCCAATCGTTAAGAACAAAGAGAAGGAAAGCACATGTGCTGCGAACAATCTCATCATTAGCATTACATCAAAAGAGCTAGCTCAACGAATGGCTA : 1604
H5                   : GAAGCTACCGCCAATCGTTAAGAACAAAGAGAAGGAAAGCACATGTGCTGCGAACAATCTCATCATTAGCATTACATCAAAAGAGCTAGCTCAACGAATGGCTA : 1604
H6                   : GAAGCTACCGCCAATCGTTAAGAACAAAGAGAAGGAAAGCACATGTGCTGCGAACAATCTCATCATTAGCATTACATCAAAAGAGCTAGCTCAACGAATGGCTA : 1523
H7                   : GAAGCTACCGCCAATCGTTAAGAACAAAGAGAAGGAAAGCACATGTGCTGCGAACAATCTCATCATTAGCATTACATCAAAAGAGCTAGCTCAACGAATGGCTA : 1541
H8                   : GAAGCTACCGCCAATCGTTAAGAACAAAGAGAAGGAAAGCACATGTGCTGCGAACAATCTCATCATTAGCATTACATCAAAAGAGCTAGCTCAACGAATGGCTA : 1577
H9                   : GAAGCTACCGCCAATCGTTAAGAACAAAGAGAAGGAAAGCACATGTGCTGCGAACAATCTCATCATTAGCATTACATCAAAAGAGCTAGCTCAACGAATGGCTA : 1659
H10                  : GAAGCTACCGCCAATCGTTAAGAACAAAGAGAAGGAAAGCACATGTGCTGCGAACAATCTCATCATTAGCATTACATCAAAAGAGCTAGCTCAACGAATGGCTA : 1577
H11                  : GAAGCTACCGCCAATCGTTAAGAACAAAGAGAAGGAAAGCACATGTGCTGCGAACAATCTCATCATTAGCATTACATCAAAAGAGCTAGCTCAACGAATGGCTA : 1586
H12                  : GAAGCTACCGCCAATCGTTAAGAACAAAGAGAAGGAAAGCACATGTGCTGCGAACAATCTCATCATTAGCATTACATCAAAAGAGCTAGCTCAACGAATGGCTA : 1568
H13                  : GAAGCTACCGCCAATCGTTAAGAACAAAGAGAAGGAAAGCACATGTGCTGCGAACAATCTCATCATTAGCATTACATCAAAAGAGCTAGCTCAACGAATGGCTA : 1640
H14                  : GAAGCTACCGCCAATCGTTAAGAACAAAGAGAAGGAAAGCACATGTGCTGCGAACAATCTCATCATTAGCATTACATCAAAAGAGCTAGCTCAACGAATGGCTA : 1586
H15                  : GAAGCTACCGCCAATCGTTAAGAACAAAGAGAAGGAAAGCACATGTGCTGCGAACAATCTCATCATTAGCATTACATCAAAAGAGCTAGCTCAACGAATGGCTA : 1568
H16                  : GAAGCTACCGCCAATCGTTAAGAACAAAGAGAAGGAAAGCACATGTGCTGCGAACAATCTCATCATTAGCATTACATCAAAAGAGCTAGCTCAACGAATGGCTA : 1532
H17                  : GAAGCTACCGCCAATCGTTAAGAACAAAGAGAAGGAAAGCACATGTGCTGCGAACAATCTCATCATTAGCATTACATCAAAAGAGCTAGCTCAACGAATGGCTA : 1748
H18                  : GAAGCTACCGCCAATCGTTAAGAACAAAGAGAAGGAAAGCACATGTGCTGCGAACAATCTCATCATTAGCATTACATCAAAAGAGCTAGCTCAACGAATGGCTA : 1532
H19                  : GAAGCTACCGCCAATCGTTAAGAACAAAGAGAAGGAAAGCACATGTGCTGCGAACAATCTCATCATTAGCATTACATCAAAAGAGCTAGCTCAACGAATGGCTA : 1586
H20                  : GAAGCTACCGCCAATCGTTAAGAACAAAGAGAAGGAAAGCACATGTGCTGCGAACAATCTCATCATTAGCATTACATCAAAAGAGCTAGCTCAACGAATGGCTA : 1568
H21                  : GAAGCTACCGCCAATCGTTAAGAACAAAGAGAAGGAAAGCACATGTGCTGCGAACAATCTCATCATTAGCATTACATCAAAAGAGCTAGCTCAACGAATGGCTA : 1604
H22                  : GAAGCTACCGCCAATCGTTAAGAACAAAGAGAAGGAAAGCACATGTGCTGCGAACAATCTCATCATTAGCATTACATCAAAAGAGCTAGCTCAACGAATGGCTA : 1568
H23                  : GAAGCTACCGCCAATCGTTAAGAACAAAGAGAAGGAAAGCACATGTGCTGCGAACAATCTCATCATTAGCATTACATCAAAAGAGCTAGCTCAACGAATGGCTA : 1604
H24                  : GAAGCTACCGCCAATCGTTAAGAACAAAGAGAAGGAAAGCACATGTGCTGCGAACAATCTCATCATTAGCATTACATCAAAAGAGCTAGCTCAACGAATGGCTA : 1613
```

```
H25                  : GAAGCTACCGCCAATCGTTAAGAACAAAGAGAAGGAAAGCACATGTGCTGCGAACAATCTCATCATTAGCATTACATCAAAAGAGCTAGCTCAACGAATGGCTA : 1568
H26_Mauritania-I_ref : GAAGCTACCGCCAATCGTTAAGAACAAAGAGAAGGAAAGCACATGTGCTGCGAACAATCTCATCATTAGCATTACATCAAAAGAGCTAGCTCAACGAATGGCTA : 1658
H27_India-VII_refseq : GAAGCTACCGCCAATCGTTAAGAACAAAGAGAAGGAAAGCACATGTGCTGCGAACAATCTCATCATTAGCATTACATCAAAAGAGCTAGCTCAACGAATGGCTA : 1541
H28_Ctg_refsequence  : GAAGCTACCGCCAATCGTTAAGAACAAAGAGAAGGAAAGCACATGTGCTGCGAACAATCTCATCATTAGCATTACATCAAAAGAGCTAGCTCAACGAATGGCTA : 1613
H29_Salvador-I_refse : GAAGCTACCGCCAATCGTTAAGAACAAAGAGAAGGAAAGCACATGTGCTGCGAACAATCTCATCATTAGCATTACATCAAAAGAGCTAGCTCAACGAATGGCTA : 1613
H30_NKorea_refsequen : GAAGCTACCGCCAATCGTTAAGAACAAAGAGAAGGAAAGCACATGTGCTGCGAACAATCTCATCATTAGCATTACATCAAAAGAGCTAGCTCAACGAATGGCTA : 1523
H31_P.vivax01_refseq : GAAGCTACCGCCAATCGTTAAGAACAAAGAGAAGGAAAGCACATGTGCTGCGAACAATCTCATCATTAGCATTACATCAAAAGAGCTAGCTCAACGAATGGCTA : 1604
H32_Brazil-I_refsequ : GAAGCTACCGCCAATCGTTAAGAACAAAGAGAAGGAAAGCACATGTGCTGCGAACAATCTCATCATTAGCATTACATCAAAAGAGCTAGCTCAACGAATGGCTA : 1532
                       GAAGCTACCGCCAATCGTTAAGAACAAAGAGAAGGAAAGCACATGTGCTGCGAACAATCTCATCATTAGCATTACATCAAAAGAGCTAGCTCAACGAATGGCTA

                           1880         *      1900         *      1920         *      1940         *      1960         *
H1_VCG               : TCATGTTCGCAAAATGGCTAGCTCCAGAAGAATACGGGTCTGTCGTCGATTTAGATAAAAGCATAGAATTAAATGTCCTCTGCGCAGGGGCACCTATCCTAGTA : 1636
H2                   : TCATGTTCGCAAAATGGCTAGCTCCAGAAGAATACGGGTCTGTCGTCGATTTAGATAAAAGCATAGAATTAAATGTCCTCTGCGCAGGGGCACCTATCCTAGTA : 1672
H3                   : TCATGTTCGCAAAATGGCTAGCTCCAGAAGAATACGGGTCTGTCGTCGATTTAGATAAAAGCATAGAATTAAATGTCCTCTGCGCAGGGGCACCTATCCTAGTA : 1744
H4                   : TCATGTTCGCAAAATGGCTAGCTCCAGAAGAATACGGGTCTGTCGTCGATTTAGATAAAAGCATAGAATTAAATGTCCTCTGCGCAGGGGCACCTATCCTAGTA : 1708
H5                   : TCATGTTCGCAAAATGGCTAGCTCCAGAAGAATACGGGTCTGTCGTCGATTTAGATAAAAGCATAGAATTAAATGTCCTCTGCGCAGGGGCACCTATCCTAGTA : 1708
H6                   : TCATGTTCGCAAAATGGCTAGCTCCAGAAGAATACGGGTCTGTCGTCGATTTAGATAAAAGCATAGAATTAAATGTCCTCTGCGCAGGGGCACCTATCCTAGTA : 1627
H7                   : TCATGTTCGCAAAATGGCTAGCTCCAGAAGAATACGGGTCTGTCGTCGATTTAGATAAAAGCATAGAATTAAATGTCCTCTGCGCAGGGGCACCTATCCTAGTA : 1645
H8                   : TCATGTTCGCAAAATGGCTAGCTCCAGAAGAATACGGGTCTGTCGTCGATTTAGATAAAAGCATAGAATTAAATGTCCTCTGCGCAGGGGCACCTATCCTAGTA : 1681
H9                   : TCATGTTCGCAAAATGGCTAGCTCCAGAAGAATACGGGTCTGTCGTCGATTTAGATAAAAGCATAGAATTAAATGTCCTCTGCGCAGGGGCACCTATCCTAGTA : 1763
H10                  : TCATGTTCGCAAAATGGCTAGCTCCAGAAGAATACGGGTCTGTCGTCGATTTAGATAAAAGCATAGAATTAAATGTCCTCTGCGCAGGGGCACCTATCCTAGTA : 1681
H11                  : TCATGTTCGCAAAATGGCTAGCTCCAGAAGAATACGGGTCTGTCGTCGATTTAGATAAAAGCATAGAATTAAATGTCCTCTGCGCAGGGGCACCTATCCTAGTA : 1690
H12                  : TCATGTTCGCAAAATGGCTAGCTCCAGAAGAATACGGGTCTGTCGTCGATTTAGATAAAAGCATAGAATTAAATGTCCTCTGCGCAGGGGCACCTATCCTAGTA : 1672
H13                  : TCATGTTCGCAAAATGGCTAGCTCCAGAAGAATACGGGTCTGTCGTCGATTTAGATAAAAGCATAGAATTAAATGTCCTCTGCGCAGGGGCACCTATCCTAGTA : 1744
H14                  : TCATGTTCGCAAAATGGCTAGCTCCAGAAGAATACGGGTCTGTCGTCGATTTAGATAAAAGCATAGAATTAAATGTCCTCTGCGCAGGGGCACCTATCCTAGTA : 1690
H15                  : TCATGTTCGCAAAATGGCTAGCTCCAGAAGAATACGGGTCTGTCGTCGATTTAGATAAAAGCATAGAATTAAATGTCCTCTGCGCAGGGGCACCTATCCTAGTA : 1672
H16                  : TCATGTTCGCAAAATGGCTAGCTCCAGAAGAATACGGGTCTGTCGTCGATTTAGATAAAAGCATAGAATTAAATGTCCTCTGCGCAGGGGCACCTATCCTAGTA : 1636
H17                  : TCATGTTCGCAAAATGGCTAGCTCCAGAAGAATACGGGTCTGTCGTCGATTTAGATAAAAGCATAGAATTAAATGTCCTCTGCGCAGGGGCACCTATCCTAGTA : 1852
H18                  : TCATGTTCGCAAAATGGCTAGCTCCAGAAGAATACGGGTCTGTCGTCGATTTAGATAAAAGCATAGAATTAAATGTCCTCTGCGCAGGGGCACCTATCCTAGTA : 1636
H19                  : TCATGTTCGCAAAATGGCTAGCTCCAGAAGAATACGGGTCTGTCGTCGATTTAGATAAAAGCATAGAATTAAATGTCCTCTGCGCAGGGGCACCTATCCTAGTA : 1690
H20                  : TCATGTTCGCAAAATGGCTAGCTCCAGAAGAATACGGGTCTGTCGTCGATTTAGATAAAAGCATAGAATTAAATGTCCTCTGCGCAGGGGCACCTATCCTAGTA : 1672
H21                  : TCATGTTCGCAAAATGGCTAGCTCCAGAAGAATACGGGTCTGTCGTCGATTTAGATAAAAGCATAGAATTAAATGTCCTCTGCGCAGGGGCACCTATCCTAGTA : 1708
H22                  : TCATGTTCGCAAAATGGCTAGCTCCAGAAGAATACGGGTCTGTCGTCGATTTAGATAAAAGCATAGAATTAAATGTCCTCTGCGCAGGGGCACCTATCCTAGTA : 1672
H23                  : TCATGTTCGCAAAATGGCTAGCTCCAGAAGAATACGGGTCTGTCGTCGATTTAGATAAAAGCATAGAATTAAATGTCCTCTGCGCAGGGGCACCTATCCTAGTA : 1708
H24                  : TCATGTTCGCAAAATGGCTAGCTCCAGAAGAATACGGGTCTGTCGTCGATTTAGATAAAAGCATAGAATTAAATGTCCTCTGCGCAGGGGCACCTATCCTAGTA : 1717
H25                  : TCATGTTCGCAAAATGGCTAGCTCCAGAAGAATACGGGTCTGTCGTCGATTTAGATAAAAGCATAGAATTAAATGTCCTCTGCGCAGGGGCACCTATCCTAGTA : 1672
H26_Mauritania-I_ref : TCATGTTCGCAAAATGGCTAGCTCCAGAAGAATACGGGTCTGTCGTCGATTTAGATAAAAGCATAGAATTAAATGTCCTCTGCGCAGGGGCACCTATCCTAGTA : 1762
H27_India-VII_refseq : TCATGTTCGCAAAATGGCTAGCTCCAGAAGAATACGGGTCTGTCGTCGATTTAGATAAAAGCATAGAATTAAATGTCCTCTGCGCAGGGGCACCTATCCTAGTA : 1645
H28_Ctg_refsequence  : TCATGTTCGCAAAATGGCTAGCTCCAGAAGAATACGGGTCTGTCGTCGATTTAGATAAAAGCATAGAATTAAATGTCCTCTGCGCAGGGGCACCTATCCTAGTA : 1717
H29_Salvador-I_refse : TCATGTTCGCAAAATGGCTAGCTCCAGAAGAATACGGGTCTGTCGTCGATTTAGATAAAAGCATAGAATTAAATGTCCTCTGCGCAGGGGCACCTATCCTAGTA : 1717
H30_NKorea_refsequen : TCATGTTCGCAAAATGGCTAGCTCCAGAAGAATACGGGTCTGTCGTCGATTTAGATAAAAGCATAGAATTAAATGTCCTCTGCGCAGGGGCACCTATCCTAGTA : 1627
H31_P.vivax01_refseq : TCATGTTCGCAAAATGGCTAGCTCCAGAAGAATACGGGTCTGTCGTCGATTTAGATAAAAGCATAGAATTAAATGTCCTCTGCGCAGGGGCACCTATCCTAGTA : 1708
H32_Brazil-I_refsequ : TCATGTTCGCAAAATGGCTAGCTCCAGAAGAATACGGGTCTGTCGTCGATTTAGATAAAAGCATAGAATTAAATGTCCTCTGCGCAGGGGCACCTATCCTAGTA : 1636
                       TCATGTTCGCAAAATGGCTAGCTCCAGAAGAATACGGGTCTGTCGTCGATTTAGATAAAAGCATAGAATTAAATGTCCTCTGCGCAGGGGCACCTATCCTAGTA
```

```
                         1980        *      2000        *      2020        *      2040        *      2060        *      2080
H1_VCG                 : CAGCAATGGAAATATTACCAAAACATGTTAGGGTTCGAAACGGGAAATGAACATGCCTTCTTAAACCTCATCGATGAGTTGCTAGTCATCGACAAGAGACACAG : 1740
H2                     : CAGCAATGGAAATATTACCAAAACATGTTAGGGTTCGAAACGGGAAATGAACATGCCTTCTTAAACCTCATCGATGAGTTGCTAGTCATCGACAAGAGACACAG : 1776
H3                     : CAGCAATGGAAATATTACCAAAACATGTTAGGGTTCGAAACGGGAAATGAACATGCCTTCTTAAACCTCATCGATGAGTTGCTAGTCATCGACAAGAGACACAG : 1848
H4                     : CAGCAATGGAAATATTACCAAAACATGTTAGGGTTCGAAACGGGAAATGAACATGCCTTCTTAAACCTCATCGATGAGTTGCTAGTCATCGACAAGAGACACAG : 1812
H5                     : CAGCAATGGAAATATTACCAAAACATGTTAGGGTTCGAAACGGGAAATGAACATGCCTTCTTAAACCTCATCGATGAGTTGCTAGTCATCGACAAGAGACACAG : 1812
H6                     : CAGCAATGGAAATATTACCAAAACATGTTAGGGTTCGAAACGGGAAATGAACATGCCTTCTTAAACCTCATCGATGAGTTGCTAGTCATCGACAAGAGACACAG : 1731
H7                     : CAGCAATGGAAATATTACCAAAACATGTTAGGGTTCGAAACGGGAAATGAACATGCCTTCTTAAACCTCATCGATGAGTTGCTAGTCATCGACAAGAGACACAG : 1749
H8                     : CAGCAATGGAAATATTACCAAAACATGTTAGGGTTCGAAACGGGAAATGAACATGCCTTCTTAAACCTCATCGATGAGTTGCTAGTCATCGACAAGAGACACAG : 1785
H9                     : CAGCAATGGAAATATTACCAAAACATGTTAGGGTTCGAAACGGGAAATGAACATGCCTTCTTAAACCTCATCGATGAGTTGCTAGTCATCGACAAGAGACACAG : 1867
H10                    : CAGCAATGGAAATATTACCAAAACATGTTAGGGTTCGAAACGGGAAATGAACATGCCTTCTTAAACCTCATCGATGAGTTGCTAGTCATCGACAAGAGACACAG : 1785
H11                    : CAGCAATGGAAATATTACCAAAACATGTTAGGGTTCGAAACGGGAAATGAACATGCCTTCTTAAACCTCATCGATGAGTTGCTAGTCATCGACAAGAGACACAG : 1794
H12                    : CAGCAATGGAAATATTACCAAAACATGTTAGGGTTCGAAACGGGAAATGAACATGCCTTCTTAAACCTCATCGATGAGTTGCTAGTCATCGACAAGAGACACAG : 1776
H13                    : CAGCAATGGAAATATTACCAAAACATGTTAGGGTTCGAAACGGGAAATGAACATGCCTTCTTAAACCTCATCGATGAGTTGCTAGTCATCGACAAGAGACACAG : 1848
H14                    : CAGCAATGGAAATATTACCAAAACATGTTAGGGTTCGAAACGGGAAATGAACATGCCTTCTTAAACCTCATCGATGAGTTGCTAGTCATCGACAAGAGACACAG : 1794
H15                    : CAGCAATGGAAATATTACCAAAACATGTTAGGGTTCGAAACGGGAAATGAACATGCCTTCTTAAACCTCATCGATGAGTTGCTAGTCATCGACAAGAGACACAG : 1776
H16                    : CAGCAATGGAAATATTACCAAAACATGTTAGGGTTCGAAACGGGAAATGAACATGCCTTCTTAAACCTCATCGATGAGTTGCTAGTCATCGACAAGAGACACAG : 1740
H17                    : CAGCAATGGAAATATTACCAAAACATGTTAGGGTTCGAAACGGGAAATGAACATGCCTTCTTAAACCTCATCGATGAGTTGCTAGTCATCGACAAGAGACACAG : 1956
H18                    : CAGCAATGGAAATATTACCAAAACATGTTAGGGTTCGAAACGGGAAATGAACATGCCTTCTTAAACCTCATCGATGAGTTGCTAGTCATCGACAAGAGACACAG : 1740
H19                    : CAGCAATGGAAATATTACCAAAACATGTTAGGGTTCGAAACGGGAAATGAACATGCCTTCTTAAACCTCATCGATGAGTTGCTAGTCATCGACAAGAGACACAG : 1794
H20                    : CAGCAATGGAAATATTACCAAAACATGTTAGGGTTCGAAACGGGAAATGAACATGCCTTCTTAAACCTCATCGATGAGTTGCTAGTCATCGACAAGAGACACAG : 1776
H21                    : CAGCAATGGAAATATTACCAAAACATGTTAGGGTTCGAAACGGGAAATGAACATGCCTTCTTAAACCTCATCGATGAGTTGCTAGTCATCGACAAGAGACACAG : 1812
H22                    : CAGCAATGGAAATATTACCAAAACATGTTAGGGTTCGAAACGGGAAATGAACATGCCTTCTTAAACCTCATCGATGAGTTGCTAGTCATCGACAAGAGACACAG : 1776
H23                    : CAGCAATGGAAATATTACCAAAACATGTTAGGGTTCGAAACGGGAAATGAACATGCCTTCTTAAACCTCATCGATGAGTTGCTAGTCATCGACAAGAGACACAG : 1812
H24                    : CAGCAATGGAAATATTACCAAAACATGTTAGGGTTCGAAACGGGAAATGAACATGCCTTCTTAAACCTCATCGATGAGTTGCTAGTCATCGACAAGAGACACAG : 1821
H25                    : CAGCAATGGAAATATTACCAAAACATGTTAGGGTTCGAAACGGGAAATGAACATGCCTTCTTAAACCTCATCGATGAGTTGCTAGTCATCGACAAGAGACACAG : 1776
H26_Mauritania-I_ref   : CAGCAATGGAAATATTACCAAAACATGTTAGGGTTCGAAACGGGAAATGAACATGCCTTCTTAAACCTCATCGATGAGTTGCTAGTCATCGACAAGAGACACAG : 1866
H27_India-VII_refseq   : CAGCAATGGAAATATTACCAAAACATGTTAGGGTTCGAAACGGGAAATGAACATGCCTTCTTAAACCTCATCGATGAGTTGCTAGTCATCGACAAGAGACACAG : 1749
H28_Ctg_refsequence    : CAGCAATGGAAATATTACCAAAACATGTTAGGGTTCGAAACGGGAAATGAACATGCCTTCTTAAACCTCATCGATGAGTTGCTAGTCATCGACAAGAGACACAG : 1821
H29_Salvador-I_refse   : CAGCAATGGAAATATTACCAAAACATGTTAGGGTTCGAAACGGGAAATGAACATGCCTTCTTAAACCTCATCGATGAGTTGCTAGTCATCGACAAGAGACACAG : 1821
H30_NKorea_refsequen   : CAGCAATGGAAATATTACCAAAACATGTTAGGGTTCGAAACGGGAAATGAACATGCCTTCTTAAACCTCATCGATGAGTTGCTAGTCATCGACAAGAGACACAG : 1731
H31_P.vivax01_refseq   : CAGCAATGGAAATATTACCAAAACATGTTAGGGTTCGAAACGGGAAATGAACATGCCTTCTTAAACCTCATCGATGAGTTGCTAGTCATCGACAAGAGACACAG : 1812
H32_Brazil-I_refsequ   : CAGCAATGGAAATATTACCAAAACATGTTAGGGTTCGAAACGGGAAATGAACATGCCTTCTTAAACCTCATCGATGAGTTGCTAGTCATCGACAAGAGACACAG : 1740
                         CAGCAATGGAAATATTACCAAAACATGTTAGGGTTCGAAACGGGAAATGAACATGCCTTCTTAAACCTCATCGATGAGTTGCTAGTCATCGACAAGAGACACAG

                                  *      2100        *      2120        *      2140        *      2160        *      2180
H1_VCG                 : TAACAATGAAGCGTATTCAAAAGTGATAAGGAAAATAAAAAAATCGAAAGCCTTTAATTATTGTACTAAAGTTATGAGGATTGCAGGGAATATTTCTTCCATTC : 1844
H2                     : TAACAATGAAGCGTATTCAAAAGTGATAAGGAAAATAAAAAAATCGAAAGCCTTTAATTATTGTACTAAAGTTATGAGGATTGCAGGGAATATTTCTTCCATTC : 1880
H3                     : TAACAATGAAGCGTATTCAAAAGTGATAAGGAAAATAAAAAAATCGAAAGCCTTTAATTATTGTACTAAAGTTATGAGGATTGCAGGGAATATTTCTTCCATTC : 1952
H4                     : TAACAATGAAGCGTATTCAAAAGTGATAAGGAAAATAAAAAAATCGAAAGCCTTTAATTATTGTACTAAAGTTATGAGGATTGCAGGGAATATTTCTTCCATTC : 1916
H5                     : TAACAATGAAGCGTATTCAAAAGTGATAAGGAAAATAAAAAAATCGAAAGCCTTTAATTATTGTACTAAAGTTATGAGGATTGCAGGGAATATTTCTTCCATTC : 1916
H6                     : TAACAATGAAGCGTATTCAAAAGTGATAAGGAAAATAAAAAAATCGAAAGCCTTTAATTATTGTACTAAAGTTATGAGGATTGCAGGGAATATTTCTTCCATTC : 1835
H7                     : TAACAATGAAGCGTATTCAAAAGTGATAAGGAAAATAAAAAAATCGAAAGCCTTTAATTATTGTACTAAAGTTATGAGGATTGCAGGGAATATTTCTTCCATTC : 1853
H8                     : TAACAATGAAGCGTATTCAAAAGTGATAAGGAAAATAAAAAAATCGAAAGCCTTTAATTATTGTACTAAAGTTATGAGGATTGCAGGGAATATTTCTTCCATTC : 1889
H9                     : TAACAATGAAGCGTATTCAAAAGTGATAAGGAAAATAAAAAAATCGAAAGCCTTTAATTATTGTACTAAAGTTATGAGGATTGCAGGGAATATTTCTTCCATTC : 1971
```

```
H10                  : TAACAATGAAGCGTATTCAAAAGTGATAAGGAAAATAAAAAAATCGAAAGCCTTTAATTATTGTACTAAAGTTATGAGGATTGCAGGGAATATTTCTTCCATTC : 1889
H11                  : TAACAATGAAGCGTATTCAAAAGTGATAAGGAAAATAAAAAAATCGAAAGCCTTTAATTATTGTACTAAAGTTATGAGGATTGCAGGGAATATTTCTTCCATTC : 1898
H12                  : TAACAATGAAGCGTATTCAAAAGTGATAAGGAAAATAAAAAAATCGAAAGCCTTTAATTATTGTACTAAAGTTATGAGGATTGCAGGGAATATTTCTTCCATTC : 1880
H13                  : TAACAATGAAGCGTATTCAAAAGTGATAAGGAAAATAAAAAAATCGAAAGCCTTTAATTATTGTACTAAAGTTATGAGGATTGCAGGGAATATTTCTTCCATTC : 1952
H14                  : TAACAATGAAGCGTATTCAAAAGTGATAAGGAAAATAAAAAAATCGAAAGCCTTTAATTATTGTACTAAAGTTATGAGGATTGCAGGGAATATTTCTTCCATTC : 1898
H15                  : TAACAATGAAGCGTATTCAAAAGTGATAAGGAAAATAAAAAAATCGAAAGCCTTTAATTATTGTACTAAAGTTATGAGGATTGCAGGGAATATTTCTTCCATTC : 1880
H16                  : TAACAATGAAGCGTATTCAAAAGTGATAAGGAAAATAAAAAAATCGAAAGCCTTTAATTATTGTACTAAAGTTATGAGGATTGCAGGGAATATTTCTTCCATTC : 1844
H17                  : TAACAATGAAGCGTATTCAAAAGTGATAAGGAAAATAAAAAAATCGAAAGCCTTTAATTATTGTACTAAAGTTATGAGGATTGCAGGGAATATTTCTTCCATTC : 2060
H18                  : TAACAATGAAGCGTATTCAAAAGTGATAAGGAAAATAAAAAAATCGAAAGCCTTTAATTATTGTACTAAAGTTATGAGGATTGCAGGGAATATTTCTTCCATTC : 1844
H19                  : TAACAATGAAGCGTATTCAAAAGTGATAAGGAAAATAAAAAAATCGAAAGCCTTTAATTATTGTACTAAAGTTATGAGGATTGCAGGGAATATTTCTTCCATTC : 1898
H20                  : TAACAATGAAGCGTATTCAAAAGTGATAAGGAAAATAAAAAAATCGAAAGCCTTTAATTATTGTACTAAAGTTATGAGGATTGCAGGGAATATTTCTTCCATTC : 1880
H21                  : TAACAATGAAGCGTATTCAAAAGTGATAAGGAAAATAAAAAAATCGAAAGCCTTTAATTATTGTACTAAAGTTATGAGGATTGCAGGGAATATTTCTTCCATTC : 1916
H22                  : TAACAATGAAGCGTATTCAAAAGTGATAAGGAAAATAAAAAAATCGAAAGCCTTTAATTATTGTACTAAAGTTATGAGGATTGCAGGGAATATTTCTTCCATTC : 1880
H23                  : TAACAATGAAGCGTATTCAAAAGTGATAAGGAAAATAAAAAAATCGAAAGCCTTTAATTATTGTACTAAAGTTATGAGGATTGCAGGGAATATTTCTTCCATTC : 1916
H24                  : TAACAATGAAGCGTATTCAAAAGTGATAAGGAAAATAAAAAAATCGAAAGCCTTTAATTATTGTACTAAAGTTATGAGGATTGCAGGGAATATTTCTTCCATTC : 1925
H25                  : TAACAATGAAGCGTATTCAAAAGTGATAAGGAAAATAAAAAAATCGAAAGCCTTTAATTATTGTACTAAAGTTATGAGGATTGCAGGGAATATTTCTTCCATTC : 1880
H26_Mauritania-I_ref : TAACAATGAAGCGTATTCAAAAGTGATAAGGAAAATAAAAAAATCGAAAGCCTTTAATTATTGTACTAAAGTTATGAGGATTGCAGGGAATATTTCTTCCATTC : 1970
H27_India-VII_refseq : TAACAATGAAGCGTATTCAAAAGTGATAAGGAAAATAAAAAAATCGAAAGCCTTTAATTATTGTACTAAAGTTATGAGGATTGCAGGGAATATTTCTTCCATTC : 1853
H28_Ctg_refsequence  : TAACAATGAAGCGTATTCAAAAGTGATAAGGAAAATAAAAAAATCGAAAGCCTTTAATTATTGTACTAAAGTTATGAGGATTGCAGGGAATATTTCTTCCATTC : 1925
H29_Salvador-I_refse : TAACAATGAAGCGTATTCAAAAGTGATAAGGAAAATAAAAAAATCGAAAGCCTTTAATTATTGTACTAAAGTTATGAGGATTGCAGGGAATATTTCTTCCATTC : 1925
H30_NKorea_refsequen : TAACAATGAAGCGTATTCAAAAGTGATAAGGAAAATAAAAAAATCGAAAGCCTTTAATTATTGTACTAAAGTTATGAGGATTGCAGGGAATATTTCTTCCATTC : 1835
H31_P.vivax01_refseq : TAACAATGAAGCGTATTCAAAAGTGATAAGGAAAATAAAAAAATCGAAAGCCTTTAATTATTGTACTAAAGTTATGAGGATTGCAGGGAATATTTCTTCCATTC : 1916
H32_Brazil-I_refsequ : TAACAATGAAGCGTATTCAAAAGTGATAAGGAAAATAAAAAAATCGAAAGCCTTTAATTATTGTACTAAAGTTATGAGGATTGCAGGGAATATTTCTTCCATTC : 1844
                       TAACAATGAAGCGTATTCAAAAGTGATAAGGAAAATAAAAAAATCGAAAGCCTTTAATTATTGTACTAAAGTTATGAGGATTGCAGGGAATATTTCTTCCATTC

                            *      2200         *      2220         *      2240         *      2260         *      2280
H1_VCG               : CATTTAATCATGAAAATAATAAAACGCCTAGCTACTCCATCATCGGGTCGTTAGGTAACCTGGTGAAGGCTCACATGGGGAACTACTACGTTGCCATTGCTAAT : 1948
H2                   : CATTTAATCATGAAAATAATAAAACGCCTAGCTACTCCATCATCGGGTCGTTAGGTAACCTGGTGAAGGCTCACATGGGGAACTACTACGTTGCCATTGCTAAT : 1984
H3                   : CATTTAATCATGAAAATAATAAAACGCCTAGCTACTCCATCATCGGGTCGTTAGGTAACCTGGTGAAGGCTCACATGGGGAACTACTACGTTGCCATTGCTAAT : 2056
H4                   : CATTTAATCATGAAAATAATAAAACGCCTAGCTACTCCATCATCGGGTCGTTAGGTAACCTGGTGAAGGCTCACATGGGGAACTACTACGTTGCCATTGCTAAT : 2020
H5                   : CATTTAATCATGAAAATAATAAAACGCCTAGCTACTCCATCATCGGGTCGTTAGGTAACCTGGTGAAGGCTCACATGGGGAACTACTACGTTGCCATTGCTAAT : 2020
H6                   : CATTTAATCATGAAAATAATAAAACGCCTAGCTACTCCATCATCGGGTCGTTAGGTAACCTGGTGAAGGCTCACATGGGGAACTACTACGTTGCCATTGCTAAT : 1939
H7                   : CATTTAATCATGAAAATAATAAAACGCCTAGCTACTCCATCATCGGGTCGTTAGGTAACCTGGTGAAGGCTCACATGGGGAACTACTACGTTGCCATTGCTAAT : 1957
H8                   : CATTTAATCATGAAAATAATAAAACGCCTAGCTACTCCATCATCGGGTCGTTAGGTAACCTGGTGAAGGCTCACATGGGGAACTACTACGTTGCCATTGCTAAT : 1993
H9                   : CATTTAATCATGAAAATAATAAAACGCCTAGCTACTCCATCATCGGGTCGTTAGGTAACCTGGTGAAGGCTCACATGGGGAACTACTACGTTGCCATTGCTAAT : 2075
H10                  : CATTTAATCATGAAAATAATAAAACGCCTAGCTACTCCATCATCGGGTCGTTAGGTAACCTGGTGAAGGCTCACATGGGGAACTACTACGTTGCCATTGCTAAT : 1993
H11                  : CATTTAATCATGAAAATAATAAAACGCCTAGCTACTCCATCATCGGGTCGTTAGGTAACCTGGTGAAGGCTCACATGGGGAACTACTACGTTGCCATTGCTAAT : 2002
H12                  : CATTTAATCATGAAAATAATAAAACGCCTAGCTACTCCATCATCGGGTCGTTAGGTAACCTGGTGAAGGCTCACATGGGGAACTACTACGTTGCCATTGCTAAT : 1984
H13                  : CATTTAATCATGAAAATAATAAAACGCCTAGCTACTCCATCATCGGGTCGTTAGGTAACCTGGTGAAGGCTCACATGGGGAACTACTACGTTGCCATTGCTAAT : 2056
H14                  : CATTTAATCATGAAAATAATAAAACGCCTAGCTACTCCATCATCGGGTCGTTAGGTAACCTGGTGAAGGCTCACATGGGGAACTACTACGTTGCCATTGCTAAT : 2002
H15                  : CATTTAATCATGAAAATAATAAAACGCCTAGCTACTCCATCATCGGGTCGTTAGGTAACCTGGTGAAGGCTCACATGGGGAACTACTACGTTGCCATTGCTAAT : 1984
H16                  : CATTTAATCATGAAAATAATAAAACGCCTAGCTACTCCATCATCGGGTCGTTAGGTAACCTGGTGAAGGCTCACATGGGGAACTACTACGTTGCCATTGCTAAT : 1948
H17                  : CATTTAATCATGAAAATAATAAAACGCCTAGCTACTCCATCATCGGGTCGTTAGGTAACCTGGTGAAGGCTCACATGGGGAACTACTACGTTGCCATTGCTAAT : 2164
H18                  : CATTTAATCATGAAAATAATAAAACGCCTAGCTACTCCATCATCGGGTCGTTAGGTAACCTGGTGAAGGCTCACATGGGGAACTACTACGTTGCCATTGCTAAT : 1948
H19                  : CATTTAATCATGAAAATAATAAAACGCCTAGCTACTCCATCATCGGGTCGTTAGGTAACCTGGTGAAGGCTCACATGGGGAACTACTACGTTGCCATTGCTAAT : 2002
```

```
H20                  : CATTTAATCATGAAAATAATAAAACGCCTAGCTACTCCATCATCGGGTCGTTAGGTAACCTGGTGAAGGCTCACATGGGGAACTACTACGTTGCCATTGCTAAT : 1984
H21                  : CATTTAATCATGAAAATAATAAAACGCCTAGCTACTCCATCATCGGGTCGTTAGGTAACCTGGTGAAGGCTCACATGGGGAACTACTACGTTGCCATTGCTAAT : 2020
H22                  : CATTTAATCATGAAAATAATAAAACGCCTAGCTACTCCATCATCGGGTCGTTAGGTAACCTGGTGAAGGCTCACATGGGGAACTACTACGTTGCCATTGCTAAT : 1984
H23                  : CATTTAATCATGAAAATAATAAAACGCCTAGCTACTCCATCATCGGGTCGTTAGGTAACCTGGTGAAGGCTCACATGGGGAACTACTACGTTGCCATTGCTAAT : 2020
H24                  : CATTTAATCATGAAAATAATAAAACGCCTAGCTACTCCATCATCGGGTCGTTAGGTAACCTGGTGAAGGCTCACATGGGGAACTACTACGTTGCCATTGCTAAT : 2029
H25                  : CATTTAATCATGAAAATAATAAAACGCCTAGCTACTCCATCATCGGGTCGTTAGGTAACCTGGTGAAGGCTCACATGGGGAACTACTACGTTGCCATTGCTAAT : 1984
H26_Mauritania-I_ref : CATTTAATCATGAAAATAATAAAACGCCTAGCTACTCCATCATCGGGTCGTTAGGTAACCTGGTGAAGGCTCACATGGGGAACTACTACGTTGCCATTGCTAAT : 2074
H27_India-VII_refseq : CATTTAATCATGAAAATAATAAAACGCCTAGCTACTCCATCATCGGGTCGTTAGGTAACCTGGTGAAGGCTCACATGGGGAACTACTACGTTGCCATTGCTAAT : 1957
H28_Ctg_refsequence  : CATTTAATCATGAAAATAATAAAACGCCTAGCTACTCCATCATCGGGTCGTTAGGTAACCTGGTGAAGGCTCACATGGGGAACTACTACGTTGCCATTGCTAAT : 2029
H29_Salvador-I_refse : CATTTAATCATGAAAATAATAAAACGCCTAGCTACTCCATCATCGGGTCGTTAGGTAACCTGGTGAAGGCTCACATGGGGAACTACTACGTTGCCATTGCTAAT : 2029
H30_NKorea_refsequen : CATTTAATCATGAAAATAATAAAACGCCTAGCTACTCCATCATCGGGTCGTTAGGTAACCTGGTGAAGGCTCACATGGGGAACTACTACGTTGCCATTGCTAAT : 1939
H31_P.vivax01_refseq : CATTTAATCATGAAAATAATAAAACGCCTAGCTACTCCATCATCGGGTCGTTAGGTAACCTGGTGAAGGCTCACATGGGGAACTACTACGTTGCCATTGCTAAT : 2020
H32_Brazil-I_refsequ : CATTTAATCATGAAAATAATAAAACGCCTAGCTACTCCATCATCGGGTCGTTAGGTAACCTGGTGAAGGCTCACATGGGGAACTACTACGTTGCCATTGCTAAT : 1948
                       CATTTAATCATGAAAATAATAAAACGCCTAGCTACTCCATCATCGGGTCGTTAGGTAACCTGGTGAAGGCTCACATGGGGAACTACTACGTTGCCATTGCTAAT

                        *      2300         *      2320         *      2340         *      2360         *      2380         *
H1_VCG               : AGGATTAACTCCTATTTTGCCTATGCCGAGAAGAGGAACAAAAAAAACAGCCCCTTAAAAGTTGTCTCCGTCTGTACTCTCCTTCATTTGACCGACATGCTCCA : 2052
H2                   : AGGATTAACTCCTATTTTGCCTATGCCGAGAAGAGGAACAAAAAAAACAGCCCCTTAAAAGTTGTCTCCGTCTGTACTCTCCTTCATTTGACCGACATGCTCCA : 2088
H3                   : AGGATTAACTCCTATTTTGCCTATGCCGAGAAGAGGAACAAAAAAAACAGCCCCTTAAAAGTTGTCTCCGTCTGTACTCTCCTTCATTTGACCGACATGCTCCA : 2160
H4                   : AGGATTAACTCCTATTTTGCCTATGCCGAGAAGAGGAACAAAAAAAACAGCCCCTTAAAAGTTGTCTCCGTCTGTACTCTCCTTCATTTGACCGACATGCTCCA : 2124
H5                   : AGGATTAACTCCTATTTTGCCTATGCCGAGAAGAGGAACAAAAAAAACAGCCCCTTAAAAGTTGTCTCCGTCTGTACTCTCCTTCATTTGACCGACATGCTCCA : 2124
H6                   : AGGATTAACTCCTATTTTGCCTATGCCGAGAAGAGGAACAAAAAAAACAGCCCCTTAAAAGTTGTCTCCGTCTGTACTCTCCTTCATTTGACCGACATGCTCCA : 2043
H7                   : AGGATTAACTCCTATTTTGCCTATGCCGAGAAGAGGAACAAAAAAAACAGCCCCTTAAAAGTTGTCTCCGTCTGTACTCTCCTTCATTTGACCGACATGCTCCA : 2061
H8                   : AGGATTAACTCCTATTTTGCCTATGCCGAGAAGAGGAACAAAAAAAACAGCCCCTTAAAAGTTGTCTCCGTCTGTACTCTCCTTCATTTGACCGACATGCTCCA : 2097
H9                   : AGGATTAACTCCTATTTTGCCTATGCCGAGAAGAGGAACAAAAAAAACAGCCCCTTAAAAGTTGTCTCCGTCTGTACTCTCCTTCATTTGACCGACATGCTCCA : 2179
H10                  : AGGATTAACTCCTATTTTGCCTATGCCGAGAAGAGGAACAAAAAAAACAGCCCCTTAAAAGTTGTCTCCGTCTGTACTCTCCTTCATTTGACCGACATGCTCCA : 2097
H11                  : AGGATTAACTCCTATTTTGCCTATGCCGAGAAGAGGAACAAAAAAAACAGCCCCTTAAAAGTTGTCTCCGTCTGTACTCTCCTTCATTTGACCGACATGCTCCA : 2106
H12                  : AGGATTAACTCCTATTTTGCCTATGCCGAGAAGAGGAACAAAAAAAACAGCCCCTTAAAAGTTGTCTCCGTCTGTACTCTCCTTCATTTGACCGACATGCTCCA : 2088
H13                  : AGGATTAACTCCTATTTTGCCTATGCCGAGAAGAGGAACAAAAAAAACAGCCCCTTAAAAGTTGTCTCCGTCTGTACTCTCCTTCATTTGACCGACATGCTCCA : 2160
H14                  : AGGATTAACTCCTATTTTGCCTATGCCGAGAAGAGGAACAAAAAAAACAGCCCCTTAAAAGTTGTCTCCGTCTGTACTCTCCTTCATTTGACCGACATGCTCCA : 2106
H15                  : AGGATTAACTCCTATTTTGCCTATGCCGAGAAGAGGAACAAAAAAAACAGCCCCTTAAAAGTTGTCTCCGTCTGTACTCTCCTTCATTTGACCGACATGCTCCA : 2088
H16                  : AGGATTAACTCCTATTTTGCCTATGCCGAGAAGAGGAACAAAAAAAACAGCCCCTTAAAAGTTGTCTCCGTCTGTACTCTCCTTCATTTGACCGACATGCTCCA : 2052
H17                  : AGGATTAACTCCTATTTTGCCTATGCCGAGAAGAGGAACAAAAAAAACAGCCCCTTAAAAGTTGTCTCCGTCTGTACTCTCCTTCATTTGACCGACATGCTCCA : 2268
H18                  : AGGATTAACTCCTATTTTGCCTATGCCGAGAAGAGGAACAAAAAAAACAGCCCCTTAAAAGTTGTCTCCGTCTGTACTCTCCTTCATTTGACCGACATGCTCCA : 2052
H19                  : AGGATTAACTCCTATTTTGCCTATGCCGAGAAGAGGAACAAAAAAAACAGCCCCTTAAAAGTTGTCTCCGTCTGTACTCTCCTTCATTTGACCGACATGCTCCA : 2106
H20                  : AGGATTAACTCCTATTTTGCCTATGCCGAGAAGAGGAACAAAAAAAACAGCCCCTTAAAAGTTGTCTCCGTCTGTACTCTCCTTCATTTGACCGACATGCTCCA : 2088
H21                  : AGGATTAACTCCTATTTTGCCTATGCCGAGAAGAGGAACAAAAAAAACAGCCCCTTAAAAGTTGTCTCCGTCTGTACTCTCCTTCATTTGACCGACATGCTCCA : 2124
H22                  : AGGATTAACTCCTATTTTGCCTATGCCGAGAAGAGGAACAAAAAAAACAGCCCCTTAAAAGTTGTCTCCGTCTGTACTCTCCTTCATTTGACCGACATGCTCCA : 2088
H23                  : AGGATTAACTCCTATTTTGCCTATGCCGAGAAGAGGAACAAAAAAAACAGCCCCTTAAAAGTTGTCTCCGTCTGTACTCTCCTTCATTTGACCGACATGCTCCA : 2124
H24                  : AGGATTAACTCCTATTTTGCCTATGCCGAGAAGAGGAACAAAAAAAACAGCCCCTTAAAAGTTGTCTCCGTCTGTACTCTCCTTCATTTGACCGACATGCTCCA : 2133
H25                  : AGGATTAACTCCTATTTTGCCTATGCCGAGAAGAGGAACAAAAAAAACAGCCCCTTAAAAGTTGTCTCCGTCTGTACTCTCCTTCATTTGACCGACATGCTCCA : 2088
H26_Mauritania-I_ref : AGGATTAACTCCTATTTTGCCTATGCCGAGAAGAGGAACAAAAAAAACAGCCCCTTAAAAGTTGTCTCCGTCTGTACTCTCCTTCATTTGACCGACATGCTCCA : 2178
H27_India-VII_refseq : AGGATTAACTCCTATTTTGCCTATGCCGAGAAGAGGAACAAAAAAAACAGCCCCTTAAAAGTTGTCTCCGTCTGTACTCTCCTTCATTTGACCGACATGCTCCA : 2061
H28_Ctg_refsequence  : AGGATTAACTCCTATTTTGCCTATGCCGAGAAGAGGAACAAAAAAAACAGCCCCTTAAAAGTTGTCTCCGTCTGTACTCTCCTTCATTTGACCGACATGCTCCA : 2133
H29_Salvador-I_refse : AGGATTAACTCCTATTTTGCCTATGCCGAGAAGAGGAACAAAAAAAACAGCCCCTTAAAAGTTGTCTCCGTCTGTACTCTCCTTCATTTGACCGACATGCTCCA : 2133
```

```
H30_NKorea_refsequen : AGGATTAACTCCTATTTTGCCTATGCCGAGAAGAGGAACAAAAAAAACAGCCCCTTAAAAGTTGTCTCCGTCTGTACTCTCCTTCATTTGACCGACATGCTCCA : 2043
H31_P.vivax01_refseq : AGGATTAACTCCTATTTTGCCTATGCCGAGAAGAGGAACAAAAAAAACAGCCCCTTAAAAGTTGTCTCCGTCTGTACTCTCCTTCATTTGACCGACATGCTCCA : 2124
H32_Brazil-I_refsequ : AGGATTAACTCCTATTTTGCCTATGCCGAGAAGAGGAACAAAAAAAACAGCCCCTTAAAAGTTGTCTCCGTCTGTACTCTCCTTCATTTGACCGACATGCTCCA : 2052
                       AGGATTAACTCCTATTTTGCCTATGCCGAGAAGAGGAACAAAAAAAACAGCCCCTTAAAAGTTGTCTCCGTCTGTACTCTCCTTCATTTGACCGACATGCTCCA

                           2400         *      2420         *      2440         *      2460         *      2480         *
H1_VCG               : CAATTGCTCTGATGAACATTTAAAGAATATTTTGGACCTCAACACTTTGAAGCTTAACATTCTGAACATGCAGGGCAGGCGCGTGCTGCAGCCCCTCGTCAAGA : 2156
H2                   : CAATTGCTCTGATGAACATTTAAAGAATATTTTGGACCTCAACACTTTGAAGCTTAACATTCTGAACATGCAGGGCAGGCGCGTGCTGCAGCCCCTCGTCAAGA : 2192
H3                   : CAATTGCTCTGATGAACATTTAAAGAATATTTTGGACCTCAACACTTTGAAGCTTAACATTCTGAACATGCAGGGCAGGCGCGTGCTGCAGCCCCTCGTCAAGA : 2264
H4                   : CAATTGCTCTGATGAACATTTAAAGAATATTTTGGACCTCAACACTTTGAAGCTTAACATTCTGAACATGCAGGGCAGGCGCGTGCTGCAGCCCCTCGTCAAGA : 2228
H5                   : CAATTGCTCTGATGAACATTTAAAGAATATTTTGGACCTCAACACTTTGAAGCTTAACATTCTGAACATGCAGGGCAGGCGCGTGCTGCAGCCCCTCGTCAAGA : 2228
H6                   : CAATTGCTCTGATGAACATTTAAAGAATATTTTGGACCTCAACACTTTGAAGCTTAACATTCTGAACATGCAGGGCAGGCGCGTGCTGCAGCCCCTCGTCAAGA : 2147
H7                   : CAATTGCTCTGATGAACATTTAAAGAATATTTTGGACCTCAACACTTTGAAGCTTAACATTCTGAACATGCAGGGCAGGCGCGTGCTGCAGCCCCTCGTCAAGA : 2165
H8                   : CAATTGCTCTGATGAACATTTAAAGAATATTTTGGACCTCAACACTTTGAAGCTTAACATTCTGAACATGCAGGGCAGGCGCGTGCTGCAGCCCCTCGTCAAGA : 2201
H9                   : CAATTGCTCTGATGAACATTTAAAGAATATTTTGGACCTCAACACTTTGAAGCTTAACATTCTGAACATGCAGGGCAGGCGCGTGCTGCAGCCCCTCGTCAAGA : 2283
H10                  : CAATTGCTCTGATGAACATTTAAAGAATATTTTGGACCTCAACACTTTGAAGCTTAACATTCTGAACATGCAGGGCAGGCGCGTGCTGCAGCCCCTCGTCAAGA : 2201
H11                  : CAATTGCTCTGATGAACATTTAAAGAATATTTTGGACCTCAACACTTTGAAGCTTAACATTCTGAACATGCAGGGCAGGCGCGTGCTGCAGCCCCTCGTCAAGA : 2210
H12                  : CAATTGCTCTGATGAACATTTAAAGAATATTTTGGACCTCAACACTTTGAAGCTTAACATTCTGAACATGCAGGGCAGGCGCGTGCTGCAGCCCCTCGTCAAGA : 2192
H13                  : CAATTGCTCTGATGAACATTTAAAGAATATTTTGGACCTCAACACTTTGAAGCTTAACATTCTGAACATGCAGGGCAGGCGCGTGCTGCAGCCCCTCGTCAAGA : 2264
H14                  : CAATTGCTCTGATGAACATTTAAAGAATATTTTGGACCTCAACACTTTGAAGCTTAACATTCTGAACATGCAGGGCAGGCGCGTGCTGCAGCCCCTCGTCAAGA : 2210
H15                  : CAATTGCTCTGATGAACATTTAAAGAATATTTTGGACCTCAACACTTTGAAGCTTAACATTCTGAACATGCAGGGCAGGCGCGTGCTGCAGCCCCTCGTCAAGA : 2192
H16                  : CAATTGCTCTGATGAACATTTAAAGAATATTTTGGACCTCAACACTTTGAAGCTTAACATTCTGAACATGCAGGGCAGGCGCGTGCTGCAGCCCCTCGTCAAGA : 2156
H17                  : CAATTGCTCTGATGAACATTTAAAGAATATTTTGGACCTCAACACTTTGAAGCTTAACATTCTGAACATGCAGGGCAGGCGCGTGCTGCAGCCCCTCGTCAAGA : 2372
H18                  : CAATTGCTCTGATGAACATTTAAAGAATATTTTGGACCTCAACACTTTGAAGCTTAACATTCTGAACATGCAGGGCAGGCGCGTGCTGCAGCCCCTCGTCAAGA : 2156
H19                  : CAATTGCTCTGATGAACATTTAAAGAATATTTTGGACCTCAACACTTTGAAGCTTAACATTCTGAACATGCAGGGCAGGCGCGTGCTGCAGCCCCTCGTCAAGA : 2210
H20                  : CAATTGCTCTGATGAACATTTAAAGAATATTTTGGACCTCAACACTTTGAAGCTTAACATTCTGAACATGCAGGGCAGGCGCGTGCTGCAGCCCCTCGTCAAGA : 2192
H21                  : CAATTGCTCTGATGAACATTTAAAGAATATTTTGGACCTCAACACTTTGAAGCTTAACATTCTGAACATGCAGGGCAGGCGCGTGCTGCAGCCCCTCGTCAAGA : 2228
H22                  : CAATTGCTCTGATGAACATTTAAAGAATATTTTGGACCTCAACACTTTGAAGCTTAACATTCTGAACATGCAGGGCAGGCGCGTGCTGCAGCCCCTCGTCAAGA : 2192
H23                  : CAATTGCTCTGATGAACATTTAAAGAATATTTTGGACCTCAACACTTTGAAGCTTAACATTATGAACATGCAGGGCAGGCGCGTGCTGCAGCCCCTCGTCAAGA : 2228
H24                  : CAATTGCTCTGATGAACATTTAAAGAATATTTTGGACCTCAACACTTTGAAGCTTAACATTCTGAACATGCAGGGCAGGCGCGTGCTGCAGCCCCTCGTCAAGA : 2237
H25                  : CAATTGCTCTGATGAACATTTAAAGAATATTTTGGACCTCAACACTTTGAAGCTTAACATTCTGAACATGCAGGGCAGGCGCGTGCTGCAGCCCCTCGTCAAGA : 2192
H26_Mauritania-I_ref : CAATTGCTCTGATGAACATTTAAAGAATATTTTGGACCTCAACACTTTGAAGCTTAACATTCTGAACATGCAGGGCAGGCGCGTGCTGCAGCCCCTCGTCAAGA : 2282
H27_India-VII_refseq : CAATTGCTCTGATGAACATTTAAAGAATATTTTGGACCTCAACACTTTGAAGCTTAACATTCTGAACATGCAGGGCAGGCGCGTGCTGCAGCCCCTCGTCAAGA : 2165
H28_Ctg_refsequence  : CAATTGCTCTGATGAACATTTAAAGAATATTTTGGACCTCAACACTTTGAAGCTTAACATTCTGAACATGCAGGGCAGGCGCGTGCTGCAGCCCCTCGTCAAGA : 2237
H29_Salvador-I_refse : CAATTGCTCTGATGAACATTTAAAGAATATTTTGGACCTCAACACTTTGAAGCTTAACATTCTGAACATGCAGGGCAGGCGCGTGCTGCAGCCCCTCGTCAAGA : 2237
H30_NKorea_refsequen : CAATTGCTCTGATGAACATTTAAAGAATATTTTGGACCTCAACACTTTGAAGCTTAACATTCTGAACATGCAGGGCAGGCGCGTGCTGCAGCCCCTCGTCAAGA : 2147
H31_P.vivax01_refseq : CAATTGCTCTGATGAACATTTAAAGAATATTTTGGACCTCAACACTTTGAAGCTTAACATTCTGAACATGCAGGGCAGGCGCGTGCTGCAGCCCCTCGTCAAGA : 2228
H32_Brazil-I_refsequ : CAATTGCTCTGATGAACATTTAAAGAATATTTTGGACCTCAACACTTTGAAGCTTAACATTCTGAACATGCAGGGCAGGCGCGTGCTGCAGCCCCTCGTCAAGA : 2156
                       CAATTGCTCTGATGAACATTTAAAGAATATTTTGGACCTCAACACTTTGAAGCTTAACATTcTGAACATGCAGGGCAGGCGCGTGCTGCAGCCCCTCGTCAAGA

                       2500         *      2520         *      2540         *      2560         *      2580         *      2600
H1_VCG               : TGAGCTTCCTCGGCGCCGCGCAGAACCCCGCCCTCAAGGAGATCTGCGAGCCGAGCAACCACCTAGTGGGTGAGTGACGGGGGGGCGACCTCCATGGGGGCT : 2260
H2                   : TGAGCTTCCTCGGCGCCGCGCAGAACCCCGCCCTCAAGGAGATCTGCGAGCCGAGCAACCACCTAGTGGGTGAGTGACGGGGGGGCGACCTCCATGGGGGCT : 2296
H3                   : TGAGCTTCCTCGGCGCCGCGCAGAACCCCGCCCTCAAGGAGATCTGCGAGCCGAGCAACCACCTAGTGGGTGAGTGACGGGGGGGCGACCTCCATGGGGGCT : 2368
H4                   : TGAGCTTCCTCGGCGCCGCGCAGAACCCCGCCCTCAAGGAGATCTGCGAGCCGAGCAACCACCTAGTGGGTGAGTGACGGGGGGGCGACCTCCATGGGGGCT : 2332
```

```
H5                    : TGAGCTTCCTCGGCGCCGCGCAGAACCCCGCCCTCAAGGAGATCTGCGAGCCGAGCAACCACCTAGTGGGTGAGTGACGGGGGGGGCGACCTCCATGGGGGGCT : 2332
H6                    : TGAGCTTCCTCGGCGCCGCGCAGAACCCCGCCCTCAAGGAGATCTGCGAGCCGAGCAACCACCTAGTGGGTGAGTGACGGGGGGGGCGACCTCCATGGGGGGCT : 2251
H7                    : TGAGCTTCCTCGGCGCCGCGCAGAACCCCGCCCTCAAGGAGATCTGCGAGCCGAGCAACCACCTAGTGGGTGAGTGACGGGGGGGGCGACCTCCATGGGGGGCT : 2269
H8                    : TGAGCTTCCTCGGCGCCGCGCAGAACCCCGCCCTCAAGGAGATCTGCGAGCCGAGCAACCACCTAGTGGGTGAGTGACGGGGGGGGCGACCTCCATGGGGGGCT : 2305
H9                    : TGAGCTTCCTCGGCGCCGCGCAGAACCCCGCCCTCAAGGAGATCTGCGAGCCGAGCAACCACCTAGTGGGTGAGTGACGGGGGGGGCGACCTCCATGGGGGGCT : 2387
H10                   : TGAGCTTCCTCGGCGCCGCGCAGAACCCCGCCCTCAAGGAGATCTGCGAGCCGAGCAACCACCTAGTGGGTGAGTGACGGGGGGGGCGACCTCCATGGGGGGCT : 2305
H11                   : TGAGCTTCCTCGGCGCCGCGCAGAACCCCGCCCTCAAGGAGATCTGCGAGCCGAGCAACCACCTAGTGGGTGAGTGACGGGGGGGGCGACCTCCATGGGGGGCT : 2314
H12                   : TGAGCTTCCTCGGCGCCGCGCAGAACCCCGCCCTCAAGGAGATCTGCGAGCCGAGCAACCACCTAGTGGGTGAGTGACGGGGGGGGCGACCTCCATGGGGGGCT : 2296
H13                   : TGAGCTTCCTCGGCGCCGCGCAGAACCCCGCCCTCAAGGAGATCTGCGAGCCGAGCAACCACCTAGTGGGTGAGTGACGGGGGGGGCGACCTCCATGGGGGGCT : 2368
H14                   : TGAGCTTCCTCGGCGCCGCGCAGAACCCCGCCCTCAAGGAGATCTGCGAGCCGAGCAACCACCTAGTGGGTGAGTGACGGGGGGGGCGACCTCCATGGGGGGCT : 2314
H15                   : TGAGCTTCCTCGGCGCCGCGCAGAACCCCGCCCTCAAGGAGATCTGCGAGCCGAGCAACCACCTAGTGGGTGAGTGACGGGGGGGGCGACCTCCATGGGGGGCT : 2296
H16                   : TGAGCTTCCTCGGCGCCGCGCAGAACCCCGCCCTCAAGGAGATCTGCGAGCCGAGCAACCACCTAGTGGGTGAGTGACGGGGGGGGCGACCTCCATGGGGGGCT : 2260
H17                   : TGAGCTTCCTCGGCGCCGCGCAGAACCCCGCCCTCAAGGAGATCTGCGAGCCGAGCAACCACCTAGTGGGTGAGTGACGGGGGGGGCGACCTCCATGGGGGGCT : 2476
H18                   : TGAGCTTCCTCGGCGCCGCGCAGAACCCCGCCCTCAAGGAGATCTGCGAGCCGAGCAACCACCTAGTGGGTGAGTGACGGGGGGGGCGACCTCCATGGGGGGCT : 2260
H19                   : TGAGCTTCCTCGGCGCCGCGCAGAACCCCGCCCTCAAGGAGATCTGCGAGCCGAGCAACCACCTAGTGGGTGAGTGACGGGGGGGGCGACCTCCATGGGGGGCT : 2314
H20                   : TGAGCTTCCTCGGCGCCGCGCAGAACCCCGCCCTCAAGGAGATCTGCGAGCCGAGCAACCACCTAGTGGGTGAGTGACGGGGGGGGCGACCTCCATGGGGGGCT : 2296
H21                   : TGAGCTTCCTCGGCGCCGCGCAGAACCCCGCCCTCAAGGAGATCTGCGAGCCGAGCAACCACCTAGTGGGTGAGTGACGGGGGGGGCGACCTCCATGGGGGGCT : 2332
H22                   : TGAGCTTCCTCGGCGCCGCGCAGAACCCCGCCCTCAAGGAGATCTGCGAGCCGAGCAACCACCTAGTGGGTGAGTGACGGGGGGGGCGACCTCCATGGGGGGCT : 2296
H23                   : TGAGCTTCCTCGGCGCCGCGCAGAACCCCGCCCTCAAGGAGATCTGCGAGCCGAGCAACCACCTAGTGGGTGAGTGACGGGGGGGGCGACCTCCATGGGGGGCT : 2332
H24                   : TGAGCTTCCTCGGCGCCGCGCAGAACCCCGCCCTCAAGGAGATCTGCGAGCCGAGCAACCACCTAGTGGGTGAGTGACGGGGGGGGCGACCTCCATGGGGGGCT : 2341
H25                   : TGAGCTTCCTCGGCGCCGCGCAGAACCCCGCCCTCAAGGAGATCTGCGAGCCGAGCAACCACCTAGTGGGTGAGTGACGGGGGGGGCGACCTCCATGGGGGGCT : 2296
H26_Mauritania-I_ref  : TGAGCTTCCTCGGCGCCGCGCAGAACCCCGCCCTCAAGGAGATCTGCGAGCCGAGCAACCACCTAGTGGGTGAGTGACGGGGGGGGCGACCTCCATGGGGGGCT : 2386
H27_India-VII_refseq  : TGAGCTTCCTCGGCGCCGCGCAGAACCCCGCCCTCAAGGAGATCTGCGAGCCGAGCAACCACCTAGTGGGTGAGTGACGGGGGGGGCGACCTCCATGGGGGGCT : 2269
H28_Ctg_refsequence   : TGAGCTTCCTCGGCGCCGCGCAGAACCCCGCCCTCAAGGAGATCTGCGAGCCGAGCAACCACCTAGTGGGTGAGTGACGGGGGGGACGACCTCCATGGGGGGCT : 2341
H29_Salvador-I_refse  : TGAGCTTCCTCGGCGCCGCGCAGAACCCCGCCCTCAAGGAGATCTGCGAGCCGAGCAACCACCTAGTGGGTGAGTGACGGGGGGGACGACCTCCATGGGGGGCT : 2341
H30_NKorea_refsequen  : TGAGCTTCCTCGGCGCCGCGCAGAACCCCGCCCTCAAGGAGATCTGCGAGCCGAGCAACCACCTAGTGGGTGAGTGACGGGGGGGGCGACCTCCATGGGGGGCT : 2251
H31_P.vivax01_refseq  : TGAGCTTCCTCGGCGCCGCGCAGAACCCCGCCCTCAAGGAGATCTGCGAGCCGAGCAACCACCTAGTGGGTGAGTGACGGGGGGGGCGACCTCCATGGGGGGCT : 2332
H32_Brazil-I_refsequ  : TGAGCTTCCTCGGCGCCGCGCAGAACCCCGCCCTCAAGGAGATCTGCGAGCCGAGCAACCACCTAGTGGGTGAGTGACGGGGGGGGCGACCTCCATGGGGGGCT : 2260
                        TGAGCTTCCTCGGCGCCGCGCAGAACCCCGCCCTCAAGGAGATCTGCGAGCCGAGCAACCACCTAGTGGGTGAGTGACGGGGGGGgCGACCTCCATGGGGGGCT

                                 *      2620         *      2640         *      2660         *      2680         *      2700
H1_VCG                : GCTATTCCAGCGCGGTGATCCACGCAGTGATCCGCGCAGCTGCTTTATTGGAAGCGCTTTACTCGAACTGCCTTCCTTCATTCGAACTTCCTTCCTTCATTCGC : 2364
H2                    : GCTATTCCAGCGCGGTGATCCACGCAGTGATCCGCGCAGCTGCTTTATTGGAAGCGCTTTACTCGAACTGCCTTCCTTCATTCGAACTTCCTTCCTTCATTCGC : 2400
H3                    : GCTATTCCAGCGCGGTGATCCACGCAGTGATCCGCGCAGCTGCTTTATTGGAAGCGCTTTACTCGAACTGCCTTCCTTCATTCGAACTTCCTTCCTTCATTCGC : 2472
H4                    : GCTATTCCAGCGCGGTGATCCACGCAGTGATCCGCGCAGCTGCTTTATTGGAAGCGCTTTACTCGAACTGCCTTCCTTCATTCGAACTTCCTTCCTTCATTCGC : 2436
H5                    : GCTATTCCAGCGCGGTGATCCACGCAGTGATCCGCGCAGCTGCTTTATTGGAAGCGCTTTACTCGAACTGCCTTCCTTCATTCGAACTTCCTTCCTTCATTCGC : 2436
H6                    : GCTATTCCAGCGCGGTGATCCACGCAGTGATCCGCGCAGCTGCTTTATTGGAAGCGCTTTACTCGAACTGCCTTCCTTCATTCGAACTTCCTTCCTTCATTCGC : 2355
H7                    : GCTATTCCAGCGCGGTGATCCACGCAGTGATCCGCGCAGCTGCTTTATTGGAAGCGCTTTACTCGAACTGCCTTCCTTCATTCGAACTTCCTTCCTTCATTCGC : 2373
H8                    : GCTATTCCAGCGCGGTGATCCACGCAGTGATCCGCGCAGCTGCTTTATTGGAAGCGCTTTACTCGAACTGCCTTCCTTCATTCGAACTTCCTTCCTTCATTCGC : 2409
H9                    : GCTATTCCAGCGCGGTGATCCACGCAGTGATCCGCGCAGCTGCTTTATTGGAAGCGCTTTACTCGAACTGCCTTCCTTCATTCGAACTTCCTTCCTTCATTCGC : 2491
H10                   : GCTATTCCAGCGCGGTGATCCACGCAGTGATCCGCGCAGCTGCTTTATTGGAAGCGCTTTACTCGAACTGCCTTCCTTCATTCGAACTTCCTTCCTTCATTCGC : 2409
H11                   : GCTATTCCAGCGCGGTGATCCACGCAGTGATCCGCGCAGCTGCTTTATTGGAAGCGCTTTACTCGAACTGCCTTCCTTCATTCGAACTTCCTTCCTTCATTCGC : 2418
H12                   : GCTATTCCAGCGCGGTGATCCACGCAGTGATCCGCGCAGCTGCTTTATTGGAAGCGCTTTACTCGAACTGCCTTCCTTCATTCGAACTTCCTTCCTTCATTCGC : 2400
H13                   : GCTATTCCAGCGCGGTGATCCACGCAGTGATCCGCGCAGCTGCTTTATTGGAAGCGCTTTACTCGAACTGCCTTCCTTCATTCGAACTTCCTTCCTTCATTCGC : 2472
H14                   : GCTATTCCAGCGCGGTGATCCACGCAGTGATCCGCGCAGCTGCTTTATTGGAAGCGCTTTACTCGAACTGCCTTCCTTCATTCGAACTTCCTTCCTTCATTCGC : 2418
```

```
H15                  : GCTATTCCAGCGCGGTGATCCACGCAGTGATCCGCGCAGCTGCTTTATTGGAAGCGCTTTACTCGAACTGCCTTCCTTCATTCGAACTTCCTTCCTTCATTCGC : 2400
H16                  : GCTATTCCAGCGCGGTGATCCACGCAGTGATCCGCGCAGCTGCTTTATTGGAAGCGCTTTACTCGAACTGCCTTCCTTCATTCGAACTTCCTTCCTTCATTCGC : 2364
H17                  : GCTATTCCAGCGCGGTGATCCACGCAGTGATCCGCGCAGCTGCTTTATTGGAAGCGCTTTACTCGAACTGCCTTCCTTCATTCGAACTTCCTTCCTTCATTCGC : 2580
H18                  : GCTATTCCAGCGCGGTGATCCACGCAGTGATCCGCGCAGCTGCTTTATTGGAAGCGCTTTACTCGAACTGCCTTCCTTCATTCGAACTTCCTTCCTTCATTCGC : 2364
H19                  : GCTATTCCAGCGCGGTGATCCACGCAGTGATCCGCGCAGCTGCTTTATTGGAAGCGCTTTACTCGAACTGCCTTCCTTCATTCGAACTTCCTTCCTTCATTCGC : 2418
H20                  : GCTATTCCAGCGCGGTGATCCACGCAGTGATCCGCGCAGCTGCTTTATTGGAAGCGCTTTACTCGAACTGCCTTCCTTCATTCGAACTTCCTTCCTTCATTCGC : 2400
H21                  : GCTATTCCAGCGCGGTGATCCACGCAGTGATCCGCGCAGCTGCTTTATTGGAAGCGCTTTACTCGAACTGCCTTCCTTCATTCGAACTTCCTTCCTTCATTCGC : 2436
H22                  : GCTATTCCAGCGCGGTGATCCACGCAGTGATCCGCGCAGCTGCTTTATTGGAAGCGCTTTACTCGAACTGCCTTCCTTCATTCGAACTTCCTTCCTTCATTCGC : 2400
H23                  : GCTATTCCAGCGCGGTGATCCACGCAGTGATCCGCGCAGCTGCTTTATTGGAAGCGCTTTACTCGAACTGCCTTCCTTCATTCGAACTTCCTTCCTTCATTCGC : 2436
H24                  : GCTATTCCAGCGCGGTGATCCACGCAGTGATCCGCGCAGCTGCTTTATTGGAAGCGCTTTACTCGAACTGCCTTCCTTCATTCGAACTTCCTTCCTTCATTCGC : 2445
H25                  : GCTATTCCAGCGCGGTGATCCACGCAGTGATCCGCGCAGCTGCTTTATTGGAAGCGCTTTACTCGAACTGCCTTCCTTCATTCGAACTTCCTTCCTTCATTCGC : 2400
H26_Mauritania-I_ref : GCTATTCCAGCGCGGTGATCCACGCAGTGATCCGCGCAGCTGCTTTATTGGAAGCGCTTTACTCGAACTGCCTTCCTTCATTCGAACTTCCTTCCTTCATTCGC : 2490
H27_India-VII_refseq : GCTATTCCAGCGCGGTGATCCACGCAGTGATCCGCGCAGCTGCTTTATTGGAAGCGCTTTACTCGAACTGCCTTCCTTCATTCGAACTTCCTTCCTTCATTCGC : 2373
H28_Ctg_refsequence  : GCTATTCCAGCGCGGTGATCCACGCAGTGATCCGCGCAGCTGCTTTATTGGAAGCGCTTTACTCGAACTGCCTTCCTTCATTCGAACTTCCTTCCTTCATTCGC : 2445
H29_Salvador-I_refse : GCTATTCCAGCGCGGTGATCCACGCAGTGATCCGCGCAGCTGCTTTATTGGAAGCGCTTTACTCGAACTGCCTTCCTTCATTCGAACTTCCTTCCTTCATTCGC : 2445
H30_NKorea_refsequen : GCTATTCCAGCGCGGTGATCCACGCAGTGATCCGCGCAGCTGCTTTATTGGAAGCGCTTTACTCGAACTGCCTTCCTTCATTCGAACTTCCTTCCTTCATTCGC : 2355
H31_P.vivax01_refseq : GCTATTCCAGCGCGGTGATCCACGCAGTGATCCGCGCAGCTGCTTTATTGGAAGCGCTTTACTCGAACTGCCTTCCTTCATTCGAACTTCCTTCCTTCATTCGC : 2436
H32_Brazil-I_refsequ : GCTATTCCAGCGCGGTGATCCACGCAGTGATCCGCGCAGCTGCTTTATTGGAAGCGCTTTACTCGAACTGCCTTCCTTCATTCGAACTTCCTTCCTTCATTCGC : 2364
                       GCTATTCCAGCGCGGTGATCCACGCAGTGATCCGCGCAGCTGCTTTATTGGAAGCGCTTTACTCGAACTGCCTTCCTTCATTCGAACTTCCTTCCTTCATTCGC

                            *      2720         *      2740         *      2760         *      2780         *      2800
H1_VCG               : ACCGCCTTCTTTCACCCCCCTCAGACGAGACCCTCTCCAAGCTGCTCAACCTGCTCTCGACCGGTTCGCACGAACTCCTCGCGGCCGAGGTGGAAAAGCGCGGC : 2468
H2                   : ACCGCCTTCTTTCACCCCCCTCAGACGAGACCCTCTCCAAGCTGCTCAACCTGCTCTCGACCGGTTCGCACGAACTCCTCGCGGCCGAGGTGGAAAAGCGCGGC : 2504
H3                   : ACCGCCTTCTTTCACCCCCCTCAGACGAGACCCTCTCCAAGCTGCTCAACCTGCTCTCGACCGGTTCGCACGAACTCCTCGCGGCCGAGGTGGAAAAGCGCGGC : 2576
H4                   : ACCGCCTTCTTTCACCCCCCTCAGACGAGACCCTCTCCAAGCTGCTCAACCTGCTCTCGACCGGTTCGCACGAACTCCTCGCGGCCGAGGTGGAAAAGCGCGGC : 2540
H5                   : ACCGCCTTCTTTCACCCCCCTCAGACGAGACCCTCTCCAAGCTGCTCAACCTGCTCTCGACCGGTTCGCACGAACTCCTCGCGGCCGAGGTGGAAAAGCGCGGC : 2540
H6                   : ACCGCCTTCTTTCACCCCCCTCAGACGAGACTCTCTCCAAGCTGCTCAACCTGCTCTCGACCGGTTCGCACGAACTCCTCGCGGCCGAGGTGGAAAAGCGCGGC : 2459
H7                   : ACCGCCTTCTTTCACCCCCCTCAGACGAGACCCTCTCCAAGCTGCTCAACCTGCTCTCGACCGGTTCGCACGAACTCCTCGCGGCCGAGGTGGAAAAGCGCGGC : 2477
H8                   : ACCGCCTTCTTTCACCCCCCTCAGACGAGACCCTCTCCAAGCTGCTCAACCTGCTCTCGACCGGTTCGCACGAACTCCTCGCGGCCGAGGTGGAAAAGCGCGGC : 2513
H9                   : ACCGCCTTCTTTCACCCCCCTCAGACGAGACCCTCTCCAAGCTGCTCAACCTGCTCTCGACCGGTTCGCACGAACTCCTCGCGGCCGAGGTGGAAAAGCGCGGC : 2595
H10                  : ACCGCCTTCTTTCACCCCCCTCAGACGAGACCCTCTCCAAGCTGCTCAACCTGCTCTCGACCGGTTCGCACGAACTCCTCGCGGCCGAGGTGGAAAAGCGCGGC : 2513
H11                  : ACCGCCTTCTTTCACCCCCCTCAGACGAGACCCTCTCCAAGCTGCTCAACCTGCTCTCGACCGGTTCGCACGAACTCCTCGCGGCCGAGGTGGAAAAGCGCGGC : 2522
H12                  : ACCGCCTTCTTTCACCCCCCTCAGACGAGACCCTCTCCAAGCTGCTCAACCTGCTCTCGACCGGTTCGCACGAACTCCTCGCGGCCGAGGTGGAAAAGCGCGGC : 2504
H13                  : ACCGCCTTCTTTCACCCCCCTCAGACGAGACCCTCTCCAAGCTGCTCAACCTGCTCTCGACCGGTTCGCACGAACTCCTCGCGGCCGAGGTGGAAAAGCGCGGC : 2576
H14                  : ACCGCCTTCTTTCACCCCCCTCAGACGAGACCCTCTCCAAGCTGCTCAACCTGCTCTCGACCGGTTCGCACGAACTCCTCGCGGCCGAGGTGGAAAAGCGCGGC : 2522
H15                  : ACCGCCTTCTTTCACCCCCCTCAGACGAGACCCTCTCCAAGCTGCTCAACCTGCTCTCGACCGGTTCGCACGAACTCCTCGCGGCCGAGGTGGAAAAGCGCGGC : 2504
H16                  : ACCGCCTTCTTTCACCCCCCTCAGACGAGACCCTCTCCAAGCTGCTCAACCTGCTCTCGACCGGTTCGCACGAACTCCTCGCGGCCGAGGTGGAAAAGCGCGGC : 2468
H17                  : ACCGCCTTCTTTCACCCCCCTCAGACGAGACCCTCTCCAAGCTGCTCAACCTGCTCTCGACCGGTTCGCACGAACTCCTCGCGGCCGAGGTGGAAAAGCGCGGC : 2684
H18                  : ACCGCCTTCTTTCACCCCCCTCAGACGAGACCCTCTCCAAGCTGCTCAACCTGCTCTCGACCGGTTCGCACGAACTCCTCGCGGCCGAGGTGGAAAAGCGCGGC : 2468
H19                  : ACCGCCTTCTTTCACCCCCCTCAGACGAGACCCTCTCCAAGCTGCTCAACCTGCTCTCGACCGGTTCGCACGAACTCCTCGCGGCCGAGGTGGAAAAGCGCGGC : 2522
H20                  : ACCGCCTTCTTTCACCCCCCTCAGACGAGACTCTCTCCAAGCTGCTCAACCTGCTCTCGACCGGTTCGCACGAACTCCTCGCGGCCGAGGTGGAAAAGCGCGGC : 2504
H21                  : ACCGCCTTCTTTCACCCCCCTCAGACGAGACTCTCTCCAAGCTGCTCAACCTGCTCTCGACCGGTTCGCACGAACTCCTCGCGGCCGAGGTGGAAAAGCGCGGC : 2540
H22                  : ACCGCCTTCTTTCACCCCCCTCAGACGAGACTCTCTCCAAGCTGCTCAACCTGCTCTCGACCGGTTCGCACGAACTCCTCGCGGCCGAGGTGGAAAAGCGCGGC : 2504
H23                  : ACCGCCTTCTTTCACCCCCCTCAGACGAGACCCTCTCCAAGCTGCTCAACCTGCTCTCGACCGGTTCGCACGAACTCCTCGCGGCCGAGGTGGAAAAGCGCGGC : 2540
H24                  : ACCGCCTTCTTTCACCCCCCTCAGACGAGACCCTCTCCAAGCTGCTCAACCTGCTCTCGACCGGTTCGCACGAACTCCTCGCGGCCGAGGTGGAAAAGCGCGGC : 2549
```

```
H25                  : ACCGCCTTCTTTCACCCCCCTCAGACGAGACCCTCTCCAAGCTGCTCAACCTGCTCTCGACCGGTTCGCACGAACTCCTCGCGGCCGAGGTGGAAAAGCGCGGC : 2504
H26_Mauritania-I_ref : ACCGCCTTCTTTCACCCCCCTCAGACGAGACCCTCTCCAAGCTGCTCAACCTGCTCTCGACCGGTTCGCACGAACTCCTCGCGGCCGAGGTGGAAAAGCGCGGC : 2594
H27_India-VII_refseq : ACCGCCTTCTTTCACCCCCCTCAGACGAGACCCTCTCCAAGCTGCTCAACCTGCTCTCGACCGGTTCGCACGAACTCCTCGCGGCCGAGGTGGAAAAGCGCGGC : 2477
H28_Ctg_refsequence  : ACCGCCTTCTTTCACCCCCCTCAGACGAGACCCTCTCCAAGCTGCTCAACCTGCTCTCGACCGGTTCGCACGAACTCCTCGCGGCCGAGGTGGAAAAGCGCGGC : 2549
H29_Salvador-I_refse : ACCGCCTTCTTTCACCCCCCTCAGACGAGACCCTCTCCAAGCTGCTCAACCTGCTCTCGACCGGTTCGCACGAACTCCTCGCGGCCGAGGTGGAAAAGCGCGGC : 2549
H30_NKorea_refsequen : ACCGCCTTCTTTCACCCCCCTCAGACGAGACCGTCTCCAAGCTGCTCAACCTGCTCTCGACCGGTTCGCACGAACTCCTCGCGGCCGAGGTGGAAAAGCGCGGC : 2459
H31_P.vivax01_refseq : ACCGCCTTCTTTCACCCCCCTCAGACGAGACCCTCTCCAAGCTGCTCAACCTGCTCTCGACCGGTTCGCACGAACTCCTCGCGGCCGAGGTGGAAAAGCGCGGC : 2540
H32_Brazil-I_refsequ : ACCGCCTTCTTTCACCCCCCTCAGACGAGACCCTCTCCAAGCTGCTCAACCTGCTCTCGACCGGTTCGCACGAACTCCTCGCGGCCGAGGTGGAAAAGCGCGGC : 2468
                       ACCGCCTTCTTTCACCCCCCTCAGACGAGACccTCTCCAAGCTGCTCAACCTGCTCTCGACCGGTTCGCACGAACTCCTCGCGGCCGAGGTGGAAAAGCGCGGC

                        *      2820         *      2840         *      2860         *      2880         *      2900         *
H1_VCG               : TTCGACGAGGACTACATCCAGGAGGAAATCAAAAACATAAACGAAAGCGACAACAACGTTCGGGACAAGGGCGAAGACGAAGTGGAAAATCTCATTTTTGAAGA : 2572
H2                   : TTCGACGAGGACTACATCCAGGAGGAAATCAAAAACATAAACGAAAGCGACAACAACGTTCGGGACAAGGGCGAAGACGAAGTGGAAAATCTCATTTTTGAAGA : 2608
H3                   : TTCGACGAGGACTACATCCAGGAGGAAATCAAAAACATAAACGAAAGCGACAACAACGTTCGGGACAAGGGCGAAGACGAAGTGGAAAATCTCATTTTTGAAGA : 2680
H4                   : TTCGACGAGGACTACATCCAGGAGGAAATCAAAAACATAAACGAAAGCGACAACAACGTTCGGGACAAGGGCGAAGACGAAGTGGAAAATCTCATTTTTGAAGA : 2644
H5                   : TTCGACGAGGACTACATCCAGGAGGAAATCAAAAACATAAACGAAAGCGACAACAACGTTCGGGACAAGGGCGAAGACGAAGTGGAAAATCTCATTTTTGAAGA : 2644
H6                   : TTCGACGAGGACTACATCCAGGAGGAAATCAAAAACATAAACGAAAGCGACAACAACGTTCGGGACAAGGGCGAAGACGAAGTGGAAAATCTCATTTTTGAAGA : 2563
H7                   : TTCGACGAGGACTACATCCAGGAGGAAATCAAAAACATAAACGAAAGCGACAACAACGTTCGGGACAAGGGCGAAGACGAAGTGGAAAATCTCATTTTTGAAGA : 2581
H8                   : TTCGACGAGGACTACATCCAGGAGGAAATCAAAAACATAAACGAAAGCGACAACAACGTTCGGGACAAGGGCGAAGACGAAGTGGAAAATCTCATTTTTGAAGA : 2617
H9                   : TTCGACGAGGACTACATCCAGGAGGAAATCAAAAACATAAACGAAAGCGACAACAACGTTCGGGACAAGGGCGAAGACGAAGTGGAAAATCTCATTTTTGAAGA : 2699
H10                  : TTCGACGAGGACTACATCCAGGAGGAAATCAAAAACATAAACGAAAGCGACAACAACGTTCGGGACAAGGGCGAAGACGAAGTGGAAAATCTCATTTTTGAAGA : 2617
H11                  : TTCGACGAGGACTACATCCAGGAGGAAATCAAAAACATAAACGAAAGCGACAACAACGTTCGGGACAAGGGCGAAGACGAAGTGGAAAATCTCATTTTTGAAGA : 2626
H12                  : TTCGACGAGGACTACATCCAGGAGGAAATCAAAAACATAAACGAAAGCGACAACAACGTTCGGGACAAGGGCGAAGACGAAGTGGAAAATCTCATTTTTGAAGA : 2608
H13                  : TTCGACGAGGACTACATCCAGGAGGAAATCAAAAACATAAACGAAAGCGACAACAACGTTCGGGACAAGGGCGAAGACGAAGTGGAAAATCTCATTTTTGAAGA : 2680
H14                  : TTCGACGAGGACTACATCCAGGAGGAAATCAAAAACATAAACGAAAGCGACAACAACGTTCGGGACAAGGGCGAAGACGAAGTGGAAAATCTCATTTTTGAAGA : 2626
H15                  : TTCGACGAGGACTACATCCAGGAGGAAATCAAAAACATAAACGAAAGCGACAACAACGTTCGGGACAAGGGCGAAGACGAAGTGGAAAATCTCATTTTTGAAGA : 2608
H16                  : TTCGACGAGGACTACATCCAGGAGGAAATCAAAAACATAAACGAAAGCGACAACAACGTTCGGGACAAGGGCGAAGACGAAGTGGAAAATCTCATTTTTGAAGA : 2572
H17                  : TTCGACGAGGACTACATCCAGGAGGAAATCAAAAACATAAACGAAAGCGACAACAACGTTCGGGACAAGGGCGAAGACGAAGTGGAAAATCTCATTTTTGAAGA : 2788
H18                  : TTCGACGAGGACTACATCCAGGAGGAAATCAAAAACATAAACGAAAGCGACAACAACGTTCGGGACAAGGGCGAAGACGAAGTGGAAAATCTCATTTTTGAAGA : 2572
H19                  : TTCGACGAGGACTACATCCAGGAGGAAATCAAAAACATAAACGAAAGCGACAACAACGTTCGGGACAAGGGCGAAGACGAAGTGGAAAATCTCATTTTTGAAGA : 2626
H20                  : TTCGACGAGGACTACATCCAGGAGGAAATCAAAAACATAAACGAAAGCGACAACAACGTTCGGGACAAGGGCGAAGACGAAGTGGAAAATCTCATTTTTGAAGA : 2608
H21                  : TTCGACGAGGACTACATCCAGGAGGAAATCAAAAACATAAACGAAAGCGACAACAACGTTCGGGACAAGGGCGAAGACGAAGTGGAAAATCTCATTTTTGAAGA : 2644
H22                  : TTCGACGAGGACTACATCCAGGAGGAAATCAAAAACATAAACGAAAGCGACAACAACGTTCGGGACAAGGGCGAAGACGAAGTGGAAAATCTCATTTTTGAAGA : 2608
H23                  : TTCGACGAGGACTACATCCAGGAGGAAATCAAAAACATAAACGAAAGCGACAACAACGTTCGGGACAAGGGCGAAGACGAAGTGGAAAATCTCATTTTTGAAGA : 2644
H24                  : TTCGACGAGGACTACATCCAGGAGGAAATCAAAAACATAAACGAAAGCGACAACAACGTTCGGGACAAGGGCGAAGACGAAGTGGAAAATCTCATTTTTGAAGA : 2653
H25                  : TTCGACGAGGACTACATCCAGGAGGAAATCAAAAACATAAACGAAAGCGACAACAACGTTCGGGACAAGGGCGAAGACGAAGTGGAAAATCTCATTTTTGAAGA : 2608
H26_Mauritania-I_ref : TTCGACGAGGACTACATCCAGGAGGAAATCAAAAACATAAACGAAAGCGACAACAACGTTCGGGACAAGGGCGAAGACGAAGTGGAAAATCTCATTTTTGAAGA : 2698
H27_India-VII_refseq : TTCGACGAGGACTACATCCAGGAGGAAATCAAAAACATAAACGAAAGCGACAACAACGTTCGGGACAAGGGCGAAGACGAAGTGGAAAATCTCATTTTTGAAGA : 2581
H28_Ctg_refsequence  : TTCGACGAGGACTACATCCAGGAGGAAATCAAAAACATAAACGAAAGCGACAACAACGTTCGGGACAAGGGCGAAGACGAAGTGGAAAATCTCATTTTTGAAGA : 2653
H29_Salvador-I_refse : TTCGACGAGGACTACATCCAGGAGGAAATCAAAAACATAAACGAAAGCGACAACAACGTTCGGGACAAGGGCGAAGACGAAGTGGAAAATCTCATTTTTGAAGA : 2653
H30_NKorea_refsequen : TTCGACGAGGACTACATCCAGGAGGAAATCAAAAACATAAACGAAAGCGACAACAACGTTCGGGACAAGGGCGAAGACGAAGTGGAAAATCTCATTTTTGAAGA : 2563
H31_P.vivax01_refseq : TTCGACGAGGACTACATCCAGGAGGAAATCAAAAACATAAACGAAAGCGACAACAACGTTCGGGACAAGGGCGAAGACGAAGTGGAAAATCTCATTTTTGAAGA : 2644
H32_Brazil-I_refsequ : TTCGACGAGGACTACATCCAGGAGGAAATCAAAAACATAAACGAAAGCGACAACAACGTTCGGGACAAGGGCGAAGACGAAGTGGAAAATCTCATTTTTGAAGA : 2572
                       TTCGACGAGGACTACATCCAGGAGGAAATCAAAAACATAAACGAAAGCGACAACAACGTTCGGGACAAGGGCGAAGACGAAGTGGAAAATCTCATTTTTGAAGA
```

```
H1_VCG                : CTTGTGA : 2579
H2                    : CTTGTGA : 2615
H3                    : CTTGTGA : 2687
H4                    : CTTGTGA : 2651
H5                    : CTTGTGA : 2651
H6                    : CTTGTGA : 2570
H7                    : CTTGTGA : 2588
H8                    : CTTGTGA : 2624
H9                    : CTTGTGA : 2706
H10                   : CTTGTGA : 2624
H11                   : CTTGTGA : 2633
H12                   : CTTGTGA : 2615
H13                   : CTTGTGA : 2687
H14                   : CTTGTGA : 2633
H15                   : CTTGTGA : 2615
H16                   : CTTGTGA : 2579
H17                   : CTTGTGA : 2795
H18                   : CTTGTGA : 2579
H19                   : CTTGTGA : 2633
H20                   : CTTGTGA : 2615
H21                   : CTTGTGA : 2651
H22                   : CTTGTGA : 2615
H23                   : CTTGTGA : 2651
H24                   : CTTGTGA : 2660
H25                   : CTTGTGA : 2615
H26_Mauritania-I_ref  : CTTGTGA : 2705
H27_India-VII_refseq  : CTTGTGA : 2588
H28_Ctg_refsequence   : CTTGTGA : 2660
H29_Salvador-I_refse  : CTTGTGA : 2660
H30_NKorea_refsequen  : CTTGTGA : 2570
H31_P.vivax01_refseq  : CTTGTGA : 2651
H32_Brazil-I_refsequ  : CTTGTGA : 2579
                        CTTGTGA
```

11.3 Anexo 3. Publicación

El trabajo fue publicado en la revista Malaria Journal. **Buitrago SP**, Garzón-Ospina D, Patarroyo MA. *Size polymorphism and low sequence diversity in the locus encoding the Plasmodium vivax rhoptry neck protein 4 (PvRON4) in Colombian isolates.* Malar J. 2016 Oct 18; 15(1):501.