MICHELLE CRISTINA DO COUTO BRANDI

CARACTERIZAÇÃO E PADRONIZAÇÃO DE MARCADORES GENÉTICOS

PARA O ESTUDO DE POLIMORFISMOS EM *Plasmodium vivax*

Dissertação apresentada ao Programa de Pós-Graduação em Biologia da Relação Patógeno-Hospedeiro do Instituto de Ciências Biomédicas da Universidade de São Paulo para obtenção do título de Mestre em Ciências.

São Paulo
2013

MICHELLE CRISTINA DO COUTO BRANDI

CARACTERIZAÇÃO E PADRONIZAÇÃO DE MARCADORES GENÉTICOS

PARA O ESTUDO DE POLIMORFISMOS EM *Plasmodium vivax*

Dissertação apresentada ao Programa de Pós-Graduação em Biologia da Relação Patógeno-Hospedeiro do Instituto de Ciências Biomédicas da Universidade de São Paulo, para obtenção de título de Mestre em Ciências.

Área de concentração: Biologia da Relação Patógeno-Hospedeiro.

Orientador: Prof. Dr. Marcelo Urbano Ferreira

Versão original

São Paulo
2013

DADOS DE CATALOGAÇÃO NA PUBLICAÇÃO (CIP)
Serviço de Biblioteca e Informação Biomédica do
Instituto de Ciências Biomédicas da Universidade de São Paulo

© reprodução total

Brandi, Michelle Cristina do Couto.

Caracterização e padronização de marcadores genéticos para o estudo de polimorfismos de *Plasmodium vivax* / Michelle Cristina do Couto Brandi. -- São Paulo, 2013.

Orientador: Prof. Dr. Marcelo Urbano Ferreira.

Dissertação (Mestrado) – Universidade de São Paulo. Instituto de Ciências Biomédicas. Departamento de Parasitologia. Área de concentração: Biologia da Relação Patógeno-Hospedeiro. Linha de pesquisa: x.

Versão do título para o inglês: Characterization and standardization of genetic markers in the study of *Plasmodium vivax*.

1. Malária 2. *Plasmodium* 3. Genética 4. Doenças parasitárias
5. Marcador molecular Amplificação de genes I. Ferreira,
Prof. Dr. Marcelo Urbano II. Universidade de São Paulo. Instituto de
Ciências Biomédicas. Programa de Pós-Graduação em Biologia da
Relação Patógeno-Hospedeiro III. Título.

ICB/SBIB021/2013

UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE CIÊNCIAS BIOMÉDICAS

Candidato(a): Michelle Cristina do Couto Brandi.

Título da Caracterização e padronização de marcadores genéticos para o estudo de polimorfismos de *Plasmodium vivax*.

Orientador(a): Prof. Dr. Marcelo Urbano Ferreira.

A Comissão Julgadora dos trabalhos de Defesa da **Dissertação de Mestrado**, em sessão pública realizada a/...../....., considerou

() **Aprovado(a)**

() **Reprovado(a)**

Examinador(a): Assinatura:
Nome:
Instituição:

Examinador(a): Assinatura:
Nome:
Instituição:

Presidente: Assinatura:
Nome:
Instituição:

**UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE CIÊNCIAS BIOMÉDICAS**

Cidade Universitária "Armando de Salles Oliveira"
Av. Prof. Lineu Prestes, 2415 - cep. 05508-000 São Paulo, SP - Brasil
Telefone : (55) (11) 3091.7733 telefax : (55) (11) 3091-8405
e-mail: cep@icb.usp.br

São Paulo, 14 de outubro de 2010.

PARECER 963/CEP

A *Comissão de Ética em Pesquisas com Seres Humanos* do ICB, na sessão de 21.09.2010, **APROVOU** o projeto intitulado: "Caracterização e padronização de marcadores genéticos no estudo de polimorfismos de *Plasmodium vivax* na Amazônia rural brasileira" sob responsabilidade de execução dos autores Prof. Dr. **MARCELO URBANO FERREIRA** e a aluna **MICHELLE CRISTINA DO COUTO BRANDI**.

Cabe aos pesquisadores executantes elaborar e apresentar a este Comitê, relatórios anuais (parciais ou final), de acordo com a resolução 196/06 do Conselho Nacional da Saúde, item IX. 2 letra c.

O primeiro relatório deverá ser encaminhado à Secretaria deste CEP em
21.09.2011

Atenciosamente,

Prof. Dr. PAOLO M. A. ZANOTTO
Coordenador da Comissão de Ética em
Pesquisas com Seres Humanos - ICB/USP

À Edna, mamãezinha querida Sol da minha vida, fonte de inspiração e exemplo, companheira de vida, por tudo tenho dela que é imensurável.... e inestimável.

AGRADECIMENTOS

Ver o fruto deste trabalho é motivo de orgulho e felicidade para mim. Ao longo destes anos de aprendizado, convivência e troca de experiências tenho muito a agradecer aos meus colegas que, de alguma forma, fizeram parte deste processo de desenvolvimento profissional e pessoal.

O meu primeiro agradecimento é ao Professor Marcelo Urbano Ferreira, meu orientador, pela oportunidade de fazer parte deste renomado grupo de pesquisa, pela confiança em mim depositada, pelo conhecimento compartilhado, pois, através desta experiência enriquecedora e deste ensejo, poderei ir muito além.

Agradeço à Pamela, por me acolher no laboratório, pelos engrandecedores ensinamentos de bancada, por dividir comigo a sua experiência e pela boa convivência no dia a dia.

Sou infinitamente grata a Melissa, que esteve presente desde o meu primeiro dia de laboratório, que me ajudou em tudo, da adaptação à nova cidade até os mais difíceis protocolos de pesquisa, sem dúvida ela foi um anjo que eu conheci.

Agradeço a Priscila, minha companheira de mestrado, que, por tantas vezes nos vimos pensando, estudando e discutindo procedimentos novos, programas e artigos. Apreendi muito com ela, e, sem dúvida isto fez minha experiência mais fácil e mais agradável.

Agradeço a Raquel e a Nath pelo companheirismo e sincera amizade em todas as oportunidades que passamos juntas, seja no laboratório, seja durante os cafezinhos da tarde, elas serão inesquecíveis e desejo que nossa amizade se prolongue e fortaleça sempre.

Obrigada às amigas Miluska, Vanessa, Bianca, Amanda, Simone, Michelle, Carol, Kézia, Camilla pelos momentos que tivemos, aprendi muito com o passamos juntas. A companhia de vocês fez esta experiência mais leve e feliz. Obrigada, Guilherme, pela ajuda com os microssatélites também de grande importância no início da minha jornada. Obrigada, Mônica, pela ajuda de moradia. Obrigada a Maria José pela ajuda com a logística do laboratório. Foi muito prazeroso também conhecer alunas de outras nacionalidades: a Diana da Colômbia, a Krystal e a Gurmat dos EUA, Ana Maria do Panamá.

Agradeço aos pesquisadores colaboradores: Scott J Westenberger, da Scripps Research Institute, Estados Unidos, pela ajuda com os microarranjos, ideia inicial do estudo, aos pesquisadores que nos

cederam as amostras para serem utilizadas neste projeto: Richard Fairhurst, do National Institute of Allergy and Infectious Diseases, Estados Unidos, Nadira Karunaweera da Universidade de Colombo, Sri Lanka e Alexandre Januário da Silva do CDC - Centers for Disease Control and Prevention, Estados Unidos.

Agradeço aos professores, funcionários e amigos do departamento de parasitologia, em especial à Carminha por toda a ajuda com as corridas de microssatélites e ao Marcinho pela ajuda com o sequenciamento e as dicas de clonagem. Muito obrigada também ao Wesley e a todo o grupo do laboratório de pesquisa do Professor Gerhard Wunderlich pela ajuda em esclarecimentos científicos e pelo socorro, sempre que necessitei. Sou muito grata a Catarina, pela dica e o contato dos genes sintéticos. Agradeço ao professor Ariel Silber pela ajuda e as dicas de clonagem e pelas demais conversas e aconselhamentos sempre bem-vindos.

Um especial agradecimento à minha mamãe querida, Sol da minha vida, por sempre torcer tanto por mim, por me guiar, me apoiar, me entender, me escutar, me aconselhar, me esclarecer, me acalmar, ser meu grande exemplo de vida e meu grande amor. Agradeço ao Roberto, meu paizão pelo apoio incondicional, pela presença de sempre que amo muito. Amo muitíssimo e agradeço a existência de minhas irmãs Lauren, Larissa, Carol e “Lolitinha” a elas devo todo meu conhecimento de convivência, tolerância, compreensão e amor. Agradeço a Dona Terezinha, minha avó por me amar tanto e rezar tanto por mim, sou apaixonada por ela! Agradeço as minhas irmãs de coração Fernanda e Flávia e ao Ronaldo, por sempre serem meus grandes companheiros e conselheiros e por eu poder sempre contar com a amizade deles. Obrigada ao Edmar meu primo pela nossa breve convivência morando juntos, ele não sabe, mas me ajudou muito!

Meu tenro agradecimento ao Pablo, que faz parte da minha vida fora do laboratório, que torna meus dias completos e meus finais de semana muito felizes, ele é um exemplo de determinação, comprometimento e lealdade. Ele faz com que a garoa de São Paulo seja mais colorida e eu, mais feliz.

Finalmente, agradeço a Deus, por acreditar que eu mereço tudo o que tenho e colocar todos estes anjos na minha vida e sem os quais nada disto seria possível.

Muito obrigada a todos!

Este trabalho contou com o apoio financeiro do National Institutes of Health (NIH), da Fundação do Amparo à Pesquisa do Estado de São Paulo (FAPESP) e do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

“Quando nasci um anjo esbelto, desses que tocam trombeta, anunciou: vai carregar bandeira(...) Vai ser coxo na vida é maldição para homem. Mulher é desdobrável. Eu sou.”

Adélia Prado

“Só se pode chamar de ciência o conjunto de receitas que funciona sempre. Todo o resto é literatura.”

Paul Valéry

RESUMO

Brandi, MCC. Caracterização e padronização de marcadores genéticos para o estudo de polimorfismos de *Plasmodium vivax*. [dissertação (Mestrado em Biologia da Relação Patógeno-Hospedeiro)]. São Paulo: Instituto de Ciências Biomédicas, Universidade de São Paulo; 2013.

Malária é uma importante endemia parasitária no Brasil com mais de 306.000 casos clínicos notificados na Amazônia brasileira em 2011. Cerca de 85% desses casos são causados por *Plasmodium vivax*. Compreender os padrões de diversificação destes patógenos humanos no espaço e no tempo é uma questão central da ecologia e da genética de populações, com claras implicações para o planejamento de estratégias de controle. Este trabalho teve como objetivo caracterizar e padronizar métodos para a tipagem molecular de marcadores genéticos, para futuro uso em estudos de genética populacional de *P. vivax* na população da Amazônia brasileira. As amostras sanguíneas utilizadas neste estudo foram colhidas de pacientes com infecção por *P. vivax*, oriundos do Acre (Brasil), Pursat (Camboja), Tricomalee (Sri Lanka) e de pacientes com infecção de malária vivax importada de diferentes localidades, diagnosticada no CDC (Centers for Disease Control and Prevention), Estados Unidos. Foram selecionados polimorfismos de única base (SNPs) distribuídos ao longo de cromossomos distintos de *P. vivax*; para estes polimorfismos, sete ensaios de tipagem de SNPs, que se mostraram polimórficos em amostras de parasitos do Brasil e de outras partes do mundo, foram padronizados com sucesso: (a) uma substituição sinônima de nucleotídeo no gene que codifica uma proteína hipotética PVX_119300, localizada no cromossomo 8; (b) uma substituição sinônima no gene que confere resistência à drogas, PvMDR-1, localizado no cromossomo 8; (c) duas substituições sinônimas no gene que codifica uma proteína de superfície do merozoíto MSP-1, localizado no cromossomo 7 (d) uma substituição sinônima em uma região intergênica, localizada no cromossomo 8 (e) duas substituições sinônimas em um íntron, localizado no cromossomo 13. Com a tipagem molecular destes polimorfismos nas amostras selecionadas neste estudo, foi possível definir haplótipos que caracterizam cada amostra, assim como identificar infecções geneticamente mistas (co-ocorrência de clones distintos na mesma amostra). No entanto, com este conjunto de marcadores não foi possível agrupar amostras de acordo com sua localização geográfica, por estes marcadores não serem suficientemente informativos do ponto de vista genético. Os sete SNPs avaliados, quando comparados a 13 marcadores de DNA microssatélite aplicados ao mesmo conjunto de amostras, revelaram menor proporção de infecções geneticamente mistas e menor diversidade genética, possivelmente devido ao número relativamente pequeno de SNPs utilizados e seu menor polimorfismo.

Palavras-chave: Malária. *Plasmodium vivax*. Genética de populações. Marcadores moleculares. Microssatélites. SNPs.

ABSTRACT

Brandi, MCC. Characterization and standardization of genetic markers in the study of *Plasmodium vivax* polymorphisms. [dissertation (Master Thesis in Biology of the Relation Pathogen-Host)]. São Paulo: Instituto de Ciências Biomédicas, Universidade de São Paulo: 2013.

Malaria is an important endemic parasitic disease in Brazil with more than 306,000 clinical cases reported in the Brazilian Amazon in 2011. About 85% of these cases are caused by *Plasmodium vivax*. Understanding the patterns of diversification of these human pathogens in space and time is a key issue of ecology and population genetics, with clear implications for planning control strategies. This study aimed to characterize and standardize methods for molecular typing of genetic markers to be used in the future in studies of population genetics of *P. vivax* population in the Brazilian Amazon. Blood samples used in this study were collected from patients with infection caused by *P. vivax*, from Acre (Brazil), Pursat (Cambodia), Tricomalee (Sri Lanka) and patients with vivax malaria infection imported from different localities and diagnosed at the CDC (Centers for Disease Control and Prevention), United States. Single nucleotide polymorphisms (SNPs), distributed over different *P. vivax* chromosomes were chosen and seven typing assays were successfully standardized, which target SNPs shown to be polymorphic in parasite samples from Brazil and other parts of the world: (a) one synonymous nucleotide substitution in the gene encoding a hypothetical protein PVX_119300, located on chromosome 8 (b) one synonymous substitution in the drug resistance gene, PvMDR-1, located on chromosome 8 (c) two synonymous substitutions in the gene encoding a merozoite surface protein MSP-1, located in chromosome 7 (d) a synonymous substitution in the intergenic region, located on chromosome 8 (e) two synonymous substitutions in an intron, located on chromosome 13. With the molecular typing of polymorphisms it was possible to define haplotypes that characterize each sample, as well as to identify genetically mixed infections (co-occurrence of different clones in the same sample). However, with this markers set, it was not possible to group samples according to their geographical location, because probably they are not sufficiently genetically informative. Compared to 13 microsatellite markers used to type the same samples, these seven SNPs revealed a lower proportion of mixed-clone infections and lower genetic diversity, possibly due to the relatively small number of SNPs assayed and their lower polymorphism.

Keywords: Malaria. *Plasmodium vivax*. Population genetics. Molecular markers. Microsatellites. SNPs

LISTA DE ILUSTRAÇÕES

Figura 1- Árvore filogenética de <i>Plasmodium</i> causadores de malária em diferentes espécies.....	23
Figura 2- Distribuição dos casos de malária no Brasil.....	30
Figura 3- Número de episódios de malária confirmados registrados por ano no Brasil entre 1960 e 2007.....	31
Figura 4- Distribuição mundial da malária vivax em 2010.....	33
Figura 5- Haplótipos de SNPs para os cinco isolados de <i>P. vivax</i>	40
Figura 6- Resultados da diversidade de SNPs dentre as espécies de <i>Plasmodium</i>	43
Figura 7- Filogramas <i>Neighbor-joining</i> de <i>P. vivax</i> e <i>P. falciparum</i> construídas a partir de SNPs presumidamente neutros que ocorrem em quadras degeneradas de sites codificadores sinônimos.....	44
Figura 8- Resultado do funcionamento das sondas do tipo TaqMan.....	45
Figura 9- Mapa dos Estados do Acre e Amazonas, noroeste do Brasil mostrando a localização das áreas de estudo, Granada, Plácido de Castro e Remansinho.....	57
Figura 10- Mapa da localização de Camboja e Sri Lanka, países dos quais foram analisadas amostras de DNA para este estudo.....	59
Figura 11- Em destaque, dois SNPs sinônimos encontrados no cromossomo 13 de <i>P. vivax</i> em isolados da Índia e Brasil.....	73
Figura 12- Exemplo de curva-padrão, obtida em PCR de tempo real, para a quantificação de parasitemias de <i>Plasmodium vivax</i>	76
Figura 13- Vetor pUC57 com o Gene Sintético 3 inserido e o resultado de seu sequenciamento.....	82
Figura 14 - Discriminação alélica, resultante do ensaio TaqMan 11.....	96

LISTA DE TABELAS

Tabela 1- Oligonucleotídeos utilizados para a amplificação e sequenciamento de 18 regiões genômicas com possíveis polimorfismos (SNPs) em isolados brasileiros de <i>Plasmodium vivax</i>	61
Tabela 2- Diferenças entre o resultado esperado e o resultado observado em três <i>loci</i> analisados de Sal-1 e de amostras do Acre: Acre5, Acre9, e Acre10.....	65
Tabela 3- Caracterização dos primeiros polimorfismos selecionados para tipagem de SNPs.....	67
Tabela 4- Sequência de oligonucleotídeos indianos envolvendo a amplificação de seis fragmentos de DNA não – codificantes com sua temperatura de anelamento.....	71
Tabela 5- Características dos 14 polimorfismos selecionados para uso em estudos populacionais.....	74
Tabela 6- Sequências de oligonucleotídeos e sondas utilizados nos ensaios TaqMan® marcadas com VIC (Sonda 1) e FAM (Sonda 2) para tipagem de SNPs de <i>P.vivax</i> selecionados.....	78
Tabela 7- SNPs tipados na plataforma StepOnePlus v.2.1. Ensaio correspondente, localização cromossômica e posição do SNP definida com base no genoma do isolado Sal-I (Carlton et al., 2008).....	80
Tabela 8- Sequências de genes sintéticos utilizados nos ensaios TaqMan® como controle positivo.....	83
Tabela 9- Oligonucleotídeos iniciadores (Oligos) dos 13 MSAT de <i>P. vivax</i>	85
Tabela 10- Dados de tipagem de 13 MSTAs analisados neste estudo.....	86
Tabela 11- Resultado da tipagem de SNPs pela plataforma Mastercycler Realplex gradient thermal cycler.....	94
Tabela 12- Descrição dos sete ensaios TaqMan padronizados.....	95
Tabela 13- Resultado da tipagem de SNPs. Identificação de gene, cromossomo e localização cromossômica	99
Tabela 14- Haplótipos resultantes das tipagens dos sete ensaios TaqMan.....	101
Tabela 15- Resultado da tipagem de polimorfismos pela metodologia TaqMan® , destacando a formação dos diferentes haplótipos e a detecção de infecções mistas.....	102
Tabela 16- Resultado da tipagem molecular com 13 microsatélites em isolados de <i>P.vivax</i>	104

LISTA DE ABREVIATURAS E SIGLAS

As abreviaturas utilizadas neste texto encontram-se aqui em ordem alfabética. Alguns termos técnicos de origem na língua inglesa não foram traduzidos por fazerem parte de glossário científico utilizado internacionalmente.

A-.....Adenina

AMA1-Anígeno de Membrana Apical –“ Apical Membrane Antigen”

ATP-..... Adenosina Trifosfato

cDNA- DNA Complementar – “complementary DNA”

CSP-..... Proteína de Circumsporozoíto – “Circumsporozoite Protein”

C-.....Citosina

CDC-..... Centro de Controle e Prevenção de Doenças “Centers for Disease Control and Prevention”

DBP-..... Proteína de Ligação Duffy –“ Duffy Biding Protein”

DNA-..... Ácido desoxirribinucleico – “Deoxyribonucleic acid”

dNTP-..... Trifosfatos de desoxinucleotídeos – “Deoxynucleotide triphosphates: dATP, dCTP, dGTP and dTTP”

EUA-..... Estados Unidos da América

F Oligonucleotídeo iniciador anterior – “forward”

G-..... Guanina

ICB-..... Instituto de Ciências Biomédicas

IPTG-..... Isopropil-d-D-galactosídeo

Kb-..... Quilo pares de bases –“ Kilo base pair”

KDa-..... Quilo Dalton – “Kilo Dalton”

L-..... Litro

LB-..... “Luria broth base”

MB-.....Megabase

MDR- Multirresistência a drogas –“Multi-resistance drugs”

MGB-..... “Minor Groove Binder”

MgCl₂- Cloreto de Magnésio

Min-..... Minutos

MR4- “Malaria Research and Reference Regent Resource Center”

MSAT- Microssatélite

MSP-..... Proteína de Superfície de Merozoíto – “Merozoite Surface Protein”
NBD-..... Domínio de Ligação de Nucleotídeo – “Nucleotide Binding Domain”
NCBI-..... “National Center for Biotechnology Information”
OMS-.....Organização Mundial de Saúde
Pb-..... Pares de bases – “Base Pairs”
PCR-..... Reação em cadeia da Polimerase – “Polymerase Chain Reaction”
Pg-.....Picogramas
Pmol-Picomol
Oligo-..... Oligonucleotídeo Iniciador
Pf - Plasmodium falciparum
Pv-..... Plasmodium vivax
R-..... Oligonucleotídeo reverso- “reverse”
RNA-.....Ácido ribonucleico
rRNA-..... Ácido ribonucléico ribossômico
Seg-..... Segundos
SNP-..... Polimorfismo de base única – “Single-Nucleotide Polymorphism”
T-..... Timina
Tm -..... Temperatura de dissociação – “Melting Temperature”
U-.....Unidade
USP-..... Universidade São Paulo
WGA-..... Amplificação genômica – “Whole Genome Amplification”
WHO-..... Organização Mundial de Saúde “World Health Organization”
YAC-..... Cromossomo Artificial de Levedura – “Yeast Artificial Chromosome”
° C-..... Grau Celsius
μL-.....Micro Litro
μM-.....Micro Molar

SUMÁRIO

1 INTRODUÇÃO.....	20
1.1 História da malária.....	21
1.2 Malária, a doença	22
1.3 Malária causada por <i>P. vivax</i>	25
1.4 Origem e características genéticas de <i>Plasmodium vivax</i>	26
1.5 A malária causada por <i>Plasmodium vivax</i> como problema de saúde pública	27
1.5.1 No Brasil	27
1.5.2 No Mundo	31
1.6 Infecções com múltiplos clones	33
1.7 Marcadores genéticos para estudos populacionais	34
1.7.1 Marcadores Neutros	35
1.7.1.1 Microsatélites	35
1.7.1.2 Polimorfismos de Base Única (SNPs)	38
1.7.1.3 Tipagem de SNPs com sondas TaqMan	44
1.7.2 Marcadores moleculares sob intensa pressão seletiva	47
1.7.2.1 PvMDR-1	47
1.7.2.2 AMA-1	48
1.7.2.3 MSP-1	48
1.7.2.4 MSP-3	48
1.7.2.5 MSP-4 e MSP-5	49
1.7.2.6 DBP	50
2 OBJETIVOS	51
3 MATERIAL E MÉTODOS	53

3.1 Aspectos éticos	54
3.2 Amostras de DNA	54
3.3 Áreas de estudo	55
3.3.1 Granada	55
3.3.2 Plácido de Castro	56
3.3.3 Pursat, Camboja	57
3.3.4 Tricomalee, Sri Lanka	58
3.3.5 Centro para Controle e Prevenção de Doenças (CDC), Estados Unidos da América	59
3.4 Marcadores moleculares	60
3.4.1 Seleção e caracterização de SNPs caracterizados por hibridização em microarranjos.....	60
3.4.2 Genotipagem e confirmação de SNPs caracterizados por microarranjos.....	63
3.4.3 Sequenciamento e Análises dos SNPs	63
3.5 Identificação de SNPs em sequências depositadas no GenBank	66
3.5.1 Critério de seleção das sequências codificadoras.....	66
3.5.2 Análises das sequências codificadoras.....	67
3.6 Seleção de SNPs previamente descritos na literatura.....	68
3.7 Análise de polimorfismos situados em região não-codificadora caracterizados por Gupta e colaboradores	69
3.7.1 Extração do DNA e amplificação genômica com os oligonucleotídeos de Gupta e colaboradores.....	70
3.7.2 Sequenciamento e análise de SNPs	71
3.8 Padronização de ensaios de tipo TaqMan para tipagem de SNPs.....	74
3.8.1 Quantificação do DNA das amostras utilizadas no estudo por PCR em tempo real.....	75

3.8.2 Padronização dos ensaios TaqMan® para tipagem de SNPs na plataforma Mastercycler Realplex gradient thermal cycler.....	76
3.8.3 Padronização dos ensaios de tipagem TaqMan® na plataforma StepOnePlus	79
3.9 Desenvolvimento de controles positivos através de genes sintetizados.....	81
3.10 Análise de microssatélites.....	84
3.10.1 Genotipagem de MSAT.....	84
3.10.2. Dados previamente descritos de resultados de MSTa.....	86
4 RESULTADOS	87
4.1 Análise da caracterização de marcadores moleculares – microarranjos.....	88
4.2 Análise de sequências de DNA de genes de cópia única disponíveis no GenBank.....	89
4.2.1 Análise das sequências codificadoras	89
4.2.2 Análise de SNPs previamente descritos na literatura.....	90
4.2.2.1 Cromossomo 8.....	90
4.2.2.2 Gene MDR-1.....	90
4.2.2.3 Proteínas hipotéticas codificadas por genes nos cromossomo 12 e 14	90
4.2.2.4 Polimorfismos situados em região não-codificadora do cromossomo 13.....	91
4.3 Análise da tipagem de SNPs na plataforma Mastercycler Realplex gradient thermal cycler	91
4.4 Análise dos ensaios de tipagem de SNPs pela na plataforma StepOnePlus v.2.1.....	95
4.5 Análise de haplótipos.....	101
4.6 Análise de tipagem molecular por microssatélites.....	103
5 DISCUSSÃO.....	106

5.1 Marcadores moleculares – microarranjos.....	107
5.2 Sequências de DNA de genes de cópia única disponíveis no GenBank e SNPs previamente descritos na literatura.....	108
5.3 Tipagem molecular de SNPs.....	108
5.4 Análise de SNPs comparada à análise de microssatélites.....	110
6 CONCLUSÃO.....	111
REFERÊNCIAS.....	113

1.1 História da malária

Denominada “malária”, originada da expressão italiana Medieval “ar ruim”, paludismo, que significa pântano em latim, impaludismo, febres terçã ou quartã, a malária pode ter acompanhado toda a história da humanidade. Parasitos próximos aos parasitos de humanos continuam comuns até os dias de hoje em chimpanzés. Referências às febres características da malária são descritas desde 2700 a.C. na China. A malária já foi comum na maioria da Europa e da América do Norte, onde já não é mais endêmica, embora casos de malária importada ainda ocorram.

A descoberta do parasito causador da malária foi feita em 1880 pelo médico francês Charles Louis Alphonse Laveran, que trabalhava em um hospital militar em Constantine, Argélia, ao observar parasitos em células vermelhas do sangue de um paciente infectado com malária. Foi então proposto que esta doença era causada por um protozoário protista (esta foi a primeira vez que um protozoário protista foi identificado como causador de uma doença). Por esta e por outras descobertas, este médico foi agraciado com o prêmio Nobel em 1907 em Fisiologia ou Medicina.

Em 1898, Sir Ronald Ross, médico da Grã-Bretanha que trabalhava em um hospital Geral da Presidência, em Calcutá, finalmente provou que a malária é transmitida por mosquitos. Por este trabalho, ele recebeu em 1902 um Prêmio Nobel em Medicina. Um ano depois, pesquisadores italianos mostraram que os anofelinos eram os vetores responsáveis pela transmissão da doença aos seres humanos.

O primeiro tratamento eficaz contra malária era extraído da casca da árvore de cinchona (que cresce nas encostas dos Andes, principalmente no Peru), que contém quinino como agente ativo, que foi isolado e nomeado pelos químicos franceses Pierre Joseph Pelletier e Bienaimé Joseph Caventou em 1820.

Os primeiros relatos dos sintomas clínicos da malária foram feitos em 400 a.C. por Hipócrates, há relatos da doença por toda a história dos impérios grego e romano. Personagens ilustres como Petrarca, Dante, Alighieri e até Alexandre, o Grande, foram infectados por este parasito que se propagou nos quatro cantos da Terra, com exceção de regiões polares e subpolares. No Novo Mundo, por falta de qualquer relato maia, asteca ou inca, uma das hipóteses é que a malária foi trazida com a colonização de espanhóis e portugueses e com a vinda de escravos trazidos da África.

Apesar de poder ser prevenida e curada, nas últimas décadas, esta infecção apresentou também um alto impacto socioeconômico, contribuindo para um aumento de custo na saúde pública, além de diminuir o produto interno bruto em até 1,3% em países com altos índices da doença (World Health Organization, 2011).

1.2 Malária, a doença

A malária é uma doença infecciosa, parasitária e endêmica na maioria dos ecossistemas tropicais e subtropicais do mundo. É caracterizada por episódios febris, precedidos por calafrios e seguidos por intensa sudorese e cefaléia. Os parasitos da malária pertencem ao filo Apicomplexa, família Plasmodiidae gênero *Plasmodium*, são parasitos intracelulares obrigatórios e infectam vários hospedeiros vertebrados, incluindo várias espécies de primatas não humanos (Cornejo, Escalante, 2006) atualmente são conhecidas cerca de 150 espécies causadoras de malária (de Carli, 2010).

A malária pode ser causada no homem por cinco espécies de plasmódios: *Plasmodium falciparum*, *Plasmodium vivax*, *Plasmodium malariae* e *Plasmodium ovale*. Mais recentemente, alguns casos humanos de malária também foram confirmados com *Plasmodium knowlesi*- uma malária de primatas que ocorre em áreas florestais do Sudeste da Ásia (WHO, 2010). A Figura 1 representa a árvore filogenética de *Plasmodium* causadores de malária em diferentes espécies.

Figura 1- Árvore filogenética de *Plasmodium* causadores de malária em diferentes espécies. Dendograma simplificado, não escalonado. Cada braço foi classificado com o nome da espécie, seguido pela representação gráfica de seus hospedeiros vertebrados naturais e o número de genomas sequenciados disponíveis publicados.

Fonte: Adaptada de Hall, 2012.

Os vetores transmissores da malária são as fêmeas dos mosquitos da ordem Díptera, da família Culicidae e do Gênero *Anopheles*. No Brasil, o principal anofelino transmissor da malária humana é o *Anopheles darlingii* que apresenta ampla dispersão no território nacional (presente em 80% do país), distribuindo-se por todo seu interior (Oliveira-Ferreira et al. 2010). A transmissão de *P. knowlesi* dá-se através da picada da fêmea do mosquito

Anopheles balabacensis. Este é um vetor importante da malária em humanos em áreas florestais do Sudeste Asiático, onde os hospedeiros naturais do *P. knowlesi* são macacos de cauda longa (*Macaca fascicularis*) (Bartoloni, Zammarchi, 2012).

A transmissão natural da malária é variável em intensidade e regularidade nas várias partes do mundo e depende de fatores locais como padrões climáticos principalmente períodos de chuvas, proximidade as fontes alimentares dos mosquitos transmissores e peculiaridade do comportamento das espécies de vetores que predominam em cada região. As fontes de infecção humana para os mosquitos são pessoas doentes ou até mesmo indivíduos assintomáticos que albergam formas sexuadas do parasita (de Oliveira, 2009).

Este protozoário tem ciclo de vida complexo que envolve reprodução sexual no mosquito vetor e estádios assexuais no hospedeiro vertebrado. Na malária humana, o hospedeiro é inoculado com esporozoítos através da picada do vetor durante seu repasto sanguíneo. Estes esporozoítos migram para o fígado, invadem células parenquimais e se diferenciam em merozoítos hepáticos. Estes se multiplicam, rompem as células e atingem a corrente sanguínea onde se ligam e invadem hemácias jovens. Uma subpopulação destes merozoítos desenvolve-se em gametócitos que são as formas que podem ser ingeridas na picada do mosquito. O ciclo de vida é completo pela reprodução sexual dentro do mosquito, seguido pelo desenvolvimento e invasão de esporozoítos haplóides em sua glândula salivar, da qual eles podem ser inoculados no hospedeiro vertebrado (Coatney et al., 1971).

Em *P. vivax*, assim como em *P. ovale*, pode haver a ocorrência de recaídas - reaparecimento da parasitemia, frequentemente com sintomas da doença (mesmo após tratamento eficaz) - de meses até mais de um ano depois do primeiro episódio de malária. Este fato deve-se à presença de hipnozoítos, formas presentes nas células do fígado que permanecem latentes até que sua divisão e reprodução se iniciem, gerando o quadro da doença. Entretanto, o motivo deste despertar dos hipnozoítos ainda não foi esclarecido.

1.3 Malária causada por *Plasmodium. vivax*

A maior parte da carga global da malária é causada por dois parasitos: *Plasmodium falciparum* que é mais letal, principalmente na África e *Plasmodium vivax* que tem tido um aumento no foco de pesquisas devido a sua maior distribuição geográfica (mais de 21 países do continente americano), maior exposição potencial e maior número de pessoas em risco (Guerra et al., 2010). Também por apresentar alto custo de tratamento assim como igual custo de diagnóstico, além de cursar com frequentes recaídas (Hay et al., 2004) .

O controle de *P. vivax* é potencialmente mais difícil do que o das demais espécies, principalmente *P. falciparum*, por motivos como: o surgimento precoce de gametócitos no sangue periférico (três dias após a infecção em *P. vivax*, enquanto que para *P. falciparum* são de oito a 10 dias), que torna possível a transmissão antes mesmo dos sintomas ocorrerem e o tratamento ser iniciado (Oliveira-Ferreira et al. 2010) e as dificuldades de prevenir recaídas, decorrentes da reativação dos hipnozoítos (Sattabongkot et al., 2003). Além disso, são comuns as infecções assintomáticas em populações endêmicas (Alves et al., 2002), reduzindo o impacto das medidas de controle baseadas exclusivamente no diagnóstico precoce e tratamento de infecções sintomáticas.

Apesar de ser classicamente descrita como uma doença benigna, com baixos níveis de mortalidade e morbidade, de tratamento relativamente simples, *P. vivax* possui uma grande diversidade fenotípica, especialmente em seus padrões de recaídas e apresenta um número crescente de infecções graves e potencialmente fatais (Antinori et al., 2012, Bartoloni, Zammarchi, 2012), além de apresentar resistência à cloroquina, o que sugere a aquisição de novas características fenotípicas por parasitos de diferentes regiões geográficas (Price et al., 2009).

1.4 Origem e características genéticas de *Plasmodium vivax*

Um estudo de Mu e colaboradores em 2005, realizado com 176 sequências de genoma mitocondrial de *P. vivax* de diversas localidades geográficas sugere que *P. vivax* originou-se no Sudeste Asiático e dispersou-se ao leste pela Índia e África. Este estudo foi corroborado por outro estudo de filogenia de genoma mitocondrial completo de Cornejo e Escalante em 2006 que combinou estes dados com mais 105 sequências de genoma mitocondrial obtidas do trabalho de Jongwutiwes e colaboradores no ano de 2005. Em todos estes estudos os parasitos do Sudeste Asiático apresentaram-se mais polimórficos, seguidos pelos da Melanésia, subcontinente Indiano, África e América do Sul respectivamente. Populações mais polimórficas são consideradas mais antigas, portanto admite-se que a invasão por *P. vivax* deu-se de leste a oeste de parasitos oriundos do Sudeste Asiático. No entanto, como “mais antigo” não é sinônimo de “aquele que dá origem a”, Culleton e colaboradores fizeram em 2011 uma análise genética de 23 genomas mitocondriais de isolados de diferentes localidades e este estudo revelou que a diversidade haplotípica diminui relativamente de forma concordante com estudos anteriores, do Leste da Ásia, para a Índia, para a América do Sul para o Norte da África e finalmente para África Sub-Saariana.

Ainda segundo Cornejo e colaboradores, em estudos filogenéticos feitos em 2006, *Plasmodium vivax* divergiu de *Plasmodium cynomolgi*, sua espécie irmã, oriunda de primatas não humanos (macacos do Sudeste Asiático). Este parentesco originou especulações que *P. vivax* emergiu como um parasito humano nesta região.

A primeira tentativa de conhecer o genoma de *P. vivax* foi feita com a criação de uma biblioteca de YACs (cromossomos artificiais de leveduras- do Inglês- “*Yeast Artificial Chromosome*”) disponível em Poptools e representativa de todo o genoma de *P. vivax* (Camargo et al., 1997). Somente em 2008 foi publicada por Carlton e colaboradores a sequência genômica de *P. vivax* a partir do método de *shotgun* do genoma total do isolado Salvador I, originário de El Salvador.

Em 2010, Dharia e colaboradores submeteram ao sequenciamento completo um isolado de campo de um paciente febril do Peru, denominado IQ07 e em 2012, Neafsey e colaboradores sequenciaram quatro genomas de *P. vivax* de localidades distintas a fim de proporcionar o primeiro estudo comparativo de variabilidade dentro desta espécie.

Os parasitos da malária no estágio sanguíneo, nas células do hospedeiro, são haplóides (Imwong et al., 2012). O genoma de *P. vivax*, como o genoma de todos os plasmódios estudados até o momento, está distribuído entre seus 14 cromossomos haplóides e 5.500 genes (Su, et al., 2009), consiste de 23 megabases (Mb) , com genoma mitocondrial (seis kilobases) e genoma de plastídeo (35 kilobases) (Conway, 2007) porém possui um conteúdo menor de AT (aproximadamente 57%) em comparação ao genoma de *P. falciparum* (aproximadamente 81%) (Neafsey et al., 2012). A conservação da sintenia, propriedade de dois ou mais genes estarem localizados no mesmo cromossomo em espécies relacionadas, é extensiva entre todas as espécies de *Plasmodium* e é reduzida quando as distâncias evolutivas entre as espécies aumentam (Feng et al., 2003).

1.5 A malária causada por *Plasmodium vivax* como problema de saúde pública

1.5.1 No Brasil

A maior parte das evidências disponíveis até o momento sugerem que a malária não estava presente no novo mundo antes da colonização européia em 1492, entretanto a introdução da malária nas Américas ainda é controversa. As linhas de pensamento estão baseadas em recentes descobertas sobre a evolução dos parasitos da malária em primatas, nas limitadas mutações genéticas causadas pela doença, na conexão entre grupos sanguíneos e malária, nas condições ambientais pouco favoráveis à transmissão da malária durante os movimentos migratórios através da Sibéria e do Alasca pelos primeiros ocupantes do continente americano e especialmente na ausência de relatos históricos de civilizações pré Colombianas (de Castro, Singer, 2004).

No entanto, Cormier em 2010 publicou um estudo que em parte concorda com a possibilidade da malária ter sido trazida ao Novo Mundo durante a era da colonização por europeus, mas também propõe uma nova hipótese que é a da malária endêmica no Novo Mundo, na qual índios da região Amazônica teriam sido infectados por plasmódios oriundos de macacos do Novo Mundo. Neste cenário, algumas formas de malária teriam sido espalhadas para o Velho Mundo durante a colonização. Esta hipótese é baseada em três linhas de pensamento: a etnomedicina, as mudanças climáticas e a genética da malária neotropical.

Ainda Neafsey e colaboradores (2012) sugerem que a alta diversidade genética de *P. vivax* no Novo Mundo, combinado com a proximidade da afinidade filogenética de três isolados oriundos do Novo Mundo com os isolados do Leste Asiático ao invés de com os isolados da África ou do sul da Ásia, pode sugerir a chegada pré-colonial de *P. vivax* no Novo Mundo, acompanhando a dispersão humana da Ásia por mar ou menos provavelmente pelo estreito de Bering durante a última era glacial. Alternativamente, este perfil pode ser explicado pela introdução recente, porém em grande escala, introdução pós-colonial de *P. vivax* no Novo Mundo via transações transoceânicas do Leste da Ásia, Pacífico ou Europa, a última que presumidamente abrigou clados de parasitos geneticamente distintos antes da eliminação da doença.

Nosso país concentra o maior número de casos de malária do continente americano, mas é na região Amazônica que são adquiridas aproximadamente 99% das infecções registradas no Brasil (Oliveira-Ferreira et al., 2010).

Ainda segundo Oliveira-Ferreira e colaboradores, no trabalho de 2010, o cenário no qual a malária se encontrava na década de 1930 no Brasil pode ser descrito inicialmente quando o Nordeste do país foi invadido por *Anopheles gambiae* vindos de navio de Senegal e um grave surto de malária atingiu uma população sem imunidade, levando a 13% de fatalidade e

alarmando malariologistas e autoridades de saúde. Este surto foi controlado nos primeiros anos de 1940, com campanhas de controle de saúde pública. Mas já nesta década a malária era um problema nacional que atingia seis milhões de pessoas, aproximadamente 20% da população infectada por ano. Em 1950 uma campanha de sucesso nacional, objetivando a erradicação de vetores (baseada na pulverização de spray nas paredes das casas), levou a diminuição da malária a 36,9 mil casos registrados em 1960. Porém nem ações contra o vetor, nem o tratamento de casos febris com cloroquina, puderam conter o rápido avanço da doença na região Amazônica, na qual a malária ainda é um sério problema de saúde (Oliveira-Ferreira et al., 2010).

De meados de 1960 em diante, o Brasil presenciou um rápido e desorganizado processo de assentamento agrícola na Amazônia patrocinado pelo governo e este movimento migratório levou a um aumento progressivo de casos reportados (Oliveira-Ferreira et al., 2010).

Nos anos entre 1970 e 1985, no Brasil, a incidência anual de malária multiplicou-se por dez estabilizando-se na década seguinte em torno de 500.000 casos anuais (Loiola et al., 2002). No entanto, novos incrementos de incidência, registraram 1999 com mais 636.000 casos, 2005 com 600.000 casos (Ladislau et al., 2006) e 2007 com 458.042 casos clínicos confirmados de malária, a grande maioria provenientes da região Amazônica (cerca de 306 mil casos por ano), conhecida como Amazônia Legal, que compreende os estados do Amazonas, Pará, Acre, Roraima, Rondônia, Amapá, Mato Grosso, Tocantins e Maranhão, onde a transmissão da doença geralmente está relacionada as condições ambientais e socioculturais (Pan American Health Organization, 2008 ; Oliveira-Ferreira et al., 2010) ; SIVEP/ Ministério da Saúde, 2010).

Em 2009, dados do Programa Nacional de Controle da Malária revelaram que apenas três estados, Amazonas, Pará e Rondônia, registraram, aproximadamente, 240 mil casos. Porém, é na região extra-amazônica que observa-se a maior letalidade da malária, seja devido ao diagnóstico tardio, ao manejo inadequado dos casos esporádicos importados de áreas endêmicas, ou mesmo autóctones em poucos estados (SIVEP/ Ministério da Saúde, 2010).

Abaixo encontra-se o mapa da distribuição espacial dos casos da doença no país.

Figura 2- Distribuição dos casos de malária no Brasil.

Fonte: Organização Mundial de Saúde, 2011.

Dentre as três espécies de *Plasmodium* que são transmitidas no País, *Plasmodium vivax* é a espécie responsável pela maioria dos casos, embora *Plasmodium falciparum* e *Plasmodium malariae* também sejam encontradas, porém, em menores proporções (Oliveira-Ferreira et al., 2010). No Brasil, tanto a ocorrência de resistência à cloroquina (de Santana Filho, 2007) como a de doença grave (Andrade et al., 2010) estão bem caracterizadas .

Das décadas de 1960 a meados da década de 1980 a transmissão de *P. falciparum* e *P. vivax* era igualmente prevalente no Brasil (Loyola et al., 2002). A partir do fim da década de 80 até os dias de hoje nota-se também uma clara mudança na distribuição das espécies de plasmódios que causam malária havendo o predomínio de transmissão de malária causada por *P. vivax* 83,7% dos casos, contra 16,7 % de infecções causadas pelos demais plasmódios. *Plasmodium falciparum* é responsável pelas formas mais graves e letais da doença, assim como pelas internações, mas há um declínio destes números de 2,6% em 2006 para 1,3% em

2008 (Ministério da Saúde, 2010). A Figura 2 apresenta o número de casos de malária entre os anos de 1960 e 2007, assim como a distribuição de *P. falciparum* e *P. vivax* entre estes anos.

Figura 3- Número de episódios de malária confirmados registrados por ano no Brasil entre 1960 e 2007.

Fonte: Ministério da Saúde, 2010. (Dados não publicados)

1.5.2 No Mundo

Metade da população mundial é estimada de estar sob o risco da malária causada por *P. vivax* especialmente devido ao seu potencial de recaída e tolerância a climas mais amenos (Guerra et al. 2010). Isto significa em números que 100 milhões de pessoas são infectadas com *P. vivax* a cada ano e esta espécie é responsável por mais da metade dos casos de malária fora do território africano (Hall, 2012).

Em 2011, 106 países foram identificados como endêmicos, tendo transmissão de malária durante todo o ano, três países (Egito, Armênia e Turquemenistão) não registram mais casos de malária e passaram da fase de eliminação para prevenção da reintrodução da doença. Enquanto seis outros (Geórgia, Azerbaijão, Quirguistão, Tadjiquistão, Uzbequistão e Turquia) caminham para mudança de estágio de pré-eliminação para eliminação da malária (WHO, 2010).

Regiões tropicais e subtropicais estão mais susceptíveis ao impacto da malária na saúde e no desenvolvimento econômico da população. No Sudeste Asiático há o maior contingente de pessoas expostas ao risco de contrair esta malária (1,345 bilhões de pessoas), na região oeste do Pacífico, 890 milhões de pessoa estão expostas ao risco, na África, 50 milhões de pessoas e nas Américas do Sul e Central, é estimado que 78 milhões de pessoas estejam expostas, um número bem inferior às demais localidades (Guerra et al., 2010). Na África, a malária vivax é rara na sua porção Ocidental e não frequente na África Oriental (Carter, 2003) esta situação é atribuída especialmente à grande prevalência do fenótipo Duffy negativo (ausência de glicoproteínas transmembrânicas presentes em células epiteliais, endoteliais e eritrócitos, que são receptores de antígeno responsáveis pela entrada de *P. vivax* nas células do hospedeiro) (Culleton et al., 2011) .

Embora a malária por *Plasmodium vivax* seja relativamente negligenciada, ela é um grande problema de saúde pública para as Américas Central e do Sul, Oriente Médio, Centro, Sul e Sudeste da Ásia, Oceania e Leste da África, onde existem 2.85 bilhões de pessoas em risco de infecção, 70-80 milhões de casos clínicos reportados cada ano (de Brito e Ferreira, 2011, Guerra et al., 2010).

Na América Latina, a Região Amazônica concentra 90% dos casos da doença registrados anualmente, mas de acordo com a Organização Mundial de Saúde (do inglês *World Health Organization -WHO*) há malária nos oito demais países que são contíguos com a floresta Amazônica : Bolívia, Colômbia, Equador, Guiana Francesa, Guiana, Suriname, e Venezuela. Destes 25% dos casos de malária são causados por *P. falciparum* e 75% por *P. vivax*, apesar de *P. malariae* ocorrer nesta região, não está documentado o valor estatístico desta transmissão. A Figura 3 apresenta a distribuição da malária causada por *P. vivax* no mundo.

estimativas podem variar de acordo com marcadores genéticos e métodos para tipagem utilizados (Havryliuk et al., 2008).

Métodos de identificação de variação genética podem ser explorados epidemiologicamente para prover meios de identificação de parasitos que infectam indivíduos e segui-los pelo curso da infecção quando drogas e outras intervenções forem aplicadas. Um exemplo é na identificação de parasitos em um indivíduo anteriormente ao início do tratamento e também depois do fim do tratamento para reconhecer se a parasitemia pós-tratamento é resultado de falha de tratamento – recrudescência ou re-infecção por outra forma do parasito. Métodos de tipagem molecular são importantes para determinar ainda a eficácia de drogas em campo (Daniels et al., 2008).

1.7 Marcadores genéticos para estudos populacionais

A dificuldade de cultura *in vitro* de *P. vivax*, as baixas parasitemias encontradas nos isolados e a dificuldade de adaptação dos isolados de campo aos macacos são fatores que dificultam muito as pesquisas genômicas em *P. vivax* (Feng et al., 2003).

Outro grande obstáculo aos estudos de genética de populações de *P. vivax* foi, até recentemente, a falta de marcadores moleculares adequados para uso em larga escala. Os polimorfismos mais estudados nessa espécie referem-se a genes que codificam antígenos (Cui et al., 2003), que podem ser informativos sobre o grau de variação intra-específica, mas apresentam a desvantagem de estarem sob intensa pressão seletiva. A seleção natural produz um padrão de diversificação nestes genes que não reflete necessariamente o que ocorre no genoma como um todo (Escalante et al., 2004).

1.7.1 Marcadores Neutros

Como primeiro passo para preencher essa lacuna, centenas de marcadores de DNA (Ácido Desoxirribonucléico – do inglês “*Deoxyribonucleic acid*”) microsatélite (MSAT) e polimorfismos de base única (SNP do inglês Single Nucleotide Polymorphism) de *P. vivax*, foram descritos ao longo da última década. Estes marcadores são úteis em estudos populacionais de diversidade genética de eucariotos, em função de sua neutralidade em termos evolutivos e seu alto grau de polimorfismo (Schlotterer, 2004).

1.7.1.1 Microsatélites

Microsatélites (MSAT) são repetições em tandem (de dezenas a milhares) de motivos conservados ou repetitivos de um a seis nucleotídeos polimórficos. São os mais populares e informativos marcadores genéticos disponíveis (Guichoux et al., 2011), tem sido utilizados desde de 1989 para genotipagem (Ellegren, 2004; Schlötterer, 2004) e são amplamente utilizados para investigar a diversidade genética de parasitos de malária por sua abundância, altas taxas de mutação e provável evolução neutra. Representam regiões instáveis do genoma, que estão sob alterações mutacionais a taxas muito maiores que as observadas nas sequências de cópia única. É esta instabilidade dos MSAT que resulta em marcadores altamente polimórficos. Os alelos diferem no número de repetições em série devido ao *crossing over* desigual durante a meiose, e principalmente pelo deslizamento (*strand-slippage*) da DNA polimerase durante a replicação do DNA (Ellegren, 2004). Em relação a *P. falciparum*, *P. vivax* possui menor abundância de MSAT. Foram descritos aproximadamente 160 marcadores polimórficos em oito isolados de referência de *P. vivax*, a maioria deles do tipo trinucleotídeo, com aproximadamente 19 unidades repetitivas por marcador e um conteúdo médio de 27,5% de C-G (Feng et al., 2003; Carlton et al., 2008).

Em 2003 foi descrito o primeiro MSAT para *P. vivax* por Gomez e colaboradores, obtido a partir da análise da região subtelomérica de um cromossomo artificial de levedura YAC (YAC

do inglês *Yeast Artificial Chromosome*). A tipagem deste MSAT em 89 amostras de Papua Nova Guiné revelou uma população de *P. vivax* muito heterogênea. Em 2004, Leclerc e colaboradores obtiveram um resultado diferente com a tipagem de 13 marcadores de DNA microssatélite e minissatélite, obtidos a partir de uma biblioteca de DNA com sequências repetitivas em 108 isolados de diferentes regiões do mundo. Nove dos 13 marcadores apresentaram-se monomórficos e, dos quatro restantes, somente um deles foi altamente polimórfico. Estes resultados sugeriam inicialmente que os MSAT de *P. vivax* eram pouco abundantes e pouco polimórficos.

No entanto, estudos recentes (Imwong et al., 2006; Imwong et al., 2007; Ferreira et al., 2007; Karunaweera et al., 2008; Joy et al., 2008; Orjuela-Sánchez et al., 2009; Van den Eede et al., 2010; Gunawardena et al., 2010) com novos marcadores revelam extenso polimorfismo de *P. vivax* em diferentes regiões. O estudo de Gunawardena e colaboradores em 2010 com a tipagem de 425 isolados de *P. vivax* de Sri Lanka, Mianmar e Etiópia, utilizando o conjunto de marcadores descritos pelo mesmo grupo no ano de 2007 comprovou que estes MSAT são uma ferramenta eficaz para mapear isolados segundo sua ancestralidade e origem continental.

Em 2012 Imwong e colaboradores realizaram um estudo prospectivo de coorte na Tailândia com 21 crianças que adquiriram malária vivax no seu primeiro ano de vida, 12 de suas mães também apresentaram malária vivax durante a gravidez ou após o parto. Foram utilizados seis marcadores moleculares tipo microssatélites e como resultado foi observado que o genótipo dos parasitas eram todos diferentes entre mães e filhos e que quando havia uma recaída, com exceção da primeira recaída da vida, os parasitos encontrados eram distintos da primeira infecção, portanto oriundos de hipnozoítos heterólogos inoculados pelo vetor ou adquiridos de infecções prévias.

Embora os MSATs tenham permitido comparar parâmetros populacionais fundamentais, como níveis de diversidade e de desequilíbrio de ligação, em amostras de *P. vivax* da Ásia, África e das Américas, seu uso em larga escala tem sérias limitações. Em primeiro lugar, as

sequências de MSAT têm uma taxa de mutação extremamente elevada (da ordem de 10^{-2} eventos por replicação). A cada evento mutacional, um número variável de cópias do motivo repetitivo pode ser inserido ou excluído. Consequentemente, novos haplótipos são frequentemente introduzidos em populações não somente por migração de novas linhagens de parasitos por ou recombinação entre haplótipos existentes, mas também por mutação nos alelos que os constituem, dificultando as análises temporais. Como resultado das altas taxas de mutação, diferenças significantes entre isolados e populações reveladas por marcadores hipervariáveis nem sempre têm algum significado biológico (Hendrick, 1999). Além disso, MSATs são sujeitos a homoplasia (identidade de sequência na ausência de um ancestral comum), fato que viola os pressupostos de modelos clássicos de análise genética populacional. Finalmente, os métodos de tipagem de MSATs estão muito sujeitos a variação entre laboratórios, reduzindo as possibilidades de comparação de resultados obtidos por diferentes grupos de pesquisa com o mesmo conjunto de marcadores.

Além disto, existem algumas dificuldades para tipagem de microssatélites em infecções de múltiplos clones em malária. Alelos principais e secundários de microssatélites co-existem na mesma infecção e podem teoricamente ser diferenciados pela medida relativa da abundância dos respectivos produtos de PCR identificados com fluoróforos, depois da separação eletroforética por seqüenciadores automatizados de DNA. Esta estratégia permite a definição do haplótipo predominante que é a combinação única dos alelos mais abundantes em cada locus, analisados em misturas clonais, mas não validado para marcadores de microssatélites disponíveis para tipagem de *P. vivax* em larga escala. Além do mais, alelos secundários podem se confundidos com artefatos ou produtos inespecíficos de PCR (Havryliuk et al., 2008).

Ainda há outras características tornam a utilização de microssatélites para genotipagem, desvantajosa em comparação aos SNPs (do inglês *Single Nucleotide Polymorphisms*), por exemplo: a) um grande número de alelos por locus em microssatélites implica que para uma acurada estimativa de frequência alélica, torna-se necessário um grande número de amostras; b) maior número de mutações espontâneas fazem os microssatélites serem fracos

indicadores de história populacional de longo prazo; c) a variabilidade entre os microssatélites altamente polimórficos pode não refletir exatamente a diversidade genômica subjacente; d) as poucas formas de detecção comumente utilizadas para análise de microssatélites (eletroforese por capilaridade associada à detecção baseada na emissão de fluorescência), enquanto que os SNPs podem ser tipados por diferentes técnicas, que permitem o desenvolvimento de métodos de densidade ultra altos; e) finalmente, com o uso de microssatélites existe a necessidade de inclusão de controles comuns entre os estudos e através do tempo, em contrapartida os SNPs podem ser replicados, realizados em diferentes laboratórios e podem ter novas amostras adicionadas ao estudo sempre que as mesmas surgirem, sem ter que haver uma padronização nova de cada passo. Até o presente momento, a reduzida portabilidade de dados de microssatélites entre laboratórios tem resultado em significativa limitação de uso de resultados (Revisado em Guichoux et al., 2011).

1.7.1.2 Polimorfismos de Base Única (SNPs)

Os polimorfismos de base única (do inglês *Single Nucleotide Polymorphisms*) são variações genéticas que correspondem às alterações mais elementares da molécula do DNA, ou seja, mutações de bases únicas nitrogenadas (Adenina-A, Citosina-C, Timina-T e Guanina-G). As mutações mais comuns são conhecidas como transições, nas quais ocorrem trocas de uma purina por outra purina ($A \rightarrow G$) ou de uma pirimidina por outra pirimidina ($C \rightarrow T$). As transversões são menos frequentes e ocorrem quando há troca de uma purina por uma pirimidina, ou vice-versa ($C-T \rightarrow A-G$). Normalmente os SNPs são bi-alélicos, mas podem ser também tri-alélicos e podem ocorrer em regiões codificadoras ou com função regulatória. No entanto, em sua maioria são encontrados em espaços intergênicos, sem função conhecida.

De 2000 em diante os SNPs tornaram-se os marcadores genéticos de escolha devido ao seu alto rendimento nos ensaios e tipagem mais reprodutível com o uso de diversas técnicas de alto rendimento (Feng et al., 2003; Neafsey et al., 2008; Mu et al., 2010). Os SNPs

representam um marcador genético mais estável, com taxa de mutação da ordem de 10^{-9} eventos/replicação. A genotipagem por SNP é baseada na substituição de nucleotídeos e pode ser utilizada em *Plasmodium vivax*. SNPs são abundantes em *Plasmodium falciparum* assim como em *P. vivax*, mas sua distribuição pode variar entre os vários cromossomos e diferentes genes (Su et al., 2009).

O estudo de Mu e colaboradores em 2002 analisou 204 genes do cromossomo três de *P. falciparum* e identificou aproximadamente um SNP para cada 900 bp ao comparar sequências de cinco isolados, a maioria encontrados em genes que codificam antígenos e possíveis transportadores ao invés de em genes que codificam proteínas essenciais ao metabolismo do parasito. Para genes *housekeeping*, a frequência de SNPs pode ser muito menor (Su et al., 2009). Em estudos de sequenciamento em larga escala, como o de Jeffares et al, (2007) e Volkman et al, (2007) sugerem que a frequência de SNPs pode ser muito maior, por exemplo, de 46,937 SNPs.

Em *P. vivax*, a análise comparativa de 100 kb de sequência de DNA do cromossomo 8 em cinco isolados naturais do parasita, identificou em média um SNP a cada 530 pb de DNA (Feng et al., 2003). Esta abundância de SNPs torna possível o desenvolvimento de mapas gênicos de alta resolução dessa espécie, facilitando os estudos de associação, ao nível genômico, entre mutações e genótipos de interesse em populações de parasitas. A Figura 5 ilustra o resultado deste trabalho apresentando os haplótipos de SNPs dos cinco isolados de *P. vivax* utilizados.

Figura 5- Haplótipos de SNPs para os cinco isolados de *P. vivax*. Azul: G, verde A, vermelho T e amarelo C. Blocos de haplótipos conservados estão marcados com barras pretas.

Fonte: Modificado de Feng *et al.*, 2003

Neafsey e colaboradores em 2008 demonstraram a eficácia da utilização de microarranjos de DNA contendo 3.000 SNPs para a tipagem de genoma rico em A-T de *P. falciparum*. Um total de 2.153 dos SNPs tipados encontram-se no cromossomo 7 que possui o gene PfCRT (*P. falciparum Chloroquine Resistance Transporter*), responsável pela resistência à cloroquina e os demais SNPs localizam-se nos 13 cromossomos restantes. A tipagem destes SNPs nos 76 isolados desta espécie de diferentes localidades confirmou níveis variáveis de diversidade genética e padrões de desequilíbrio de ligação atribuídos à origem geográfica dos isolados.

Em nosso laboratório, a análise de 85 SNPs ao longo do cromossomo 8 em 240 isolados de *P. vivax* provenientes de diversas áreas geográficas permitiu comparar as taxas de recombinação efetiva entre populações de parasitas com diferentes níveis de endemicidade e caracterizar segmentos cromossômicos com alta atividade recombinatória (Orjuela-Sánchez *et al.*, 2010). Entretanto, os padrões observados dizem respeito a um único segmento de somente um dos 14 cromossomos do parasito, deixando de fornecer um panorama mais geral sobre variação genética em nível genômico.

Dharia e colaboradores (2010) realizaram um estudo que propunha que a análise do genoma de *P. vivax* poderia ser feita diretamente através de parasitos isolados *ex vivo*, sem a necessidade de propagação *in vitro*. Para estes pesquisadores o maior objetivo de

sequenciar isolados de *P. vivax* é encontrar SNPs que possam ser utilizados para revelar genes sob pressão imune ou de drogas ou para encontrar marcadores para estudos de genética de populações. Para tanto, foi utilizado uma amostra de DNA de um paciente com malária do Peru, denominada IQ07. Este DNA foi sequenciado com a plataforma Illumina com uma cobertura de 30X e sua análise revelou 18.261 SNPs, 0,81 SNP para cada 1000 bases nos principais genes cromossômicos quando comparado ao isolado de referência Salvador I. Os resultados deste trabalho foram comparados a outra técnica de varredura genômica por hibridação, que utiliza cobertura por microarranjos (pequenas sondas derivadas da sequencia caracterizada no isolado Sal-1. Descrita por Westenberger *et al.* (2010), esta é uma técnica fácil de ser utilizada e mais barata que a técnica de sequenciamento de genoma completo. Para ter acesso a diversidade pelos microarranjos foram utilizados os genomas de Salvador I e IQ07. As sondas utilizadas tinham seis pares de bases distantes 2-3 nucleotídeos umas das outras. Como resultado a técnica de hibridização revelou menos SNPs que o sequenciamento do genoma completo (somente 8.118 dos 18.261) mostrando-se, portanto uma técnica menos eficaz e confiável.

Um estudo realizado na Índia por Gupta e colaboradores (2010) selecionou aproximadamente 200kb de DNA de uma região localizada no cromossomo 13 de *P. vivax* (sintênico ao cromossomo 14 de *P. falciparum*) que contém dois dos mais importantes genes de sobrevivência de parasito, protease aspártica (PvPM4) e proteína de circumsporozoíto (CSP) para identificação de SNPs. Um total de 17 fragmentos (10 intergênicos e sete íntrons) foi amplificado e sequenciados, num total de 7012 pares de bases. Posteriormente o sequenciamento de DNA e a análise das sequências confirmaram 11 fragmentos polimórficos (sete intergênicos e quatro íntrons) em um total de 18 SNPs, que revelaram que, em geral, íntrons são menos variáveis que regiões intergênicas.

O trabalho mais recente sobre este tema foi publicado por Neafsey e colaboradores em 2012 demonstrando que os parasitos de *P. vivax* possuem diversidade genética maior que os de *P. falciparum*. Para tal foram sequenciados em Plataforma Illumina, quatro isolados: um do Brasil, um da Coréia do Norte, um da Índia e um da Mauritânia- África. Foi observado o

dobro de diversidade de SNPs em todo o genoma dos isolados de *P. vivax* analisados em comparação aos isolados de *P. falciparum*, (Figura 7) confirmando resultados de outros estudos (Feng et al., 2003; Mu et al., 2005). Também foi estimada a ação da pressão seletiva em diferentes genes, como resultado, genes com maior razão dn/ds (razão de polimorfismos não-sinônimos por sinônimos) são aqueles genes mais provavelmente envolvidos na interação parasito-hospedeiro. Outro dado importante deste estudo é o alto grau de diversidade dentre os três isolados de *P. vivax* do Novo Mundo: o isolado de referência Salvador I e os demais IQ07 e Brasil I, como ilustrado na Figura 6, sugerindo uma história distinta de chegada destes parasitos no Novo Mundo se comparado a *P. falciparum*, que exibe baixa diversidade nos isolados do Novo Mundo e é considerado ter sido introduzido no Novo Mundo nos últimos 500 anos com o comércio de escravos africanos. Este dado também nega que um possível efeito gargalo tenha acontecido a *P. vivax*, assim como o que aconteceu recentemente a *P. falciparum*.

Figura 6- Resultados da diversidade de SNPs dentre as espécies de *Plasmodium*. A cor vermelha representa *P. falciparum* e a azul, *P. vivax*. O eixo x representa: 1- taxas de sequências codificadoras (CDS), 2- genoma completo, 3- sequências flanqueadoras 3', 4-sequências flanqueadoras 5', 5- todas sequências intergênicas, 6- introns e 7- quadras degeneradas de sites codificadores sinônimos (4D).

Fonte: Adaptado de Neafsey et al., 2012.

Figura 7- Filogramas *Neighbor-joining* de *P. vivax* e *P. falciparum* construídas a partir de SNPs presumidamente neutros que ocorrem em quadras degeneradas de sites codificadores sinônimos. Linhagens estão coloridas de acordo com a origem geográfica. O comprimento das linhas indica a diversidade considerável de *P. vivax* no Novo Mundo e não mostram associação clara entre os isolados do Novo mundo e o da África. Filogramas foram construídos de 471.543 sites em *P. vivax* e de 359.901 sites em *P. falciparum*.

Fonte: Neafsey et al., 2012.

1.7.1.3 Tipagem de SNPs com sondas TaqMan®

Estudos recentes desenvolveram um novo método de genotipagem baseado em SNPs com ensaios práticos, rápidos e de baixo custo, identificando qualquer parasita de *Plasmodium* através de uma pequena quantidade de DNA. Esses ensaios, conhecidos como TaqMan®, podem ser utilizados para distinguir recrudescência de re-infecção, em teste com drogas, para monitorar a frequência e distribuição de parasitas específicos em populações que fazem tratamentos com drogas, ou para identificar amostras e determinar a pureza de isolados de laboratório (Daniels et al., 2008). Ensaios do tipo TaqMan® utilizam sondas de hidrólise designadas para aumentar a especificidade do PCR. Este método foi primeiramente reportado em 1991 por pesquisadores Cetus Corporation e a tecnologia foi subsequentemente desenvolvida pela Applied Biosystems para aplicações na pesquisa. O

princípio desta técnica baseia-se na atividade 5'-3' exonuclease da Taq polymerase para separar uma sonda duplamente marcada durante a hibridação para a sequência alvo complementar e a detecção do fluoróforo. Como em outros PCRs Real time, o resultado do sinal da fluorescência permite mensurar quantitativamente a acumulação de produtos durante os estágios exponenciais do PCR como ilustra a Figura 8.

Figura 8- Resultado do funcionamento das sondas do tipo TaqMan®.

Fonte: Livack *et al.*, 1995.

Wilson e colaboradores (2005) determinaram, por meio de ensaios com sondas TaqMan, a prevalência de mutações K76T do gene PfCRT (do inglês *P. falciparum* Chloroquine Resistance Repórter Gene) em isolados de *P. falciparum* de pacientes da República Democrática do Congo (RDC) e de Maláui. Esta mutação foi encontrada em 93% dos isolados clínicos da RDC onde a resistência à cloroquina é alta e nenhuma mutação foi encontrada nos isolados de Maláui onde a cloroquina agora é pouco utilizada no tratamento da malária.

Daniels e colaboradores (2008) descreveram o primeiro ensaio de tipagem molecular para *P.*

falciparum utilizando sondas MGB-TaqMan (Minor Groove Binder). Eles selecionaram dos 112.000 SNPs descritos de 18 isolados sequenciados de *P. falciparum*, 2.100 SNPs os quais eram comuns nos isolados, se segregavam independentemente e estavam amplamente distribuídos ao longo do genoma. Foram inicialmente desenvolvidos 88 ensaios com sondas TaqMan, dos quais apenas 24 foram escolhidos para serem utilizados na identificação e localização geográfica de parasitos de *P. falciparum*. Esta técnica pode ser aplicada para tipar *locis* genéticos importantes na resistência de drogas, invasão, virulência e imunidade e pode ser utilizada a partir de isolados de pacientes quanto de isolados de cultura de laboratório.

Divis e colaboradores (2010) e Chaudry et al. (2012) padronizaram ensaios com sondas do tipo TaqMan para detectar diferentes tipos de plasmódios em isolados de campo. No estudo de Divis foram realizadas duas reações de PCR separadas, a primeira para selecionar todos os plasmódios e a segunda com sondas específicas para detectar *P. Knowlesi*. Chaudry e colaboradores em 2012 utilizaram um total de 60 amostras de pacientes com suspeita de malária em papel filtro e compararam resultados de microscopia com aqueles provenientes de PCR em tempo real para o gene da subunidade menor (SSU) de RNA ribossômico (sRNA), gene encontrado em todos os parasitos de malária. A sensibilidade e a especificidade dos dois métodos foram comparadas, e a conclusão é de que o PCR com sondas TaqMan é mais sensível e mais específico ao detectar 33 amostras enquanto 30 foram detectadas pela microscopia.

No ano de 2011 Muhamad e colaboradores utilizaram os ensaios com sondas do tipo TaqMan® (método padrão) para comparar resultados de ensaios com Sybr Green I na determinação e amplificação do número de cópias do gene PfMDR-1 (do inglês *P. falciparum* *Multi Drug Resistance Gene*) em 96 isolados de quatro áreas endêmicas da Tailândia. Como resultado foi encontrado que a eficiência de ambos os ensaios é muito similar na sensibilidade e especificidade.

1.7.2 Marcadores moleculares sob intensa pressão seletiva

Genes sob pressão seletiva são aqueles sujeitos à seleção pelo sistema imune do hospedeiro, e o uso de antimaláricos, não são assim, considerados os melhores em estudos populacionais, no entanto são os únicos sequenciados e disponíveis em bancos de dados como o GeneBank.

O gene associado à resistência a drogas utilizados neste estudo foi o PvMRD-1 (*Plasmodium vivax* Multidrug-Resistance Gene -1), ortólogo do gene PfMDR-1 associado a resistência a cloroquina em *P. falciparum*. Os genes que codificam antígenos disponíveis para tipagem de *P. vivax* são o antígeno de membrana apical 1 (AMA-1) aqueles que codificam para as proteínas de superfície de merozoítos 1 e 3, 4 e 5 (MSP-1, MSP-3, MSP-4 e MSP-5) e a proteína de ligação ao eritrócito DBP (Duffy Binding Protein).

1.7.2.1 PvMDR-1

Homólogo a PfMDR-1 PvMDR-1 é composto por um único éxon de 4.395 pb que codifica para uma proteína de 465 aminoácidos. PvMDR-1 possui dois domínios transmembrânicos hidrofóbicos que atravessam a membrana múltiplas vezes. Cada domínio consiste em seis segmentos transmembrana (alfa-hélice putativas) com um total de 12 segmentos por molécula. Os outros dois domínios são hidrofóbicos e possuem um domínio de ligação a nucleotídeo que acopla a energia da hidrólise de adenosina trifosfato (ATP) ao processo de transporte. Todas estas características estruturais caracterizam PvMDR-1 como membro da superfamília de proteínas transportadores ABC (Brega et al., 2005 e Sá et al., 2005). No ano de 2005, Sá e colaboradores sequenciaram o gene PvMDR-1 em dez isolados de diferentes regiões do mundo, incluindo amostras comprovadamente resistentes à cloroquina. Estes resultados evidenciaram que estas amostras poderiam ser agrupadas de acordo com sua origem geográfica, porém não é possível uma associação do fenótipo a esta característica.

1.7.2.2 AMA-1

AMA-1 (Apical Membrane Antigen), antígeno de membrana apical, é um dos mais promissores antígenos candidatos à vacina contra malária. É uma proteína integral de membrana tipo 1 altamente conservada, com ortólogos em pelo menos duas espécies do filo Apicomplexa, *Babesia bovi* e *Toxoplasma gondii*. Em parasitos de estádios sanguíneos de *Plasmodium*, AMA-1 está inicialmente presente em micronemas e posteriormente na superfície de merozoítos maduros. Esta proteína contém um ectodomínio rico em cisteína, um domínio transmembrânico e uma região citoplasmática. A função biológica precisa de AMA-1 ainda não foi elucidada, no entanto várias linhas de estudo sugerem seu envolvimento na invasão de eritrócitos por plasmódios, durante o qual ela formaria um complexo com outras proteínas uma junção móvel entre o parasito e as membranas das células do hospedeiro (Gentil et al., 2010)

1.7.2.3 MSP-1

Proteínas de superfície de merozoítos MSP-1 (Merozoite Surface Protein) incluindo MSP-1, MSP-2 e MSP-3 têm um importante papel na fase de invasão de eritrócitos. Estas proteínas também são estudadas como potenciais candidatas a alvos de desenvolvimento de vacina. A MSP-1 de *Plasmodium vivax* (PvMSP-1) é expressa como uma proteína de 200kDa na superfície do merozoíto e já foi clonada e sequenciada. A estrutura primária de PvMSP-1 foi originalmente caracterizada nos isolado Belém e Salvador-1 (Sal-1); o gene apresenta regiões conservadas, semi-conservadas e polimórficas. Comparações de sequencias indicam a conservação de 10 regiões entre as espécies de *Plasmodium*. Esta natureza altamente polimórfica pode indicar uma nova utilidade para MSP-1 como uma potencial alternativa de marcador molecular epidemiológico para *P.vivax* (Hwang et al.,2009).

1.7.2.4 MSP-3

A MSP- 3 de *Plasmodium vivax* MSP-3 (PvMSP-3) é uma proteína de 150 kDa que

compreende três domínios característicos: um domínio central rico em alanina, um domínio altamente conservado C- e uma região N-terminal. A família MSP-3 consiste de genes intra espécie dentre eles: PvMSP-3 α , PvMSP-3 β , e PvMSP-3 γ que exibem um alto grau de diversidade, o que os torna úteis marcadores polimórficos de áreas endêmicas e potenciais candidatos a vacina contra malária (Hwang et al.,2009).

1.7.2.5 MSP-4 e MSP-5

MSP-4 é homólogo a MSP-5, contém 230 aminoácidos, enquanto MSP-5 contém 387. Sua massa molar é 23.3 kDa e a de MSP-5 de 41kDa. Estão associados aos micronemas ou distribuídos sobre o *polar cap*, uma organela apical que é crucial para a entrada do merozoíto na célula do hospedeiro. O gene que codifica MSP-5 é caracterizado por duas estruturas exon de 861bp e 300 bp separados por um íntron de 402bp. O locus de MSP -5 e do seu gene parálogo, MSP-4 está arranjado em tandem no cromossomo quatro. A porção C-terminal de MSP-5 possui um domínio de fator de crescimento epidermal (EGF) e uma sequência de âncora de glicosil fosfatidil inositol (GPI) que também é encontrada em MSP-1 de *P. vivax* e *P. falciparum* (Putaporntip et al., 2010).

MSP-5 é antígeno candidato a vacina, recentemente utilizado em triagens com humanos e é relativamente conservado em isolado de *Plasmodium falciparum* e altamente polimórfica em isolados de *P.vivax*. Sua natureza relativamente conservada sugere que respostas imunes com alvo nesta região podem conferir proteção cruzada. MSP-5 é expresso não só em merozoítos como em esporozoítos e em hepatócitos infectados e é, portanto um alvo potencial para uma vacina combinada de estágio sanguíneo e hepático (Woodberry et al.,2008).

1.7.2.6 DBP

A proteína de ligação Duffy de *Plasmodium vivax* (Duffy Binding Protein – PvDBP), uma das proteínas de ligação de eritrócitos com 140kDa, contém duas regiões funcionais conservadas ricas em cisteína, regiões II e IV. A natureza polimórfica deste antígeno, particularmente na região II (aminoácidos 291-515), que contém o domínio central de ligação necessário para aderência ao eritrócito é o maior impedimento ao sucesso na produção de uma vacina contra malária vivax. A maior parte das sequências de aminoácidos da região II são bem conservadas e mostram um alto grau de similaridade sequência com os isolados Belem e Salvador-1 (Hwang et al.,2009).

2.1 Padronizar uma estratégia simples e reprodutível de caracterização genética de isolados de campo de *P. vivax*. Esta estratégia baseia-se na tipagem molecular de um conjunto de polimorfismos de base única (SNPs), predominantemente substituições sinônimas em regiões codificadoras ou substituições de nucleotídeos observadas em regiões intergênicas ou introns.

2.2 Comparar os resultados preliminares obtidos com a tipagem de MSATs com aqueles obtidos com a análise de SNPs nos mesmos isolados.

3. Material e Métodos

3.1 Aspectos éticos

Os protocolos das pesquisas de campo que originaram (Granada e Plácido de Castro) as amostras analisadas neste projeto foram avaliados e aprovados pela Comissão de Ética em Experimentação com Seres Humanos do Instituto de Ciências Biomédicas da USP. Os protocolos de pesquisa desta dissertação foram também analisados e aprovados pela Comissão de Ética em Experimentação com Seres Humanos do Instituto de Ciências Biomédicas (ICB) da Universidade de São Paulo (USP) (parecer 936/CEP de 14 de outubro de 2010).

3.2 Amostras de DNA

As amostras sanguíneas de origem brasileira utilizadas neste estudo foram coletas de pacientes com infecção por *P. vivax*, moradores do assentamento agrícola conhecido como Granada, e do município de Plácido de Castro, no estado do Acre. Estas populações foram estudadas em 2004-2005 e 2008 respectivamente. Outras amostras de DNA utilizadas neste estudo são provenientes da Ásia: (a) da cidade de Pursat, no Camboja, colhidas de indivíduos que contraíram malária vivax em áreas florestais próximas à cidade, no ano de 2008 e (b) de Tricomalee, Sri Lanka, colhidas durante um surto de malária vivax em 2007. Os demais isolados são provenientes do CDC (do inglês *Centers for Disease Control and Prevention*), de amostras de pacientes com infecção por malária vivax importada, diagnosticadas no próprio CDC nos Estados Unidos durante os anos de 2004 a 2008.

O diagnóstico da espécie infectante das amostras de sangue provenientes do Brasil, utilizadas neste estudo, foi primeiramente realizado por meio de microscopia convencional de gota espessa corada com Giemsa, posteriormente confirmado por meio de PCR (do inglês *Polymerase Chain Reaction*) aninhada pelo método de Kimura (Kimura et al., 1997) com o alvo o gene 18S rRNA (Ferreira et al., 2007). Todas as amostras de sangue coletadas no Brasil por punção venosa foram mantidas a - 20° C até a extração de DNA.

Para obter o DNA das amostras foram coletados 5mL de sangue venoso de cada paciente, após um consentimento por escrito e a partir de 200 µL deste material, foi extraído o DNA template com o auxílio de kits comerciais de extração de DNA como GFX (Amersham Pharmacia Biotec) e Wisard (Promega, Madison, WI), seguindo as instruções do fabricante, além do protocolo fenol-clorofórmio, descrito por Moll e colaboradores (2008).

Para obter a concentração adequada e o volume necessário de DNA para os procedimentos realizados neste estudo, PCR em tempo real para obtenção da parasitemia dos isolados e tipagem molecular por ensaios tipo TaqMan[®], todas as amostras utilizadas neste estudo tiveram seu DNA pré-amplificado através de uma técnica denominada WGA (do inglês *Whole-Genome Amplification*). Esta técnica tem como finalidade a amplificação total do genoma por deslocamento múltiplo utilizando uma DNA polimerase capaz de replicar até 100 kb sem dissociar-se do DNA molde. Esta enzima possui atividade *proof-reading* (revisão exonuclease 3'→5') garantindo a fidelidade dos produtos durante a replicação. O WGA foi feito utilizando 10 ng de DNA genômico, com o Kit a REPLI-g Minikit (Qiagen, Valencia, EUA) seguindo as instruções do fabricante. De acordo com estudo de Orjuela-Sánchez e colaboradores (2010) nenhum alelo minoritário ou majoritário é modificado ou perdido durante o procedimento do WGA, portanto a utilização desta técnica não gera vieses ao processo de tipagem molecular. Além deste procedimento, foram realizados testes para confirmar que todos os marcadores utilizados neste trabalho são espécie-específicos de *P. vivax* e não apresentam amplificação cruzada com DNA humano.

3. 3 Áreas de estudo

3.3.1 Granada

Parte das amostras de *P. vivax* utilizadas neste estudo foi coletada entre março de 2004 e dezembro de 2005, em um assentamento agrícola conhecido como Granada, no extremo

leste do estado do Acre, Amazônia Ocidental brasileira (da Silva-Nunes *et al.*, 2008), situado a 30-50 Km noroeste de Acrelândia (Figura 9).

O Granada é parte da fronteira colonizada conhecida como Pedro Peixoto Colono, teve seu loteamento inicial realizado pelo Instituto Nacional de Colonização e Reforma Agrária (INCRA) em 1982, época em que chegaram as primeiras famílias de assentados. Trata-se de uma área de população estável, dedicada à agricultura comercial e de subsistência e à pecuária, em que os níveis de transmissão de malária se mantiveram em patamares relativamente baixos durante o seguimento longitudinal (30 casos de malária vivax/100 pessoas-ano e 16,3 casos de malária falciparum/100 pessoas-ano) (da Silva-Nunes *et al.*, 2008). As amostras de *P. vivax* provenientes do Granada foram obtidas no contexto de dois estudos interligados. O primeiro foi um estudo de coorte prospectiva de 509 indivíduos, que envolveu busca ativa e passiva de casos, realizado entre 2004 e 2005, durante o qual foram colhidas amostras de parasitos de casos incidentes de malária vivax que tiveram o diagnóstico inicial, feito por microscopia convencional, confirmado por PCR (Ferreira *et al.*, 2007; da Silva-Nunes *et al.*, 2008). O segundo estudo, realizado na mesma área entre 2005 e 2007, consistiu no seguimento, pelo período mínimo de seis meses, de uma coorte de 77 indivíduos com malária vivax tratados com o esquema convencional de cloroquina-primaquina (coorte B descrita por Orjuela-Sánchez *et al.*, 2009). Todas essas amostras tiveram marcadores de DNA microssatélite previamente tipados (Orjuela-Sánchez *et al.*, 2009).

3.3.2 Plácido de Castro

Criado em 1976, ganhou este nome em homenagem a José Plácido de Castro, militar, líder da Revolução Acreana e posterior Presidente do Estado Independente do Acre. Sua população em 2007 era de 17.258 habitantes e sua área é de cerca de 2.047 km² (8,4 hab./km²). A sede do município localiza-se a cerca de cem quilômetros da capital. Plácido de Castro está a uma distância de 50 km de Granada e limita ao norte com o município de Senador Guimard, ao sul com a Bolívia, a leste com o município de Acrelândia, a oeste com

o município de Capixaba e a noroeste com o município de Rio Branco (Orjuela-Sánchez et al., 2010) (Figura 9). Este município apresenta transmissão de malária hipoendêmica durante todo o ano. As amostras deste estudo foram colhidas no ano de 2008 de pacientes que receberam atendimento na clínica de malária da cidade (Orjuela-Sánchez et al., 2010). Todas essas amostras tiveram marcadores de DNA microsatélite previamente tipados (Orjuela-Sánchez et al., 2010).

Figura 9- Mapa dos Estados do Acre e Amazonas, noroeste do Brasil mostrando a localização das áreas de estudo, Granada, Plácido de Castro e Remansinho.

3.3.3 Pursat, Camboja

As amostras oriundas do Camboja utilizadas neste estudo (Figura 10) foram coletadas entre junho de dezembro de 2008 em Pursat, estas amostras são de indivíduos que contraíram malária vivax em áreas florestais próximas à cidade (ClinicalTrials.gov identifier, NCT00663546). A infecção por *P. vivax* foi confirmada por microscopia

convencional de gota espessa corada com Giemsa e por PCR com alvo o gene que codifica a subunidade menor (18S) do RNA ribossômico (18S rRNA) (Snounou et al., 1993). Essas amostras foram utilizadas para a tipagem de microssatélites (Orjuela-Sánchez et al., artigo submetido¹); e para a análise de padrões regionais de distribuição dos polimorfismos descritos neste estudo.

3.3.4 Tricomalee, Sri Lanka

As amostras provenientes de Sri Lanka (Figura 10) foram coletadas durante um surto de malária em Tricomalee, Província do Leste. Entre Janeiro e Agosto de 2007, um total de 87 casos de malária vivax foi relatado (dados não publicados). Nos últimos anos, a transmissão de malária tem gradualmente diminuído neste país, de 210.039 casos de malária em 2001 para 196 casos em 2007. Em contraste, nenhuma redução drástica na transmissão da malária foi documentada na última década para os demais locais na Ásia, incluindo no Camboja. (Orjuela-Sanchez et al., 2010). Essas amostras foram gentilmente cedidas por Nadira Karunaweera (University of Colombo, Sri Lanka) para este estudo. Todas essas amostras tiveram marcadores de DNA microssatélite previamente tipados (Gunawardena et al., 2010).

1 Orjuela-Sanchez, P.; Sá, J. M.; Brandi, M. C. C.; Rodrigues, P. R.; Bastos M. S.; Amaratunga, C.; Duong, S.; Fairhurst, R. S.; Ferreira, M. U. Higher microsatellite diversity in *Plasmodium vivax* than in sympatric *Plasmodium falciparum* populations in Pursat, western Cambodia. 2012. Artigo submetido.

Figura 10- Mapa da localização de Camboja e Sri Lanka, países dos quais foram analisadas amostras de DNA para este estudo.

3.3.5 Centro para Controle e Prevenção de Doenças (CDC), Estados Unidos da América

Alguns isolados de *P. vivax* utilizados neste estudo são oriundos de um projeto realizado no CDC (do inglês *Centers for Disease Control and Prevention*) em Atlanta, nos Estados Unidos, dentre os anos de 2004 e 2008. Estes isolados são de casos de malária vivax importada com diagnóstico laboratorial (microscopia convencional de esfregaço e gota espessa, corados com GIEMSA) feito pelo CDC a pedido de outros laboratórios de saúde pública nos Estados Unidos. Os pacientes foram interrogados a fim de se obter dados como histórico de viagens, migração de região endêmica de malária. Desta forma foi possível inferir a localização geográfica destes parasitos. Estas amostras foram gentilmente cedidas pelo Pesquisador Alexandre Januário da Silva. Essas amostras também foram utilizadas para a tipagem de microsatélites, e para a análise de padrões regionais de distribuição dos polimorfismos descritos neste estudo.

3.4 Caracterização de marcadores moleculares baseados em microarranjos

3.4.1 Seleção e caracterização de SNPs identificados por hibridação em microarranjos

Este estudo teve início com a identificação de SNPs baseada nos experimentos de hibridação genômica comparativa, realizados por Scott Westenberger (Scripps Research Institute, La Jolla, EUA), com três isolados (Acre5, Acre9 e Acre10) coletados no Acre no segundo semestre de 2009. Os experimentos consistiram em hibridizar DNA altamente purificado desses isolados em microarranjos de alta densidade contendo o genoma completo do isolado Salvador-I de *P. vivax* (Westenberger et al., 2009). A ausência de sinal de hibridação foi interpretada como indicativo de substituição ou deleção de um ou mais nucleotídeos na região genômica do isolado de campo correspondente ao alvo da sonda empregada no microarranjo. Com base nesses experimentos de hibridação, foram selecionados SNPs presentes em regiões codificadoras de genes associados ao metabolismo do parasito (*housekeeping genes*) (Anderson et al., 2005), evitando-se marcadores em regiões intergênicas ou introns (cuja extensa variabilidade poderia dificultar o desenho de iniciadores para a sua amplificação) e polimorfismos em proteínas de superfície do parasito, potencialmente expostas ao sistema imune e sujeitas a seleção diversificadora. Optou-se também por SNPs distribuídos em diferentes cromossomos e foram selecionados aqueles que se mostraram polimórficos no conjunto de três isolados de campo (Acre5, Acre9, Acre10) submetidos à hibridação genômica.

A Tabela 1 relaciona os 18 SNPs selecionados com base nesses critérios, com suas respectivas localizações cromossômicas. Um fragmento de cerca de 400 pares de bases (pb) em torno de cada um desses SNPs foi amplificado com os Oligos (*oligonucleotídeos*) relacionados na Tabela 1. Os oligonucleotídeos foram desenhados com o auxílio do programa Oligo3 (disponível em: http://frodo.wi.mit.edu/cgi-bin/Oligo3/Oligo3_www.cgi). As temperaturas de dissociação (T_m) situam-se em torno de 60° C. Cada oligonucleotídeo compreende uma cauda na extremidade 5' (GTAAAACGACGGCCAGT para os anteriores e CAGGAAACAGCTATGACC para iniciadores reversos), derivada do vetor M13, para permitir o

sequenciamento dos fragmentos amplificados com iniciadores universais correspondentes a esse vetor.

Tabela 1- Oligonucleotídeos utilizados para a amplificação e sequenciamento de 18 regiões genômicas com possíveis polimorfismos (SNPs) em isolados brasileiros de *Plasmodium vivax*. A identificação dos *loci* e sua localização cromossômica foram baseadas na anotação do genoma do isolado Salvador-I (Carlton et al., 2008). A posição aproximada de cada SNP (no genoma de Salvador-I) foi deduzida com base nos experimentos de hibridação genômica.

(continua)

Cromossomo	Locus	Oligonucleotideo (5'→ 3')	Tm (° C)	Posição SNP
1	PVX_093615	Oligo 1 F 5'GAAGCCTCCAAAATCCCAAC 3' Oligo 1 R 5'AAGGTGGACTACCGGCAGTA 3'	60.81	689630
1	PVX_093680	Oligo 2 F 5'CCCAAGAATGAGAGGTGGAA 3' Oligo 2 R 5'CAACCCGGACACCATGAT 3'	60.04	773225, 773423
2	PVX_096960	Oligo 3 F 5'AAAGAATCACCGGGGCTAAC 3' Oligo 3 R 5'AAATTCTCCCATCCCTCTGA 3'	60.32	61219
2	PVX_081280	Oligo 4 F 5'AACCCATCCGTTTGT TTTTG 3' Oligo 4 R 5'ATGGAAGCAACAGTGGGAAG 3'	59.97	228074
2	PVX_081270	Oligo 5 F 5'GAAATGGGGCACCTGGAAT 3' Oligo 5 R 5'GATTTGGATACCTATTCGGATG 3'	61.00	216116
4	PVX_003735	Oligo 6 F 5'GCCTCAGCATTAGGGATTACA 3' Oligo 6 R 5'TGGAAAAGGAGTTTAAGCAGGA 3'	59.20	658799
4	PVX_003590	Oligo 7 F 5'GGGCGAAGAACAGGAAAACT 3' Oligo 7 R 5'TATGGGTCTGTGAAGGTGTCC 3'	60.61	753330

Tabela 1- Oligonucleotídeos utilizados para a amplificação e sequenciamento de 18 regiões genômicas com possíveis polimorfismos (SNPs) em isolados brasileiros de *Plasmodium vivax*. A identificação dos *loci* e sua localização cromossômica foram baseadas na anotação do genoma do isolado Salvador-I (Carlton et al., 2008). A posição aproximada de cada SNP (no genoma de Salvador-I) foi deduzida com base nos experimentos de hibridação genômica.

(conclusão)				
Cromossomo	Locus	Oligonucleotídeos (5'→3')	Tm (° C)	Posição SNP
5	PVX_089720	Oligo 8 F 5'GACCCCCAGCAATAAAGAAA 3' Oligo 8 R 5'CTATGCACCACCCCTGGTT 3'	59.02	787926
7	PVX_099005	Oligo 9 F 5'GAACTCCAAGACGTGTTTGCT 3' Oligo 9 R 5'TGACTACAACACTTCGAACGT AAA 3'	59.39	438255, 438571
8	PVX_095265	Oligo 10 F 5'ACTCCCTGAAGACGTGAGG 3' Oligo 10 R 5'AGGTTGGAAGTGTCTGCCC 3'	59.30	946038
8	PVX_119610	Oligo 11 F 5'TCCACCTGTGCATAACTCCA 3' Oligo 11 R 5'CGTGTGCTATCTTTGCCTT 3'	60.11	1344682
9	PVX_091230	Oligo 12 F 5'GCGCCCAAATAGAAAAGCTA 3' Oligo 12 R 5'ACTGTCTGGGTATGGGGATG 3'	59.48	351724
9	PVX_091465	Oligo 13 F 5'CAAGTCTTCCCCAGCTGTT 3' Oligo 13 R 5'CGTTTCTAAAAATTCGTTTCATC 3'	59.33	539865
10	PVX_097985	Oligo 14 F 5'TCCATTCCCATTAACTGTG 3' Oligo 14 R 5'GAGTGATCGACGTGTATGCG 3'	60.05	998797
11	PVX_113345	Oligo 15 F 5'GAAGGCTACCCAGGGTCATT 3' Oligo 15 R 5'TACACTGGTTTGGCGGCTAC 3'	60.33	1894702
12	PVX_118575	Oligo 16 F 5'GTATTTTGGCGGAAGCAAA 3' Oligo 16 R 5'AGCACTGGATGGAAAAATAAAA 3'	60.07	2846101
13	PVX_085525	Oligo 17 F 5'GGGACTCCCACTGAGCATA 3' Oligo 17 R 5'TAAAAGAAAACAAGAATAAGAGGAGA 3'	60.07	1240627
14	PVX_123885	Oligo 18 F 5'AGACACTTCAACGCGGACAT 3' Oligo 18 R 5'CACGAGAGGAAGTCCCTGAG 3'	60.72	1872561

Nota: Tm = temperatura de dissociação.

3.4.2 Genotipagem e confirmação de SNPs caracterizados por microarranjos

Para confirmar a existência de SNPs, foram selecionados três isolados brasileiros, Acre5, Acre9 e Acre10, os mesmos utilizados no experimento de hibridação genômica realizados por Scott Westenberger. O DNA destes isolados foi primeiramente amplificado por WGA para que se pudesse obter o volume necessário de DNA para os ensaios. Após a técnica de WGA, foram padronizadas as condições para amplificação (por PCR) e sequenciamento das 18 regiões genômicas relacionadas na Tabela 1.

O protocolo utilizado para o mix da amplificação por PCR consistiu em 0,25 µL de Recombinant *Taq* DNA polymerase 5U/µL (Fermentas, Burlington, Canadá), 1,5 µL 10× *Taq* buffer com (NH₄)₂SO₄ (Fermentas - este tampão contém 750 mM Tris-HCl (pH 8,8)), 200 mM (NH₄)₂SO₄ e 0.1% (v/v) Tween 20), 1,2 µL de MgCl₂ 25 mM (Fermentas), 0,6 µL de dNTP mix (dATP, dCTP, dGTP, e dTTP em ddH₂O, 10 mM cada), 0,7 µL de cada oligonucleotídeo – anterior e reverso na solução de trabalho de 5 µM/ µL, 3 µL de DNA genômico e completar o volume de água para 15 µL. Os parâmetros do ciclo do PCR são: (1) 94° C por 2 min; (2) 94° C por 30 seg; (3) 58° C por 40 seg; (4) 72° C por 50 seg; (5) 72° C por 5 min; (6) repetir passos 2–4 por 35 vezes (7) manter a 15° C.

Com o produto da amplificação por PCR, foi realizada a técnica de eletroforese em gel de agarose, com o objetivo de separar, identificar e caracterizar os fragmentos gênicos.

3.4.3 Sequenciamento e Análises dos SNPs

Dentre os 18 oligonucleotídeos, somente três (Oligo4, Oligo5 e Oligo17) permitiram uma rápida padronização da amplificação e visualização de apenas uma banda específica em gel de agarose. Para os demais oligonucleotídeos foram testadas diferentes variações de temperatura de anelamento (temperatura variando entre 56° C a 60° C), sem sucesso. Os

fragmentos amplificados com os pares de oligonucleotídeos: Oligo4, Oligo5 e Oligo17 foram purificados utilizando o kit *GFX PCR and Gel band purification* (GE Technology Infrastructure - Uppsala/Suécia), e sequenciados.

Os oligonucleotídeos utilizados na reação de sequenciamento foram os mesmos utilizados para amplificar os fragmentos de interesse (Tabela 1). As reações de sequenciamento foram realizadas em ambas as fitas de DNA, em termociclador GeneAmp PCR 9700 (Applied Biosystems, EUA), conforme o protocolo: a) mix da reação: (volume final 10 µL), 2-4 µL de DNA (~100-200ng), oligonucleotídeo (5 µM), 3,25 µL de Tampão de sequenciamento 5X Save money (Tris-HCl 0,2 mM; MgCl₂ 5 mM) e 0,75 µL de Big Dye *Terminator Cycle Sequencing Ready Reaction Kit* v. 3.1 (Applied Biosystems, Foster City, CA). b) programa de amplificação consistiu em: desnaturação inicial a 96 ° C por 1 min, seguido de 30 ciclos com desnaturação a 96 ° C por 15 seg, anelamento a 50 ° C por 15 seg, extensão a 60 ° C por 4 min.

O produto da reação de sequenciamento foi precipitado adicionando-se 90 µl de Isopropanol 66% em cada amostra, a placa foi selada, homogeneizada e incubada em temperatura ambiente por 15 min. protegida da luz. Em seguida a placa foi centrifugada por 60 min. a 2464 g por 60 minutos. O sobrenadante foi descartado por inversão da placa e o sedimento lavado com 150 µL de isopropanol a 75%. A placa foi então centrifugada a 2464 g por 30 minutos. O sobrenadante foi descartado e a placa colocada para um pulso de 45 g por 30 segundos, com a placa invertida sobre papel absorvente. O DNA foi seco em temperatura ambiente, protegido da luz por 10 min. e pode ser armazenado a -20° C. O produto foi ressuspensão a 10 µl de tampão de amostra contendo Formamida Hi-Di (Applied Biosystems, EUA). Após serem ressuspensas, as amostras foram aquecidas a 95° C por 5 min, transferidas para o gelo e então colocadas no sequenciador capilar ABI310 (Applied Biosystems, EUA). As sequências resultantes foram obtidas dos cromatogramas pelo programa Sequence Scanner Software v1.0 (Applied Biosystems, EUA) foram alinhadas para obtenção da sequência consenso utilizando-se o programa BIOEDIT, em seguida foram alinhadas utilizando-se os programas Mega 4.0.2 e ClustalX1.83.

O sequenciamento permitiu identificar um SNP em posição próxima àquela estimada com base nos experimentos de hibridação, quando comparada com as sequências dos isolados de campo alinhadas com Sal-I. Porém, nesta análise, foi possível evidenciar a ocorrência de resultados falso-negativos, ou seja, a hibridação deixou de demonstrar, em algumas amostras, diferenças entre o genoma de Sal-I e dos isolados acreanos. A Tabela 2 mostra, para três *loci* analisados inicialmente, as discrepâncias entre os resultados esperados, com base nos experimentos de hibridação de DNA dos isolados Acre5, Acre9, Acre10 e aqueles obtidos por sequenciamento do DNA desses mesmos isolados.

Tabela 2- Diferenças entre o resultado esperado e o resultado observado em três *loci* analisados de Sal-1 e de amostras do Acre: Acre5, Acre9, e Acre10.

Lócus	Resultado Esperado	Resultado observado	SNP	Posição SNP	Cromossomo
PVX_081280	Acre5=Sal-I	Acre5(G)≠Sal-I(A)	G/A	228069	2
	Acre9=Sal-I	Acre9(G)≠Sal-I(A)			
	Acre10≠Sal-I	Acre10(G)≠Sal-I(A)			
PVX_081270	Acre5≠Sal-I	Acre5(C)≠Sal-I(T)	C/T	216116	2
	Acre9=Sal-I	Acre9(C)≠Sal-I(T)			
	Acre10=Sal-I	Acre10(C)≠Sal-I(T)			
PVX_085525	Acre5≠Sal-I	Acre5(T)≠Sal-I(C)	T/C	1240623	13
	Acre9=Sal-I	Acre9(T)≠Sal-I(C)			
	Acre10=Sal-I	Acre10(T)≠Sal-I(C)			

Nota: os sinais = e ≠ indicam, respectivamente, semelhança ou diferença de sequência em relação ao isolado Sal-I.

A partir dos dados analisados, foi possível observar que os microarranjos deixaram de identificar discrepâncias de sequência entre Sal-I e isolados Acre5/Acre9 no *loci* PVX_081280 e entre Sal-I e Acre9/Acre10 no *loci* PVX_081270 e PVX_085525. Com base no sequenciamento, todos os isolados acreanos apresentavam sequências idênticas entre si e distintas de Sal-I, enquanto que a hibridação realizada por Scott Westenberger sugeria que pelo menos um deles apresentava sequência idêntica à de Sal-I e diferente dos demais

isolados da mesma área. Assim, nenhum destes polimorfismos preenchia os pré-requisitos que haviam sido propostos em nosso projeto inicial para a seleção dos marcadores. Enquanto os resultados preliminares eram analisados, o grupo de Elisabeth Winzeler (*Institute for Childhood and Neglected Diseases*) publicou, em novembro de 2010, o sequenciamento genômico completo de um isolado de campo de *P. vivax* (IQ7) proveniente de Iquitos/Peru. O DNA genômico foi sequenciado utilizando uma plataforma Illumina (Dharia et al., 2010). O DNA desse isolado havia sido submetido à hibridação com o mesmo microarranjo contendo o genoma de Sal-I utilizado na análise dos isolados Acre5, Acre9 e Acre10. Deste modo, foi possível verificar se os possíveis polimorfismos identificados pela estratégia de hibridação eram confirmados por sequenciamento. Nos genes mais polimórficos, incluindo aqueles com regiões repetitivas de sequenciamento mais difícil, a estratégia de hibridação (utilizando o mesmo valor estrigente de $p = 1 \times 10^{-15}$ que adotamos para selecionar nossos possíveis SNPs) identificou um número *muito superior* de SNPs que o sequenciamento de DNA.

Os dados preliminares obtidos neste estudo e, especialmente, aqueles publicados por Dharia et al., (2010), sugerem que a hibridação com o microarranjo contendo o genoma completo de Sal-I não produz resultados confiáveis para a identificação de possíveis SNPs em isolados de campo de *P. vivax*. Portanto, não fornecem resultados úteis para a seleção de SNPs para posterior padronização e tipagem. Foi decidido após este resultado, mudar a estratégia de seleção de SNPs através duas abordagens alternativas: (a) a análise de sequências de DNA de genes de cópia única disponíveis no GenBank (2008); (b) a análise de SNPs previamente descritos na literatura.

3.5 Identificação de SNPs em sequências de isolados de *P.vivax* depositadas no GenBank

Com o intuito de identificar outros SNPs passíveis de uso em estudos populacionais de *P. vivax*, foi realizado o alinhamento e análise de sequências parciais ou completas de genes de cópia única do parasito depositadas no GenBank.

3.5.1 Critério de seleção das sequências codificadoras

O critério de seleção de sequências codificadoras (em vez de regiões intergênicas ou íntrons) deveu-se meramente à sua disponibilidade, pois há raros estudos de polimorfismo de sequências não-codificadoras nesse banco de dados. Foram incluídos para a análise conjuntos de dados com mais de 30 sequências provenientes de diferentes regiões endêmicas com a intenção de evitar a seleção de polimorfismos com claro viés regional e que incluíssem pelo menos alguns isolados brasileiros (para confirmar a ocorrência daquele polimorfismo no Brasil). Não é surpreendente que os conjuntos de dados de sequenciamento com essas características fossem referentes, em sua grande maioria, aos genes que codificam antígenos de superfície, potenciais candidatos ao desenvolvimento de vacinas, portanto sujeitos a intensa pressão seletiva: *AMA-1*, *MSP-1*, *MSP-3 α* , *MSP-4*, *MSP-5*, *DBP*. Por isso, a busca foi direcionada para *substituições sinônimas* nesses genes, potencialmente livres de pressão seletiva.

3.5.2 Análises das sequências codificadoras

Nesta análise foram incluídos inicialmente todos os genes acima mencionados (*AMA-1*, *MSP-1*, *MSP-3 α* , *MSP-4*, *MSP-5*, *DB*). Foi possível analisar nove sequências de *MSP-3 α* , a maioria proveniente de isolados coreanos, com polimorfismos localizados próximos ou no interior de regiões repetitivas, e 42 sequências de *MSP-4*, sendo 10 referentes ao Brasil, com polimorfismo limitado e enviesado regionalmente. Após análise, as sequências de *MSP-3* e *MSP-4* foram descartadas, pois não obedeciam aos critérios de seleção de inclusão neste estudo.

Na continuação das análises, foram identificadas substituições sinônimas bem distribuídas geograficamente nos seguintes genes: (a) *MSP-1*, localizado no cromossomo 7, com

aproximadamente 1000 sequências parciais ou completas analisadas, sendo cerca de 50 provenientes de isolados brasileiros; (b) *AMA-1*, localizado no cromossomo 9, com 104 sequências analisadas, sendo 16 referentes ao Brasil; (c) *MSP-5*, localizado no cromossomo 4, com 47 sequências analisadas, com 12 referentes ao Brasil e (d) *DBP*, localizado no cromossomo 6, com 30 sequências analisadas, a maioria do Brasil estas sequências foram escolhidas apesar de incluírem somente polimorfismos não-sinônimos. A Tabela 3 relaciona os primeiros polimorfismos selecionados para tipagem de SNPs.

Tabela 3- Caracterização dos primeiros polimorfismos selecionados para tipagem de SNPs.

<i>Locus</i>	Cromossomo	Sequências Analisadas	pb (Pares de bases)	SNPs	SNPs Selecionados	S-NS
<i>AMA-1</i>	9	104	1740	190	1	S
<i>MSP-1</i>	7	cerca de 1000	5511	836	3	S
<i>MSP-5</i>	4	42	1677	183	1	S
<i>DBP</i>	6	30	3213	49	1	NS
<i>MSP-3α</i>	10	9	3210	370	0	-
<i>MSP-4</i>	4	47	978	20	0	-

Nota: S= SNPs sinônimos e NS= SNPs não sinônimos.

3.6 Seleção de SNPs previamente descritos na literatura

Com base em estudo recente de 85 polimorfismos, originalmente identificados por Feng et al., (2003) ao longo do cromossomo 8 de *P. vivax*, em 240 isolados de campo (152 do Brasil) (Orjuela-Sánchez et al., 2010), foram selecionados três SNPs para tipagem: uma substituição sinônima em uma proteína hipotética (PVX-119300) na posição 1577813 (com base no genoma de Sal-I) e mais duas substituições em uma região intergênica, uma na posição 1623958 e outra na posição 1584352. Todos foram bem distribuídos geograficamente, mostraram-se polimórficos em isolados brasileiros e asiáticos. Embora estejam localizados no mesmo cromossomo, distam entre si mais de 150 kb.

Foi selecionada também uma substituição sinônima no gene *MDR-1*, que codifica a *glicoproteína-P*, um transportador de membrana dependente de energia pertencente à superfamília de transportadores ABC (do inglês *ATP-binding cassette*). Esse polimorfismo, originalmente identificado por Orjuela-Sanchez et al., (2009) em amostras brasileiras, também mostrou-se bem distribuído geograficamente em 254 amostras analisadas do Brasil e da Ásia (Sri Lanka, Vietnã e Camboja) (Orjuela-Sánchez et al., 2010). Em amostras asiáticas, esse gene pode ser encontrado eventualmente em múltiplas cópias; este é um genótipo aparentemente selecionado, nas três últimas décadas, por proporcionar resistência à mefloquina, fato que pode dificultar a interpretação de resultados de tipagem. Entretanto, no amplo estudo de Orjuela-Sánchez et al., (2010), esse polimorfismo não se destaca por apresentar resultados de tipagem dúbios, que pudessem ser atribuídos a existência de parálogos com sequências gênicas distintas entre si.

Finalmente, foram selecionamos para análise duas substituições sinônimas identificadas por Dharia et al. (2010) em genes que codificam proteínas hipotéticas (PVX_083040, cromossomo 12, e PVX_123750, cromossomo 14) da família do fator de transcrição AP2. Esses polimorfismos, confirmados por sequenciamento em plataforma Illumina e em sequenciamento tradicional de tipo Sanger, foram observados na comparação entre os genomas de um isolado de campo do Peru (IQ7) e do isolado Sal-I. Não se conhece, até o momento, sua distribuição geográfica; por isso, seu emprego em estudos populacionais será precedido de uma análise de distribuição geográfica com base em nossas amostras brasileiras e asiáticas (descritas em Orjuela-Sánchez et al., 2010).

3.7 Seleção de polimorfismos situados em região não-codificadora caracterizados por Gupta e colaboradores

Foi analisado um conjunto de seis polimorfismos em regiões não-codificadoras (P1 e P14 em introns e P3, P4, P13 e P15 em regiões intergênicas) descritos em isolados indianos por Gupta et al. (2010). Esses polimorfismos situam-se ao longo do cromossomo 13, tendo sido

identificados mediante o sequenciamento parcial de uma região de aproximadamente 200 Kb em número variável (entre 8 e 16) isolados da região de Assam, nordeste da Índia (Gupta et al., 2010). O objetivo foi determinar se esses polimorfismos eram compartilhados por amostras brasileiras, sugerindo tratar-se de substituições amplamente distribuídas geograficamente.

3.7.1 Extração do DNA e amplificação genômica com os oligonucleotídeos de Gupta e colaboradores

Foram utilizados para esta análise um total de 20 amostras sanguíneas coletadas de pacientes com malária vivax. Dentre estas, cinco eram procedentes do Ramal do Granada, sete de Plácido de Castro, localidades pertencentes ao estado do Acre, e oito amostras do Remansinho, Amazonas. A extração do DNA das amostras citadas acima, foi realizada a partir de 200 µL de sangue venoso, utilizando o kit QIAamp DNA Mini Kit (Qiagen- Hilden, Alemanha) através do aparelho de extração automatizada QIAcube comercializado pela Qiagen. Os oligonucleotídeos utilizados para as amplificações estão listados na Tabela 4.

A reação de amplificação e o programa da amplificação foram realizados conforme o protocolo original do artigo: a) mix da reação: (volume final 25 µL) DNA (~100-200ng), Oligonucleotídeo (0,3 µM), 2,5 µL de 10X Tampão de *Taq* (NH₄)₂SO₄, 1U de *Taq*, MgCl₂ (2,5 µM) dNTPs (0,2 µM). b) programa de amplificação consistiu em: (1) 94° C por 5 min; (2) 94° C por 50 seg; (3) 62° C por 50 seg; (4) 72° C por 50 seg; (5) 72° C por 5 min; (6) repetir passos 2-4 por 35 vezes (7) manter a 4° C.

Tabela 4- Sequência de oligonucleotídeos indianos envolvendo a amplificação de seis fragmentos de DNA não – codificantes com sua temperatura de anelamento.

Fragmento de DNA	Sequência do <i>primer</i> (5'→ 3')	Temperatura de Anelamento (° C)
Intron		
P1	F 5' GATGGGGGCCTTCCTAAT 3' R 5' ACAGTCGCAAATCGTAAAATC 3'	62
P14	F 5' TGGTCATTCATCAGCATCAG 3' R 5' CGACAAACGGAAGAAGCAGA 3'	62
Região intergênica		
P3	F 5' GTTTGGCCGTTTGACAC 3' R 5' TATAATTGACGTAAGACACC 3'	54
P4	F 5' GATTCGGGGTGTCTTACG 3' R 5' ATACGGCACGGATACGAT 3'	62
P13	F 5' GCCACCCCTTGTCTCTTC 3' R 5' AGACGGCCATTAAACACGAT 3'	63
P15	F 5' GTGCCGGTCGACTGGGGTGTTC 3' R 5' GCGGTCATGGGCAAGTCAAA 3'	66

3.7.2 Sequenciamento e análise de SNPs

A purificação do produto de PCR foi realizada utilizando o kit *GFX PCR and Gel band purification* (GE Technology Infrastructure - Uppsala/Suécia), os oligonucleotídeos utilizados na reação de sequenciamento são os mesmos utilizados para amplificar os fragmentos de interesse (Tabela 4). As reações de sequenciamento foram realizadas em ambas as fitas de DNA, em termociclador GeneAmp PCR 9700 (Applied Biosystems, EUA), conforme o protocolo: (a) mix da reação: (volume final 10 µL), 2-4 µL de DNA (100-200 ng), oligonucleotídeo (5 µM), 3,25 µL de Tampão de sequenciamento 5X Save money (Tris-HCl 0,2 mM; MgCl₂ 5 mM) e 0,75 µL de Big Dye *Terminator Cycle Sequencing Ready Reaction Kit* v. 3.1 (Applied Biosystems, Foster City, CA). (b) programa de amplificação consistiu em: desnaturação inicial a 96 ° C por 1 min, seguido de 30 ciclos com desnaturação a 96 ° C por

15 seg, anelamento a 50 ° C por 15 seg, extensão a 60 ° C por 4 min.

O produto da reação de sequenciamento foi precipitado adicionando-se 90 µl de Isopropanol 66% em cada amostra, a placa foi selada, homogeneizada e incubada em temperatura ambiente por 15 min. protegida da luz. Em seguida a placa foi centrifugada por 60 min. a 2464 *g* por 60 minutos. O sobrenadante foi descartado por inversão da placa e o sedimento lavado com 150 µL de isopropanol a 75%. A placa foi então centrifugada a 2464 *g* por 30 minutos. O sobrenadante foi descartado e a placa colocada para um pulso de 45 *g* por 30 segundos, com a placa invertida sobre papel absorvente. O DNA foi seco em temperatura ambiente, protegido da luz por 10 min. e pode ser armazenado a -20° C. O produto foi ressuspendido a 10 µl de tampão de amostra contendo Formamida Hi-Di (Applied Biosystems, EUA). Após serem ressuspendidas, as amostras foram aquecidas a 95° C por 5 minutos, transferidas para o gelo e então colocadas no sequenciador capilar ABI310 (Applied Biosystems, EUA). As sequências resultantes foram obtidas dos cromatogramas pelo programa Sequence Scanner Software v1.0 (Applied Biosystems, EUA) foram alinhadas para obtenção da sequência consenso utilizando-se o programa BIOEDIT, em seguida foram alinhadas utilizando-se os programas Mega 4.0.2 e ClustalX1.83.

Após esta etapa, o produto deste ciclo, foi transferido para uma placa de 96 poços e submetido ao sequenciamento direto por capilar contendo polímero (POP-7™ Polymer for 3500/3500xL Genetic Analyzer- Applied Biosystems, CA/EUA). Para o sequenciamento, foi utilizando o kit Big Dye (Applied Biosystems, CA/EUA) no sequenciador automático ABI Prism 3100 (Applied Biosystems, CA/EUA). As sequências resultantes foram alinhadas para obtenção da sequência consenso utilizando-se o programa BIOEDIT, em seguida foram alinhadas utilizando-se os programas Mega 4.0.2 e ClustalX1.83.

Figura 11- Em destaque, dois SNPs sinônimos encontrados no cromossomo 13 de *P. vivax* em isolados da Índia e Brasil. O alinhamento foi realizado no software Clustal X 1.83 durante a análise do oligonucleotídeo P1, proposto por Gupta et al., 2010.

	190	200	210	220	230	240	250	260
BR-22P1R	TTCAAGCGTA	TAAGTGGTGATAACTTTCTATTTATGGTGTAA	TCTCTCCAAATAATGCAAATGGTTAACGAG	AGTA				
BR-179P1R	TTCAAGCGTA	TAAGTGGTGATAACTTTCTATTTATGGTGTAA	TCTCTCCAAATAATGCAAATGGTTAACGAG	AGTA				
BR-23P1R	TTCAAGCGTA	TAAGTGGTGATAACTTTCTATTTATGGTGTAA	TCTCTCCAAATAATGCAAATGGTTAACGAG	AGTA				
S2	TTCAAGCGTA	TAAGTGGTGATAACTTTCTATTTATGGTGTAA	TCTCTCCAAATAATGCAAATGGTTAACGAG	AGTA				
BR-183P1R	TTCAAGCGTA	TAAGTGGTGATAACTTTCTATTTATGGTGTAA	TCTCTCCAAATAATGCAAATGGTTAACGAG	AGTA				
S13	TTCAAGCGTA	TAAGTGGTGATAACTTTCTATTTATGGTGTAA	TCTCTCCAAATAATGCAAATGGTTAACGAG	AGTA				
S9	TTCAAGCGTA	TAAGTGGTGATAACTTTCTATTTATGGTGTAA	TCTCTCCAAATAATGCAAATGGTTAACGAG	AGTA				
S10	TTCAAGCGTA	TAAGTGGTGATAACTTTCTATTTATGGTGTAA	TCTCTCCAAATAATGCAAATGGTTAACGAG	AGTA				
S16	TTCAAGCGTA	TAAGTGGTGATAACTTTCTATTTATGGTGTAA	TCTCTCCAAATAATGCAAATGGTTAACGAG	AGTA				
BR-160P1R	TTCAAGCGTA	TAAGTGGTGATAACTTTCTATTTATGGTGTAA	TCTCTCCAAATAATGCAAATGGTTAACGAG	AGTA				
BR-28P1R	TTCAAGCGTA	TAAGTGGTGATAACTTTCTATTTATGGTGTAA	TCTCTCCAAATAATGCAAATGGTTAACGAG	AGTA				
BR-10P1R	TTCAAGCGTA	TAAGTGGTGATAACTTTCTATTTATGGTGTAA	TCTCTCCAAATAATGCAAATGGTTAACGAG	AGTA				
BR-51P1R	TTCAAGCGTA	TAAGTGGTGATAACTTTCTATTTATGGTGTAA	TCTCTCCAAATAATGCAAATGGTTAACGAG	AGTA				
S3	TTCAAGCGTA	TAAGTGGTGATAACTTTCTATTTATGGTGTAA	TCTCTCCAAATAATGCAAATGGTTAACGAG	AGTA				
S4	TTCAAGCGTA	TAAGTGGTGATAACTTTCTATTTATGGTGTAA	TCTCTCCAAATAATGCAAATGGTTAACGAG	AGTA				
S7	TTCAAGCGTA	TAAGTGGTGATAACTTTCTATTTATGGTGTAA	TCTCTCCAAATAATGCAAATGGTTAACGAG	AGTA				
S14	TTCAAGCGTA	TAAGTGGTGATAACTTTCTATTTATGGTGTAA	TCTCTCCAAATAATGCAAATGGTTAACGAG	AGTA				
BR-31P1R	TTCAAGCGTA	TAAGTGGTGATAACTTTCTATTTATGGTGTAA	TCTCTCCAAATAATGCAAATGGTTAACGAG	AGTA				
S12	TTCAAGCGTA	TAAGTGGTGATAACTTTCTATTTATGGTGTAA	TCTCTCCAAATAATGCAAATGGTTAACGAG	AGTA				
S5	TTCAAGCGTA	TAAGTGGTGATAACTTTCTATTTATGGTGTAA	TCTCTCCAAATAATGCAAATGGTTAACGAG	AGTA				
S11	TTCAAGCGTA	TAAGTGGTGATAACTTTCTATTTATGGTGTAA	TCTCTCCAAATAATGCAAATGGTTAACGAG	AGTA				
BR-80P1R	TTCAAGCGTA	TAAGTGGTGATAACTTTCTATTTATGGTGTAA	TCTCTCCAAATAATGCAAATGGTTAACGAG	AGTA				
S6	TTCAAGCGTA	TAAGTGGTGATAACTTTCTATTTATGGTGTAA	TCTCTCCAAATAATGCAAATGGTTAACGAG	AGTA				
BR-27P1R	TTCAAGCGTA	TAAGTGGTGATAACTTTCTATTTATGGTGTAA	TCTCTCCAAATAATGCAAATGGTTAACGAG	AGTA				
S15	TTCAAGCGTA	TAAGTGGTGATAACTTTCTATTTATGGTGTAA	TCTCTCCAAATAATGCAAATGGTTAACGAG	AGTA				
BR-54p1R	TTCAAGCGTA	TAAGTGGTGATAACTTTCTATTTATGGTGTAA	TCTCTCCAAATAATGCAAATGGTTAACGAG	AGTA				
BR-36P1R	TTCAAGCGTA	TAAGTGGTGATAACTTTCTATTTATGGTGTAA	TCTCTCCAAATAATGCAAATGGTTAACGAG	AGTA				
BR-25P1R	TTCAAGCGTA	TAAGTGGTGATAACTTTCTATTTATGGTGTAA	TCTCTCCAAATAATGCAAATGGTTAACGAG	AGTA				
S8	TTCAAGCGTA	TAAGTGGTGATAACTTTCTATTTATGGTGTAA	TCTCTCCAAATAATGCAAATGGTTAACGAG	AGTA				
S1	TTCAAGCGTA	TAAGTGGTGATAACTTTCTATTTATGGTGTAA	TCTCTCCAAATAATGCAAATGGTTAACGAG	AGTA				
BR-92P1R	TTCAAGCGTA	TAAGTGGTGATAACTTTCTATTTATGGTGTAA	TCTCTCCAAATAATGCAAATGGTTAACGAG	AGTA				
BR-7P1F	TTCAAGCGTA	TAAGTGGTGATAACTTTCTATTTATGGTGTAA	TCTCTCCAAATAATGCAAATGGTTAACGAG	AGTA				

Nota: A cor azul e vermelha representam o SNP T/C, cuja posição do polimorfismo é 194 e a cor amarelo e azul representam o SNP G/A, cuja posição é 257. Como evidenciado, em ambos isolados brasileiros (com iniciais BR) e em isolados indianos (com iniciais S) observa-se a presença do polimorfismo.

O critério de seleção de SNPs foi de utilizar as sequências que apresentavam substituições sinônimas, polimorfismos compartilhados entre isolados indianos e isolados brasileiros, e ausência de regiões repetitivas ao redor do polimorfismo. Essas análises permitiram a identificação de dois polimorfismos (T /C -194 e G/A- 257) em P1, como mostrado na Figura 11.

Os dois polimorfismos adicionais (localizados no cromossomo 13 de *P. vivax*) foram incluídos na seleção de potenciais marcadores a serem utilizados na tipagem de SNPs de *P.vivax* e estão listados na Tabela 5.

Tabela 5- Características dos 14 polimorfismos selecionados para uso em estudos populacionais.

Locus	Cromossomo	Posição
Proteína hipotética PVX_119300	8	1577813
Região Intergênica	8	1623958
<i>PvMDR-1</i>	10	362033
Proteína hipotética PVX_083040	12	3379
Proteína hipotética PVX_123750	14	6858
<i>AMA-1</i>	9	122951
<i>MSP-1</i>	7	158123
<i>MSP-5</i>	4	629859
Região Intergênica	8	1584352
<i>MSP-1</i>	7	1161690
<i>MSP-1</i>	7	1160451
<i>DBP</i>	6	388126
Intron	13	1753649
Intron	13	1753768

Nota: posição definida com base no genoma do isolado Sal-I (Carlton et al., 2008).

3.8 Padronização dos ensaios de tipagem TaqMan®

Neste estudo foram realizadas duas metodologias de tipagem de SNPs, a primeira na plataforma Mastercycler Realplex gradient thermal cycler e a segunda, na plataforma StepOnePlus v.2.1.

3.8.1 Quantificação do DNA das amostras utilizadas no estudo por PCR em tempo real

Antes de iniciar a tipagem, foi feita uma amplificação total do genoma de todas as amostras pela técnica de WGA (descrita no item 3.2), em seguida foi feito um PCR em tempo real (quantitativo) para que se soubesse o número exato de cópias do parasito no DNA molde e estimar sua concentração dentro do DNA total (humano + parasito) este dado é útil uma vez que a concentração de DNA do parasito é normalmente baixa dentro do DNA total. O PCR em tempo real tem como alvo o gene 18S rRNA, que possui 5 cópias no genoma do parasito e tem como intuito quantificar a parasitemia de *P. vivax* em amostras clínicas. Cada 15 µL de reação contem 2 µL de DNA genômico, 7,5 µL de 2x Maxima SYBR Green qPCR master mixture (Fermentas, Vilnius, Lithuania) e 0,5 µM de cada oligonucleotídeo. Foi utilizado oligonucleotídeo gênero específico P1 (ACG ATC AGA TAC CGT CGT AAT CTT) combinados com o oligonucleotídeo espécie- específico, V1 (CAA TCT AAG AAT AAA CTC CGA AGA GAA A). Estes oligonucleotídeos permitem a amplificação de fragmentos espécie específicos de ~100 kb do gene 18S rRNA (Gonçalves et al, 2010). Curvas padrão (Figura 12) foram preparadas com diluições seriadas (dez vezes) da sequência alvo, clonada em um vetor pGEM-T Easy (Promega, WI, EUA). A amplificação por PCR foi feita em triplicatas, no aparelho Mastercycler Realplex gradient thermal cycler (Eppendorf, Hamburgo, Alemanha) consistindo da desnaturação do DNA molde a 95 ° C por 10 min e 40 ciclos de 15 seg a 95 ° C e 1 min a 60 ° C, com a aquisição de fluorescência ao final de cada passo da extensão. A amplificação foi imediatamente seguida por uma curva de dissociação que consistiu de 15 seg a 95 ° C, 15 seg a 60 ° C e um passo de aumento de temperatura gradual de 0.2 ° C s⁻¹ até 95 ° C , com a aquisição de fluorescência a cada transição de temperatura. O resultado obtido pelo programa é dado em número de cópias do genoma do parasito em uma amostra de DNA, que permite estimar o numero correspondente de parasitos na amostra.

Figura 12- Exemplo de curva-padrão, obtida em PCR de tempo real, no aparelho Mastercycler Realplex gradient thermal cycler (Eppendorf, Hamburgo, Alemanha), para a quantificação de parasitemias de *Plasmodium vivax*. Na parte superior da figura são mostrados valores de fluorescência a cada ciclo de amplificação, enquanto na parte inferior da figura mostra-se a curva-padrão obtida com esses dados.

3.8.2 Padronização dos ensaios TaqMan® para tipagem de SNPs na plataforma Mastercycler Realplex gradient thermal cycler

Em um primeiro momento, a tipagem dos SNPs foi realizada através da técnica de PCR em tempo real, na plataforma Mastercycler (Eppendorf, Alemanha), com a utilização de ensaios do tipo TaqMan® (Applied Biosystems, CA/EUA) com sondas alelo-específicas marcadas com fluorocromos. Sondas tipo TaqMan® são sondas de hidrólise designadas para aumentar e especificidade do PCR. Este método foi primeiramente reportado em 1991 por pesquisadores Cetus Corporation e a tecnologia foi subsequentemente desenvolvida pela Applied Biosystems para aplicações na pesquisa. O princípio desta técnica baseia-se na atividade 5'-3' exonuclease da Taq polymerase para separar uma sonda duplamente marcada durante a hibridação para a sequência alvo complementar e a detecção do fluoróforo. Como em outros PCRs Real time, o resultado do sinal da fluorescência permite

mensurar quantitativamente a acumulação de produtos durante os estágios exponenciais do PCR.

A escolha desta estratégia foi devida à facilidade de reprodutibilidade destes ensaios em outros laboratórios de pesquisa que utilizem estes reagentes e marcadores, o que possibilitará também futuras comparações e o compartilhamento de dados obtidos de formas independentes.

Após a caracterização dos 14 polimorfismos, oligonucleotídeos e sondas marcadas foram desenhadas para a tipagem de cada SNP, com o auxílio do software Oligo Express, disponibilizado pela Applied Biosystems, (Custom SNP Genotyping Assays - Oligos and Probes Assay-by-Design). Depois de realizado o desenho dos oligonucleotídeos pelo software Oligo Express, foi possível customizar os ensaios do tipo TaqMan® com sonda MGB (Minor Groove Binder) através da Applied Biosystems (CA/EUA).

Alguns critérios foram seguidos para o desenho dos ensaios TaqMan®, entre eles: (a) tamanho do fragmento dos oligonucleotídeos entre 18 a 30 pares de bases; (b) conteúdo de CG de 30 a 80%; (c) temperatura de anelamento em torno de 59° C; (d) a variação da temperatura de melting entre os oligonucleotídeos *forward* e *reverse* menor que 4° C; (e) ausência de repetição de nucleotídeos, especialmente para três ou mais Gs ou Cs no final 3'; (f) ausência de complementaridade dentre dos oligonucleotídeos para evitar *hairpins*; (g) ausência de complementaridade da sonda com ambos os oligonucleotídeos; (h) ausência de nucleotídeo G no final 5' da sonda, pois a presença deste inativa o desempenho do fluoróforo e (i) a posição do polimorfismo deve estar no centro da sonda. A Tabela 6 apresenta as sequências dos oligonucleotídeos e das sondas marcadas para a tipagem de cada SNP.

Tabela 6- Sequências de oligonucleotídeos e sondas utilizados nos ensaios TaqMan® marcadas com VIC (Sonda 1) e FAM (Sonda 2) para tipagem de SNPs de *P.vivax* selecionados.

Ensaio	Gene	Oligonucleotídeo (5'→3')	Sondas 1 e 2
Taqman1	Proteína hipotética PVX_119300	F:5'GCTCCCCTTGCCTAGCTACTAA3' R:5'CACAAAATTGTATACACATGCAAAATACC3' F:5'CGGCCAGTGTGAAAACCTGAA3'	TGCTCTCTTAGAGGAA TGTGCTCTCTTAAGG AACGTTAATATGCTCCAAT
TaqMan2	Região Intergênica	R:5'CTCTATACGCATATACACGTGAACATGT3' F:5'CGAATCAGTACGACACCAACGT3'	ATAACGTTAATATGCTCCA ACCTATGAAAAGAG
TaqMan3	Pvmdr-1 Proteína hipotética IQ -	R:5'CAATGGCAATCCTTTGTTCTG3' F:5'CGCCTTTACCTTCACATTTGAAA3'	TGGAAAAGCTTATCC CGCTACGATCTGACGG
TaqMan4	PVX_083040 Proteína hipotética IQ -	R:5'CGAATTGGTAAACCCTAATGATGA3' F:5'TCTATCGTTGCTCAGTTGGTTGTT3'	TGCGCTACGATATGA TCATCAGTTTGTTGTAAAGT
TaqMan5	PVX_123750	R:5'CGAAAAGGTGGAGGAATCAAAG3' F:5'TCGTCATAGAGAATTCCGACGTT3'	CAGTTTGTTATAAGTTGTG TTCTTAAGCCCTGTGGC
TaqMan6	AMA-1	R:5'CCTCCATCCTTCAGCTTCTGA3' F:5'AGAGCTACGAAGAATTCAATCAACTG3'	CTTCTTAAGACCTGTGGCT ACGTGATAAATTCCACTA
TaqMan7	MSP-1	R:5'AGCTTCGCCCTCAACAAATC3' F:5'GACCAAGTGATAAGCAAGCTGAAG3'	CACGTGATCAACTTCCACTA TTGATTCTACTGTGACAGGT
TaqMan8	MSP-5	R:5'GCACTGAGGCTTGTCGTCTT3' F:5'GCAACGCAGCTAGTCAATTATCA R:5'CCCAATGGCAGTTACGAGGAT	TTGATTCTACCGTGACAGG TCGTCCAGTTCC CTCGTCCAGCTTC
TaqMan9	Região intergênica	F:5'GAAGAAAACGTAGCAGCAAAAGC R:5'CCCAGTGCTCATCGTTGTTG F:5'TGCGATGAATTGGACCTCCTA	ACAAAGTAATCGCACTC TACGACAAAGTAATCGCGCT TACTCCATATATGACTCCAT
TaqMan10	MSP-1	R:'CCATTCCTTCTGGTACAACTCAA F:5'GGAAGGCATCGGATATTCCA	ATATATGACTCTATGAGCAA TAGTGGAAGATAATTTGCGCAG
TaqMan11	MSP-1	R:5'CTGTTGGGCATTTTTACCAGTTC F:5'TGTAACGCATGAAGGAGCACTC	AAGATAATTTGCGTAGCAT TTCAAGCGTATAAGTGGTGA
TaqMan12	DBP	R:5'CCATTTGCATTATTTGGAGAGAATT F:5'ATTCTCTCAAATAATGCAAATGGT	AAGCGTACAAGTGGT AACGAGGTAATAAAA
TaqMan13	Intron	R:5'CCAAGTGAACATACCCCCATGT	AACGAGAGTAATAAAC

Nota: A base destacada na cor vermelha indica a localização do SNP em cada sonda.

Para a padronização dos ensaios foram testadas diferentes concentrações de sonda até ser encontrada a ideal. Foi utilizado o protocolo: 7,5 µl de Master Mix Universal (Applied Biosystems, CA/EUA), 0,135 µL de oligonucleotídeo Forward e Reverse (100 µM), sonda (variações de: 0,2 ou 0,3 ou 0,6 µM), 5 µL de DNA genômico (DNA do parasito variou entre 0,6 a 0,07 pg/µl) e água ultra pura para completar o volume final de 15 µl. O programa de amplificação consistiu em: (1) 95° C por 10 min; (2) 92° C por 15 seg; (3) 60° C por 1min. (4) repetir passos 2 e 3 por 40 vezes.

Esta metodologia de tipagem foi inicialmente feita com os oito primeiro ensaios TaqMan devido a disponibilidade no início da padronização. No entanto, não foi possível padronizar todos os oito ensaios e a primeira etapa de padronização da tipagem de SNPs apresentou algumas lacunas, como por exemplo a impossibilidade de analisar em conjunto as duas sondas, uma vez que o resultado era dado em número de cópias do parasito com um alelo específico e a ausência de controles positivos para validação de cada ensaio.

3.8.3 Padronização dos ensaios de tipagem TaqMan® na plataforma StepOnePlus v.2.1

Para solucionar estas questões foi adotada outra metodologia de tipagem de SNPs desenvolvida através de um PCR em tempo real com a plataforma StepOnePlus v.2.1 (Applied Biosystems, EUA).

Nesta etapa de tipagem de SNPs foram utilizados somente os ensaios que obtiveram sucesso na padronização na plataforma Mastercycler Realplex gradient thermal cycler (Eppendorf, Hamburg, Germany) e os ensaios TaqMan 9, TaqMan 10, TaqMan11, TaqMan 12, TaqMan 13 e Taqman 14 que não haviam sido padronizados anteriormente como descrito na Tabela 7, totalizando 10 ensaios.

Tabela7- SNPs tipados na plataforma StepOnePlus v.2.1. Ensaio correspondente, localização cromossômica e posição do SNP definida com base no genoma do isolado Sal-I (Carlton et al., 2008).

Ensaio	SNP	Cromossomo	Posição
TaqMan 1	G/A	8	1577813
TaqMan 2	C/G	8	1623958
TaqMan 3	G/A	10	362033
TaqMan 7	A/C	7	158123
TaqMan 9	T/C	8	1584352
TaqMan 10	A/G	7	1161690
TaqMan 11	C/T	7	1160451
TaqMan 12	C/T	6	388126
TaqMan13	T/C	13	1753649
TaqMan14	G/A	13	1753768

Para os novos PCRs foi utilizado o protocolo: 7,5 µl de master mix *TaqMan Genotyping Master Mix* (Applied Biosystems – EUA) conjunto de reagentes otimizados para a genotipagem de SNPs, incluindo AmpliTaq Gold DNA polimerase, dNTPs, referência passiva (ROX) e mix de componentes otimizados, uma variação de 0,25 µM a 0,35 µM de cada sonda MGB alelo-específica marcadas com VIC e FAM, uma variação de 0,9 µM a 1,5 µM de cada oligonucleotídeo anterior e reverso descrito na Tabela 6, 5 µl de DNA pré-amplificado com a parasitemia quantificada (DNA do parasito variou entre 0,6 a 0,07 pg/µl em cada amostra) e água ultra pura para completar o volume final de 20 µl. O programa de amplificação foi: 60 ° por 30 seg, 95 ° C por 10 min (ativação da AmpliTaq Gold), 40 ciclos a 95 ° C por 15 seg (desnaturação), 60° C por 1 min (anelamento/extensão), utilizando o aparelho StepOnePlus Real-Time PCR Systems (Applied Biosystems, EUA).

A determinação deu-se através do software StepOne Software 2.2.2 (Applied Biosystems, EUA). O programa utilizado normaliza a fluorescência do *Reporter dye* com o fluorocromo da referência passiva (ROX) em cada amostra e em seguida adiciona as intensidades normalizadas (Rn) de VIC e FAM em um gráfico de discriminação alélica. Cada sonda (VIC e FAM) corresponde a um único alelo e quando há a presença das duas sondas em uma mesma amostra, esta é considerada infecção mista, ou seja, o individuo foi infectado por

dois clones distintos de *P. vivax*.

Para cada reação de PCR de cada ensaio, foram adicionados um controle negativo e três controles positivos: homozigoto para o alelo 1 (alelo1/alelo1), homozigoto para o alelo 2 (alelo2/alelo2) e heterozigoto (alelo1/alelo2) simulando uma infecção mista. Foi também adicionado uma amostra de DNA contendo infecção por *P. falciparum* e uma amostra de DNA humano negativo para infecção de *P.vivax* para garantir que os oligonucleotídeos e sondas MGB marcadas eram específicas para amplificar apenas o genoma de *P. vivax*. O controle positivo foi usado na concentração que apresentou um melhor desempenho de genotipagem, na maneira recomendada pelo fabricante.

3.9 Desenvolvimento de controles positivos através de genes sintetizados.

A primeira estratégia de validação dos resultados da tipagem dos ensaios TaqMan® foi a clonagem dos fragmentos resultantes do PCR de cada ensaio com dois isolados previamente conhecidos por ter cada um, um alelo distinto. Esta técnica não apresentou os resultados esperados, portanto optou-se pela confecção de genes sintetizados.

Foram sintetizados *in vitro* quatro genes com a sequência completa de DNA de cada um dos ensaios padronizados apresentados na Tabela 6. Foram unidos em um só gene quatro sequências de quatro ensaios distintos de forma que nunca os dois polimorfismos do mesmo ensaio estariam no mesmo gene sintético, como descreve a Tabela 8. Esta sequência foi clonada no vetor pUC57, o DNA foi purificado, liofilizado e posteriormente sequenciado como protocolo de qualidade da empresa como mostra a Figura 13.

Figura 13- Vetor pUC57 com o Gene Sintético 3 inserido e o resultado de seu sequenciamento.

Os genes foram ressuspensos conforme as instruções do fabricante (GenScript, EUA). Os microtubos foram centrifugados a 6000 x *g* por 1 minuto a 4 ° C. Foi então adicionado 20 µl de água estéril livre de DNA e RNA e aplicou-se um vortex durante 1 minuto. O material foi incubado a 50 ° C por 15 minutos para certificar que todo o DNA liofilizado foi dissolvido. A solução mãe foi diluída para obter solução de trabalho com uma concentração de 200 pg/µl. Os controles foram utilizados na reação com a presença de apenas um dos alelos (homozigoto para o alelo 1 (1/1) e homozigoto para o alelo 2 (2/2)) ou com a presença de ambos os alelos para simular uma infecção mista (alelo 1 e alelo 2 “infecção mista” (1/2)).

Tabela 8- Sequências de genes sintéticos utilizados nos ensaios TaqMan® como controle positivo. As bases em destaque representam o SNP encontrado no fragmento amplificado.

Gene Sintético	Sequência dos genes sintéticos	Pares de bases
Gene1 1A2C3A7A9C	GCTCCCCTTGCCTAGCTACTAAATGTGTGCTCTCTTAAGGAAGAACAGAAAGCATAACGCATTTGTGTGT GCCCCTGCGTGGGTATTTTGCATGTGTATACAATTTGTGCGGCCAGTGTGAAACTGAAGCACCACACG TCAGTAGCGATAACGTTAATATGCTCCAATATATGTTTATATGAGCATGTAACATGTTACGTGTATAT GCGTATAGAGCGAATCAGTACGACACCAACGTGGGACCCTATGGAAAAGCTTATCCGGTGGACAGAAAC AAAGGATTGCCATTGAGAGCTACGAAGAATTCAATCAACTGATGCACGTGATAAACTTCCACTATGATTT GTTGAGGGCGAAGCTGCAACGCAGCTAGTCAATTATCAGCTCGTCCAGTTCCAATCCTCGTAACTGCCA TTGGG	425
Gene2 1G2T3G7C9T	GCTCCCCTTGCCTAGCTACTAAATGTGTGCTCTCTTAAGGAAGAACAGAAAGCATAACGCATTTGTGTGT GCCCCTGCGTGGGTATTTTGCATGTGTATACAATTTGTGCGGCCAGTGTGAAACTGAAGCACCACACG TCAGTAGCGATAACGTTAATATGCTCCAATATATGTTTATATGAGCATGTAACATGTTACGTGTATAT GCGTATAGAGCGAATCAGTACGACACCAACGTGGGACCCTATGGAAAAGCTTATCCGGTGGACAGAAAC AAAGGATTGCCATTGAGAGCTACGAAGAATTCAATCAACTGATGCACGTGATCAACTTCCACTATGATTT GTTGAGGGCGAAGCTGCAACGCAGCTAGTCAATTATCAGCTCGTCCAGTTCCAATCCTCGTAACTGCCA TTGGG	425
Gene3 10A11C12C13C14A	GAAGAAAACGTAGCAGCAAAAGCGCAGGACTACGCCGAGGACTACGACAAAGTAATCGCACTCCCTCTGT TCGGCAACAACGATGACGACTGGGTGCGATGAATTGGACCTCTATTCAATGTTTCAAGCAAACTTGCCAG CCATGTACTCCATATATGACTCATGAGCAACGAGCTACAGAACCTTTACATTGAGTTGTACCAGAAGGA AATGGGGAAGGCATCGGATATTCCAAAGTAGTGGAAGATAATTTGCGCAGCATCTTTGGAAGTGGTAAAA ATGCCCAACAGTGTAACGCATGAAGGAGCACTCACGCAGATATGTTCAAGCGTCAAGTGGTGAATAACT TTCTATTTATGGTGAATTCTCTCCAATAATGCAAATGGATTCTCTCCAATAATGCAAATGGTTAACG AGAGTAATAAACACAACATGGGGGTATGTTCACTTGG	457
Gene4 10G11T12T13T14G	GAAGAAAACGTAGCAGCAAAAGCGCAGGACTACGCCGAGGACTACGACAAAGTAATCGCGCTCCCTCTGT TCGGCAACAACGATGACGACTGGGTGCGATGAATTGGACCTCTATTCAATGTTTCAAGCAAACTTGCCAG CCATGTACTCCATATATGACTCATGAGCAACGAGCTACAGAACCTTTACATTGAGTTGTACCAGAAGGA AATGGGGAAGGCATCGGATATTCCAAAGTAGTGGAAGATAATTTGCGTAGCATCTTTGGAAGTGGTAAAA ATGCCCAACAGTGTAACGCATGAAGGAGCACTCACGCAGATATGTTCAAGCGTAAAGTGGTGAATAACT TTCTATTTATGGTGAATTCTCTCCAATAATGCAAATGGATTCTCTCCAATAATGCAAATGGTTAACG AGGTAATAAACACAACATGGGGGTATGTTCACTTGG	457

Nota: as bases em destaque representam o SNP encontrado no fragmento amplificado.

3.10 Análise de microssatélites

3.10.1 Genotipagem de MSAT

Neste estudo foram utilizados dados prévios de tipagem de microssatélites em amostras do Brasil, Sri Lanka, Camboja e CDC. Estes dados foram empregados a fim de se comparar os padrões regionais de polimorfismos com aqueles obtidos com a tipagem de SNPs. Estes resultados são especialmente importantes na comparação da identificação das infecções mistas por ambas metodologias.

Para este fim, foram utilizados 13 MSATs altamente polimórficos descritos por Karunaweera (2007), ilustrados na Tabela 9, e foi realizada a amplificação individual de cada marcador utilizando *Taq* DNA polimerase e com o complemento do sistema de tampões da Fermentas (Vilnius, Lithuania). Para cada PCR foram utilizados 3 µl de DNA genômico com 2mM de $MgCl_2$, 6 pM de cada *oligonucleotídeo forward e reverse* amplificação, 0,1 mM de cada dNTP, 1U de *Taq* polimerase recombinante e 1,5 µl 10x de tampão de *Taq* polimerase e água para completar o volume final de 15 µl. Todos os reagentes são providos pela Fermentas, com exceção dos oligonucleotídeos que são da Applied Biosystems (Foster City, CA- EUA). Os parâmetros do ciclo de amplificação são: 1 ciclo a 94° C por 2 min, 40 ciclos a 94 ° C por 30 s, 58 °C por 40 s e 72 ° C por 50 s e um ciclo final de 72° C por 5 min.

Os produtos do PCR dos MSAT foram diluídos 10 vezes em água ultrapura. A partir desta diluição 1µl de cada marcador foi adicionado a 13µl de Formamida HI-Di Applied Biosystems (Foster City, CA- EUA). Os MSATs foram agrupados de acordo com os diferentes fluoróforos e tamanho aproximados, como descrito por Havryliuk et al., (2008) e Orjuela-Sánchez et al., (2009). Como marcador de peso molecular foram adicionados 0,25µl de Gene-Scan ROX 500 Applied Biosystems (Foster City, CA- EUA). Este conjunto de reagentes é colocado em tubos individualizados e foi sequenciado no aparelho de sequenciamento automático ABI310 Applied Biosystems (Foster City, CA- EUA).

A análise dos fragmentos obtidos do sequenciador foi feita pelo programa Stand (Tornem; Hughes, 2001). Este programa fornece o tamanho dos fragmentos amplificados e a altura dos picos (que corresponde á abundância relativa de parasitos com cada alelo na amostra), importante na identificação de infecções mistas. A cada isolado foi atribuído um haplótipo correspondente ao conjunto de alelos encontrado para cada marcador. Uma infecção foi considerada geneticamente mista quando o cromatograma revelava dois picos, um dos quais era o majoritário que correspondia à maior parasitemia e o outro minoritário. De acordo com o estudo de Anderson e colaboradores (2000), o pico minoritário deve ter no mínimo um quarto do tamanho do pico majoritário. Foram considerados somente picos acima de 200 unidades de fluorescência. Em infecções mistas, somente os haplótipos predominantes foram analisados.

Tabela 9- Oligonucleotídeos iniciadores (*primers*) dos 13 MSAT de *P. vivax*.

MSAT	Cromossomo de <i>P.vivax</i>	Oligonucleotídeos (5'-3')
MS1	3	F:6-FAMTCAACTGTTGGAAGGGCAAT R:ctgtcttTTGCTGCGTTTTTGTCTG
MS2	6	F:VICGAGCTAGCCAAAGGTTCAACA R:ctgtcttTGGGGAGAGACTCCCTTTTC
MS3	4	F:NEDGAAGATCCTGTGGAGGAGCA R:ctgtcttCTCCTCGCTCCTTTCCTT
MS4	6	F:6-FAMCGATTTACTGTTGACGCTGAA R:ctgtcttCAAAGGAACATGCTCGATGA
MS5	6	F:NEDCGTCCTCTATCGCGTACACA R:ctgtcttAAAGGGAGAGGAGCGAAAAC
MS6	11	F:VICGTTCTTCGGTGATCTCTGC R:ctgtcttGGAGGAÇATCAACGGGATT
MS7	12	F:6-FAMTTGCAGAAAATGCAGAGAGC R:ctgtctAGGGTCTTCAGCGTGTGTT
MS8	12	F:NEDAGAGGAGGCAGAAATGCAGA R:ctgtcttAGCCCCTTGCGTTCTTAT
MS9	8	F:6-FAMAGATGCCTACACGTTGACGA R:ctgtcttGAAGCTGCCCATGTGGTAAT
MS10	13	F:NEDTTATCCCTGCTGGATGTGAA R:ctgtcttCCTTCAGGTGGGACTTGTT
MS12	5	F:6-FAMAATGCGCATCCTATGTCTCC R:ctgtcttCTGCTGTTGTTGTTGCTGCT
MS15	5	F:6-FAMTGTGTTGCAAAGGAATCCACA R:ctgtcttCGGCCAGATGAAAAGGATAA
MS20	10	F:NEDTGTGTTGGTTGTTGATGGTGA R:ctgtcttGTGGCAGTGGCTCATCTTCT

Para facilitar a visualização dos fragmentos amplificados no sequenciador de DNA, foram incorporados os marcadores fluorescentes 6-FAM, NED, VIC aos oligonucleotídeos anteriores: F (do inglês *forward*). As letras minúsculas nos oligonucleotídeos reversos (R) (do inglês *reverse*) referem-se a uma sequência de nucleotídeos que inibe a adição de adenosina aos produtos amplificados, garantindo a reprodutibilidade de tamanho dos fragmentos.

3.10.2. Dados previamente descritos de resultados de MSTA

Os dados da literatura referentes a tipagem dos 13 MSATs analisados neste estudo foram extraídos dos trabalhos de Orjuela-Sánchez et al., 2009 (amostras do Brasil), e uma tabela existente que reúne os resultados de 841 amostras tipadas previamente por diferentes autores. Os dados dos isolados utilizados neste estudo foram relacionados na Tabela 10.

Tabela 10- Dados de tipagem de 13 MSATs analisados neste estudo.

Origem	Autor
Brasil	Karunaweera et al., 2008; Orjuela-Sánchez et al., 2009
Sri Lanka	Gunawardena et al., 2010
Camboja	Orjuela-Sánchez et al., artigo submetido
CDC (vários países)	dados não publicados

4.1 Análise da caracterização de marcadores moleculares – microarranjos

Foram analisados possíveis SNPs selecionados para tipagem genética de isolados de *P.vivax* baseados em experimentos de hibridação com microarranjos descritos na literatura. Os microarranjos proporcionaram uma varredura genômica por hibridação de pequenas sondas derivadas da sequência caracterizada no isolado Sal-1.

Inicialmente foram selecionados 18 SNPs, e para a confirmação deste resultado foram escolhidos três isolados brasileiros (os mesmos utilizados nos experimentos de hibridação), Acre5, Acre9 e Acre10 para padronizar a reação de PCR e sequenciamento. Dentre os 18 oligonucleotídeos selecionados, somente três (Oligo4, Oligo5 e Oligo17) permitiram uma boa padronização e o seu sequenciamento evidenciou SNPs com localização aproximada aos esperados, quando comparados ao experimento de hibridação, no entanto esta técnica também apresentou resultados falso-negativos, ou seja, a hibridação não localizou SNPs nos isolados acreanos quando comparados ao isolado Sal-1, como era o esperado.

O resultado esperado era que pelo menos um dos isolados acreanos apresentasse a sequência idêntica a do isolado Sal-1 e fosse diferente dos demais isolados da mesma área. Mas o resultado encontrado mostrou que com o oligo5 no *loci* PVX_081280 os isolados Acre5 e Acre9 e Acre10 apresentaram em sua sequência o polimorfismo G, diferente do isolado Sal-1 que tinha o polimorfismo A. Com o oligo5, *loci* PVX_081270, o polimorfismo encontrado nos três isolados acreanos foi C, todos diferentes da sequência do isolado Sal-1. e por fim no oligo17 *loci* PVX_085525 a sequência encontrada nos isolados acreanos foi T enquanto a sequência do isolado Sal-1 foi C. Por esta razão, todos os polimorfismos selecionados foram descartados por não preencherem os pré-requisitos que haviam sido propostos em nosso projeto inicial para a seleção dos marcadores.

4.2 Análise de sequências de DNA de genes de cópia única disponíveis no GenBank

Nesta segunda etapa foi realizada a análise do alinhamento de sequências disponíveis de isolados de *P.vivax* depositadas no GenBank. Foi adotado um critério de seleção destes isolados de forma que fossem utilizadas todas as sequências possíveis descritas. Primeiramente, foram escolhidos dados de no mínimo 30 sequências, de diferentes regiões endêmicas a fim de se evitarem polimorfismos com viés regional. Os conjuntos de dados de sequenciamento com essas características foram referentes, em sua grande maioria, aos genes que codificam antígenos de superfície, potenciais candidatos ao desenvolvimento de vacinas, portanto sujeitos a intensa pressão seletiva: *AMA-1*, *MSP-1*, *MSP-3 α* , *MSP-4*, *MSP-5*, *DBP*. Por tanto, direcionou-se a busca para *substituições sinônimas* nesses genes, potencialmente livres de pressão seletiva.

4.2.1 Análise das sequências codificadoras

Foram identificadas substituições sinônimas com boa distribuição geográfica nos seguintes genes: (a) *MSP-1*, localizado no cromossomo 7, com aproximadamente 1000 sequências parciais ou completas analisadas, sendo cerca de 50 provenientes de isolados brasileiros. Foram analisados 5511 pares de bases, identificados 836 SNPs e dentre estes, 3 foram selecionados para dar continuidade ao estudo; (b) *AMA-1*, localizado no cromossomo 9, com 104 sequências analisadas, sendo 16 oriundas do Brasil. Neste gene foram analisados 1740 pares de bases, identificados 190 SNPs e um SNP foi selecionado para a tipagem de polimorfismos; (c) *MSP-5*, localizado no cromossomo 4, com 47 sequências analisadas, com 12 referentes ao Brasil foram analisados 1677 pares de bases neste gene, com a identificação de 183 SNPs e a seleção de um SNP e (d) *DBP*, localizado no cromossomo 6, com 30 sequências analisadas, a maioria do Brasil. Foram analisados neste gene 3213 pares de bases, identificados 49 SNPs (todos não sinônimos) e escolhido um para tipagem.

Foi ainda possível analisar nove sequências de *MSP-3 α* , a maioria proveniente de isolados

coreanos, com polimorfismos localizados próximos ou no interior de regiões repetitivas, e 42 sequências de *MSP-4*, sendo 10 referentes ao Brasil, com polimorfismo limitado e enviesado regionalmente. Estas sequências foram descartadas após análise, pois não obedeciam aos critérios de seleção de inclusão neste estudo.

4.2.2 Análise de SNPs previamente descritos na literatura

4.2.2.1 Cromossomo 8

Foram selecionados três SNPs descritos por Orjuela-Sánchez e colaboradores (2010) para tipagem: uma substituição sinônima em uma proteína hipotética (PVX-119300) na posição 1577813 (com base no genoma de *Sal-I*) e duas substituição em uma região intergênica, uma na posição 1623958 e outra na posição 1584352 todos localizadas no cromossomo 8. Estes polimorfismos foram baseados em um estudo prévio de Feng e colaboradores (2003) com 85 polimorfismos identificados no sequenciamento de 240 isolados de campo (152 do Brasil). Os SNPs mostraram-se bem distribuídos geograficamente e polimórficos em isolados brasileiros e asiáticos e estão a mais de 150kb distantes entre si.

4.2.2.2 Gene *MDR-1*

Foi selecionada uma substituição sinônima no gene *MDR-1*, localizado no cromossomo 10, este polimorfismo foi originalmente identificado e descrito por Orjuela-Sanchez et al., (2010) em amostras brasileiras, mas também mostrou-se bem distribuído geograficamente entre amostras analisadas da Ásia (Sri Lanka, Vietnã e Camboja).

4.2.2.3 Proteínas hipotéticas codificadas por genes nos cromossomo 12 e 14

Foram selecionadas duas substituições sinônimas identificadas por Dharia et al. (2010) em genes que codificam proteínas hipotéticas (PVX_083040, cromossomo 12, e PVX_123750, cromossomo 14) da família do fator de transcrição AP2. Esses polimorfismos foram confirmados por sequenciamento em plataforma Illumina e pelo sequenciamento tradicional de tipo Sanger, quando comparados o genoma de IQ7 (Peru) e o genoma do isolado Sal-1.

4.2.2.4 Polimorfismos situados em região não-codificadora do cromossomo 13

Foi analisado um conjunto de seis polimorfismos em regiões não-codificadoras (P1 e P14 em introns e P3, P4, P13 e P15 em regiões intergênicas) identificados em isolados indianos por Gupta et al. (2010). Esses polimorfismos localizam-se ao longo do cromossomo 13 e foram localizados após o sequenciamento de aproximadamente 200kb de número variável (entre 8 e 16) isolados da região de Assam, nordeste da Índia. Estes polimorfismos foram selecionados a fim de se determinar se eram compartilhados por amostras brasileiras, sugerindo tratar-se de substituições amplamente distribuídas geograficamente. Após a análise de 20 isolados brasileiros, selecionaram-se dois SNPs compartilhados entre isolados indianos e brasileiros e sinônimos: T194C e G257A localizados na região P1, região não codificadora de proteínas.

4.3 Análise da genotipagem de SNPs na plataforma Mastercycler Realplex gradient thermal cycler

Para a padronização dos primeiros ensaios TaqMan® na plataforma Mastercycler Realplex gradient thermal cycler foram utilizadas 39 amostras de DNA genômico de pacientes com diagnóstico positivo para *P. vivax*, procedentes de regiões endêmicas de malária no Brasil. Foram escolhidos isolados de diferentes localidades do estado do Acre com o intuito de encontrar haplótipos diferentes e assim hibridizarem com sondas distintas. Com o mesmo propósito foram selecionadas amostras do Camboja (45 isolados) e Sri Lanka (12 isolados). As reações de PCR foram realizadas com um volume final de 15 µl e com concentrações de

DNA que variaram de 0,07 a 0,6 pg/ μ l (devido a variação da parasitemia de cada amostra, quantificadas pelo PCR em tempo real quantitativo descrito no item 3.8.1). Os ensaios Taqman na plataforma Mastercycler Realplex gradient thermal cycler foram realizados com as duas sondas (contendo cada uma um alelo diferente) marcadas com o mesmo fluorocromo FAM. Portanto as reações de PCR precisaram ser feitas de forma separada com cada uma das amostras. Os ensaios TaqMan1, Taqman3, Taqman 4 e Taqman 7 foram padronizados com 02 μ M de cada sonda, os ensaios Taqman2 e Taqman5 foram padronizados com 0,6 μ M de cada sonda e finalmente, o ensaio Taqman 8 foi padronizado com 0,3 μ M de cada sonda.

O ensaio TaqMan1- Proteína Hipotética PVX_119300, SNP (G/A)- evidenciou polimorfismo nas amostras do Brasil, com seis amostras hibridizando com a sonda 1 (alelo A), 17 amostras hibridizando com a sonda 2 de alelo G, e uma amostra apresentou sinal de hibridização com ambas as sondas, sendo assim uma infecção mista. Nas amostras do Camboja, uma hibridizou com a sonda 1 (alelo A), 38 hibridizaram com a sonda 2 (alelo G), e sete amostras apresentaram sinal de hibridização com as sondas 1 e 2. Todas as amostras do Sri Lanka apresentaram sinal de hibridização com a sonda 2 alelo G .

O ensaio TaqMan2 - (Região Intergênica, cromossomo 8), SNP C/G- não mostrou polimorfismo nas amostras do Brasil, todas as amostras hibridizaram com a sonda 2 (alelo G). Das amostras do Camboja, três apresentaram sinal de hibridização com as sondas 1 (alelo C) e 2 (alelo G) e 42 amostras hibridizaram com a sonda 2 (alelo G). Nas amostras do Sri Lanka, metade hibridizou com as sondas 1 e 2 (infecção mista) e a outra metade somente com a sonda 2 (alelo G).

O ensaio TaqMan 3 - Pvmdr-1, SNP (G/A) identificou polimorfismo nas amostras do Brasil, 11 hibridizaram com a sonda 1 (alelo G), duas amostras com a sonda 2 (alelo A) e 25 amostras com as sondas 1 e 2 (infecção mista). Todas as amostras do Camboja hibridizaram com as sondas 1 e 2. As amostras do Sri Lanka não foram tipadas com esse ensaio.

O ensaio TaqMan4 - PVX_083040, cromossomo 12, SNP (C/A) também não evidenciou polimorfismos nas amostras do Brasil: todas hibridizaram com a sonda 1 (alelo C). No Camboja, 18 amostras hibridizaram com a sonda 1 (alelo C) ; as demais amostras do Camboja e as amostras do Sri Lanka ainda não foram tipadas com este ensaio.

O ensaio TaqMan5 -PVX_123750, cromossomo 14 – SNP (T/A) apontou polimorfismos nas amostras do Brasil: 11 hibridizaram com a sonda 1 (alelo T) e 28 foram identificadas como infecção mista, hibridizando com as sondas 1 e 2. Para as amostras do Camboja, uma hibridizou com a sonda 1 (alelo T), 2 com a sonda 2 (alelo A) e as demais com as sondas 1 e 2. As amostras do Sri Lanka não foram tipadas com este ensaio.

O ensaio TaqMan6- AMA1, SNP (G/A)- não pode ser padronizado em nenhuma das amostras selecionadas mesmo com variações de concentração dos oligonucleotídeos, sondas e temperatura de anelamento.

O ensaio TaqMan7- MSP-1, SNP (A/C) - demonstrou hibridização com a sonda 1 (alelo A) em 36 amostras e com a sonda 2 (alelo C) em três amostras do Brasil. Das amostras do Camboja, 10 hibridizaram com a sonda 1 (alelo C), duas amostras com a sonda 2 (alelo A) e as demais apresentaram sinal de hibridização com as duas sondas. Todas as amostras do Sri Lanka hibridizaram com as duas sondas.

Finalmente o ensaio TaqMan 8- MSP-5 (T/C) – apontou, no Brasil, 30 amostras hibridizando com a sonda 1 (alelo T) e três amostras com as sondas 1 e 2 (infecção mista). Das amostras do Camboja, 33 apresentaram sinal de hibridização com a sonda 1 (alelo T) e três com as sondas 1 e 2. Das amostras do Sri Lanka, três hibridizaram a sonda 1 (alelo T) e duas com as sondas 1 e 2. Estes resultados estão resumidos na Tabela 11.

Tabela 11- Resultado da tipagem de SNPs pela plataforma Mastercycler Realplex gradient thermal cycler.

Ensaio	Gene	Origem das Amostras	Sonda
			Hibridizada
Taqman1	PVX_119300	Brasil, Camboja, Sri Lanka	1 e 2
Taqman2	Região Intergênica	Brasil, Camboja, Sri Lanka	1e 2
Taqman3	Pvmdr-1	Brasil, Camboja, Sri Lanka	1 e 2
Taqman4	PVX_083040	Brasil, Camboja, Sri Lanka	1
Taqman5	PVX_123750	Brasil, Camboja, Sri Lanka	1e 2 todas
Taqman6	AMA1	Brasil, Camboja, Sri Lanka	-
Taqman7	MSP-1	Brasil, Camboja, Sri Lanka	1 e 2
Taqman8	MSP-5	Brasil, Camboja, Sri Lanka	1

Estes resultados são oriundos de reações de PCR que não utilizaram nenhuma forma de controle positivo e a hibridação das sondas era medida pelo ponto no qual a curva passa a linha de *threshold* (limiar), quanto menor era este número, maior era considerada a parasitemia das amostras, portanto maior era a concentração de DNA do parasito para hibridizar com uma das sondas. As amostras que hibridizaram com as duas sondas podem conter dois clones distintos de *P. vivax* (infecção mista); outra explicação seria um resultado falso-positivo para uma das sondas, o que significa que uma sonda pode hibridizar com maior facilidade que a outra.

Para dar continuidade ao estudo foram selecionados os ensaios que mais se mostraram polimórficos entre as amostras de cada localidade assim como entre as amostras das

diferentes localidades. Para uma melhor interpretação desses resultados, foram desenvolvidos genes sintéticos controle. Esses controles foram utilizados também para avaliar o desempenho dos ensaios de tipagem em infecções mistas com diferentes proporções de cada alelo.

4.4 Análise dos ensaios de genotipagem de SNPs pela na plataforma StepOnePlus v.2.1

Foram utilizadas, para a padronização dos ensaios de tipagem de SNPs, um total de 55 amostras: 24 amostras do Brasil, 16 amostras do Camboja, 10 amostras do Sri Lanka e 5 amostras do CDC (de casos de malária importada), com o intuito de aumentar a probabilidade de encontro dos dois diferentes haplótipos entre as amostras testadas.

O desenho dos ensaios Taqman® utilizados neste estudo, foi adaptado do trabalho de Daniels e colaboradores, (2008) que descreveu o primeiro ensaio com 24 SNPs para a genotipagem de amostras de *P. falciparum*. Foram selecionados dez polimorfismos para a padronização de ensaios de tipo TaqMan®. Entretanto, três destes ensaios não puderam ser padronizados, mesmo depois de terem sido variados diversos parâmetros da reação (concentrações dos oligonucleotídeos, sondas e DNA, bem como a temperatura de anelamento). Os outros sete ensaios de tipagem de SNPs, entretanto, foram padronizados adequadamente, gerando dados confiáveis para análise. A Tabela 12 apresenta os sete ensaios padronizados neste estudo.

Tabela 12- Descrição dos sete ensaios TaqMan padronizados.

Ensaio	SNP	Cromossomo	Gene	Tipo de SNP
TaqMan 1	G/A	8	Proteína Hipotética PVX_119300	Sinônimo
TaqMan 3	G/A	10	PvMDR-1	Sinônimo
TaqMan 7	A/C	7	MSP-1	Sinônimo
TaqMan 9	T/C	8	Região Intergênica	Sinônimo
TaqMan 11	C/T	7	MSP-1	Sinônimo
TaqMan13	T/C	13	Intron	Sinônimo
TaqMan14	G/A	13	Intron	Sinônimo

As reações de PCR foram realizadas com um volume final de 20 µl e com concentrações de DNA que variaram de 0,07 a 0,6 pg/µl (devido a variação da parasitemia de cada amostra), quantificadas pelo PCR em tempo real quantitativo descrito no item 3.8.1. Todas as sondas foram testadas utilizando controles positivos (genes sintéticos) com homozigoto VIC (alelo 1), homozigoto FAM (alelo 2) e heterozigoto VIC/FAM (alelo 1/alelo 2), controle negativo (sem DNA) e controle com *P. falciparum* e outro com DNA humano para garantir que os *Oligos* e as sondas são específicos para amplificar somente o genoma de *P. vivax*. Uma vez amplificados, os produtos foram analisados através do programa StepOnePlus v.2.2.2 (Applied Biosystems, EUA) com a função *Allelic Discrimination Plot* ilustrada na Figura 14.

Figura 14 - Discriminação alélica, resultante do ensaio TaqMan 11.

Nota: A cor vermelha representa as amostras que hibridizaram com a sonda marcada com o fluoróforo VIC (G/G), a cor azul representa as amostras que hibridizaram com a sonda marcada com o fluoróforo FAM (A/A), cor verde representa as amostras que hibridizaram com as sondas marcadas com os fluoróforos sondas marcadas com VIC e FAM (G/A), a cor preta são as amostras negativas ou indeterminadas.

O ensaio TaqMan1- Proteína Hipotética PVX_119300, SNP (G/A)- mostrou-se pouco polimórfico, uma vez que evidenciou a hibridação de todas as amostras do Brasil, do Camboja, do Sri Lanka e do CDC com a sonda marcada com o fluoróforo FAM, que possuía o

alelo A. Nenhuma amostra hibridizou com a sonda marcada com o fluoróforo VIC, alelo G e três amostras do Sri Lanka não puderam ser genotipadas (o motivo pode ter sido a baixa concentração da parasitemia das amostras).

O ensaio TaqMan3- PvMDR-1, SNP (G/A), também mostrou-se pouco polimórfico, ele identificou a hibridação de um total de 47 amostras com a sonda de alelo A, marcada com o fluoróforo FAM, sendo 20 delas do Brasil, 16 do Camboja, sete do Sri Lanka e quatro do CDC e somente uma amostra do Brasil hibridizou com a sonda de alelo G (fluoróforo VIC) e não foi possível a tipagem de sete amostras, três do Brasil, três do Sri Lanka e uma do CDC.

No ensaio TaqMan7 - MSP-1, SNP (A/C) – foi evidenciada a hibridação de 36 amostras com a sonda de alelo C, marcada com o fluoróforo (FAM), sete todas as amostras hibridizaram com a sonda de alelo A (marcada com o fluoróforo VIC), neste ensaio foram identificadas dez amostras com infecção mista, ou seja, amostras que hibridizaram com as duas sondas, nove delas do Camboja e uma do Sri Lanka. Uma amostra do Sri Lanka não pode ser tipada.

No ensaio TaqMan9- Região Intergênica, cromossomo 8, SNP (T/C) - foi identificada a hibridação de 27 amostras com a sonda que possuía o alelo C (marcada com o fluoróforo FAM), sendo 18 do Brasil, seis do Camboja, todas as dez do Sri Lanka e três do CDC. Ainda, 14 amostras hibridizaram com a sonda marcada com VIC (alelo T), sendo cinco do Brasil, oito do Camboja e uma do CDC. Este ensaio foi capaz de identificar quatro infecções mistas, uma do Brasil, duas do Camboja e uma do CDC.

O ensaio TaqMan11- MSP-1, SNP (C/T) - detectou a hibridação de 24 amostras com a sonda de alelo C (marcada com o fluoróforo VIC) um total de dez amostras do Brasil, 11 do Camboja, duas do Sri Lanka e uma do CDC. Também identificou que 14 amostras hibridizaram com a sonda T (marcada com o fluoróforo FAM), sete do Brasil, três do Camboja, seis do Sri Lanka e três do CDC. Este ensaio também detectou quatro amostras com infecção mista do Brasil, duas do Camboja. Cinco amostras falharam na tipagem, não hibridizando com nenhuma das duas sondas.

O ensaio TaqMan13 – Intron, SNP (T/C) – apontou a hibridação de 35 amostras com a sonda de alelo T (marcada com o fluoróforo VIC), 14 do Brasil, 16 do Camboja, oito do Sri Lanka e cinco do CDC. Já com a sonda de alelo C (marcada com o fluoróforo FAM) houve a hibridação de oito amostras do Brasil, duas do Sri Lanka, e este ensaio não identificou nenhuma infecção mista.

Finalmente, o ensaio TaqMan14 -Intron, SNP (C/A) – constatou a hibridação de 13 amostras brasileiras com a sonda de alelo A (marcada com o fluoróforo FAM) e todas as amostras do Camboja, do Sri Lanka e do CDC também, com a sonda de alelo G (marcada com o fluoróforo VIC) houve a hibridação de seis amostras brasileiras e cinco hibridizaram com as duas sondas, revelando serem infecções mistas. Estes resultados revelam ser este um ensaio pouco polimórfico.

Os ensaios TaqMan10 e TaqMan12 não puderam ser padronizados, mesmo após a variação de diversos parâmetros da reação de PCR (concentrações dos oligonucleotídeos, sondas e DNA, bem como a temperatura de anelamento). Todos estes resultados estão descritos na Tabela 13.

Tabela 13-Resultado da tipagem de SNPs. Identificação de gene, cromossomo e localização cromossômica.

Origem	Amostra	Proteína Hipotética PVX_119300		PvMDR-1		MSP-1		Região Intergênica		MSP-1		Intron		Intron	
		Cr 08-1577813		Cr10-362033		Cr 07-1158166		Cr 08-1584352		Cr 07-1160451		Cr 13-1753649		Cr13-1753768	
		G	A	G	A	A	C	T	C	C	T	T	C	G	A
Brasil	4V		X		X		X		X		X	X			X
Brasil	12V		X		X		X		X	X		X			X
Brasil	20V		X		X	X			X	X		X			X
Brasil	21V		X		X		X		X	X		X			X
Brasil	22V		X		X		X		X		X	X			X
Brasil	23V		X		X		X	X		X	X		X	X	
Brasil	24V		X		X	X			X	X			X	X	
Brasil	25V		X		X		X		X	X		X			X
Brasil	26V		X		X		X		X		X	X			X
Brasil	27V		X		X		X		X	X		X	X	X	X
Brasil	28V		X	X			X		X	X	X	X	X	X	X
Brasil	42V		X		X		X		X	X	X	X			X
Brasil	35V		X		IND		X	X		IND			X	X	X
Brasil	52V		X		X		X		X	X	X	X			X
Brasil	53V		X		X		X		X		X	X		X	X
Brasil	56V		X		X		X		X	X			X		X
Brasil	58V		X		X		X	X	X	X		X		X	X
Brasil	77V		X		X		X		X	X			X	X	
Brasil	9FBO		X		X		X		X		X	X			X
Brasil	10MBS		X		X	X		X		X		X			X
Brasil	11JIS		X		X		X		X		X		X		X
Brasil	24MS		X		X		X		X		X		X	X	
Brasil	28DS		X		IND		X	X		IND		X		X	
Brasil	49ACL		X		IND		X	X		IND			X	X	
Camboja	PV002		X		X	X	X		X	X		X			X
Camboja	PV006		X		X		X	X	X	X		X			X

Tabela 13-Resultado da tipagem de SNPs. Identificação de gene, cromossomo e localização cromossômica.

(Conclusão) 100

Origem	Amostra	Proteína Hipotética PVX_119300 Cr 08-1577813		PvMDR-1 Cr10-362033		MSP-1 Cr 07-1158166		Região Intergênica Cr 08-1584352		MSP-1 Cr 07-1160451		Intron Cr 13-1753649		Intron Cr13-1753768	
		G	A	G	A	A	C	T	C	C	T	T	C	G	A
Camboja	PV007		X		X		X		X	X		X			X
Camboja	PV016		X		X	X	X	X		X	X	X			X
Camboja	PV018		X		X	X			X		X	X			X
Camboja	PV020		X		X	X	X	X	X	X	X	X			X
Camboja	PV024		X		X	X			X		X	X			X
Camboja	PV029		X		X	X		X		X		X			X
Camboja	PV030		X		X		X	X		X		X			X
Camboja	PV034		X		X	X	X		X	X		X			X
Camboja	PV037		X		X	X	X	X		X		X			X
Camboja	PV040		X		X	X	X	X	X	X		X			X
Camboja	PV045		X		X	X			X		X	X			X
Camboja	PV048		X		X	X	X	X		X		X			X
Camboja	PV050		X		X	X	X	X		X		X			X
Camboja	PV060		X		X	X	X	X		X		X			X
SriLanka	SL16		IND		IND		IND		X		X	X			X
SriLanka	SL27		IND		IND		X		X		X	X			X
SriLanka	SL29		X		X		X		X		X	X			X
SriLanka	SL32		X		X		X		X		X	X			X
SriLanka	SL33		X		X		X		X		X	X			X
SriLanka	SL35		X		X		X		X		X	X			X
SriLanka	SL36		X		X		X		X	X			X		X
SriLanka	SL38		IND		X		X		x	X			X		X
SriLanka	SL41		X		X		X		X	IND		X			X
SriLanka	SL44		X		IND		X		X		X	X			X
CDC	13CDC		X		X	X			X		X	X			X
CDC	29CDC		X		X		X	X		IND		X			X
CDC	30CDC		X		IND	X	X	X	X		X	X			X
CDC	49CDC		X		X		X		X		X	X			X
CDC	56CDC		X		X		X		X	X		X			X

4.5 Análise de haplótipos

Os resultados obtidos com a tipagem dos polimorfismos com a metodologia TaqMan® apontaram haver grande variação genética no conjunto de amostras analisadas do Brasil, Camboja, Sri Lanka e CDC. Foi obtido um total de 24 haplótipos distintos com os sete ensaios tipados, considerou-se cada infecção mista não idêntica, como um haplótipo distinto, como demonstra a tabela 14. As amostras que não puderam ser tipadas por todos os ensaios não foram adicionadas a esta contagem.

Tabela 14- Haplótipos resultantes das tipagens dos sete ensaios TaqMan.

Origem	Nº de Amostras	Haplótipo						
Brasil	1	A	A	C	C	T	C	G
Brasil, CDC	4	A	A	C	C	T	T	A
Brasil, Camboja, Sri Lanka	4	A	A	C	C	C	T	A
Brasil	1	A	A	A	C	C	C	G
Brasil	1	A	A	A	C	C	T	A
Brasil	2	A	A	C	C	C	C	A
Brasil	1	A	A	C	C	T	C	A
Brasil	1	A	A	A	T	T	T	A
Camboja	1	A	A	A	C	C	T	A
Brasil, Camboja, Sri Lanka, CDC	4	A	A	C	C	C	T	A
Camboja, CDC	2	A	A	A	C	T	T	A
Camboja	1	A	A	A	C	T	T	A
Camboja	5	A	A	A/C	T	C	T	A
Camboja	1	A	A	A/C	C	C	T	A
Camboja	1	A	A	A/C	T/C	C	T	A
Camboja	1	A	A	A/C	T/C	C/T	T	A
Brasil	1	A	A	C	T/C	C	T	G/A
Camboja	1	A	A	A	T/C	C	T	A
Brasil	3	A	A	C	C	C/T	T	A
Brasil	2	A	A	C	T	C/T	C	G
Brasil	1	A	G	C	C	C/T	T/C	G/A
Brasil	1	A	A	C	C	C	T/C	G/A
Brasil	1	A	A	C	C	T	T	G/A
Brasil	1	A	G	C	C	C/T	T/C	G/A

A tipagem dos sete polimorfismos pela metodologia TaqMan identificou um total de oito

infecções mistas nas amostras do Brasil, nove nas amostras do Camboja, uma nas amostras do CDC e nenhuma nas amostras do SriLanka. Estes resultados estão descritos na Tabela 15.

Tabela 15: Resultado da tipagem de polimorfismos pela metodologia TaqMan® , destacando a formação dos diferentes haplótipos e a detecção de infecções mistas.

(continua)

Origem	Amostra	Haplótipo							I. Mista
Brasil	4V	A	A	C	C	T	T	A	N
Brasil	12V	A	A	C	C	C	T	A	N
Brasil	20V	A	A	A	C	C	T	A	N
Brasil	21V	A	A	C	C	C	T	A	N
Brasil	22V	A	A	C	C	T	T	A	N
Brasil	23V	A	A	C	T	C/T	C	G	S
Brasil	24V	A	A	A	C	C	C	G	N
Brasil	25V	A	A	C	C	C	T	A	N
Brasil	26V	A	A	C	C	T	T	A	N
Brasil	27V	A	A	C	C	C	T/C	G/A	S
Brasil	28V	A	G	C	C	C/T	T/C	G/A	S
Brasil	42V	A	A	C	C	C/T	T	A	S
Brasil	35V	A	IND	C	T	T	IND	G/A	S
Brasil	52V	A	A	C	T	C/T	T	A	S
Brasil	53V	A	A	C	C	T	T	G/A	S
Brasil	56V	A	A	C	C	C	C	A	N
Brasil	58V	A	A	C	T/C	C	T	G/A	S
Brasil	77V	A	A	C	C	C	C	A	N
Brasil	9FBO	A	A	C	C	T	T	A	N
Brasil	10MBS	A	A	A	T	T	T	A	N
Brasil	11JIS	A	A	C	C	T	C	A	N
Brasil	24MS	A	A	C	C	T	C	G	N
Brasil	28DS	A	IND	C	T	IND	T	G	N
Brasil	49ACL	A	IND	C	T	IND	C	G	N
Camboja	PV002	A	A	A	C	C	T	A	N
Camboja	PV006	A	A	A	T/C	C	T	A	S
Camboja	PV007	A	A	C	C	C	T	A	N
Camboja	PV016	A	A	A/C	T	C/T	T	A	S
Camboja	PV018	A	A	A	C	T	T	A	N
Camboja	PV020	A	A	A/C	T/C	C/T	T	A	S
Camboja	PV024	A	A	A	C	T	T	A	N
Camboja	PV029	A	A	C	T	C	T	A	N

Tabela 15: Resultado da tipagem de polimorfismos pela metodologia TaqMan® , destacando a formação dos diferentes haplótipos e a detecção de infecções mistas.

(conclusão)

Origem	Amostra	Haplótipo						I. Mista	
Camboja	PV030	A	A	C	T	C	T	A	N
Camboja	PV034	A	A	A/C	C	C	T	A	S
Camboja	PV037	A	A	A/C	T	C	T	A	S
Camboja	PV040	A	A	A/C	T/C	C	T	A	S
Camboja	PV045	A	A	A	C	T	T	A	N
Camboja	PV048	A	A	A/C	T	C	T	A	S
Camboja	PV050	A	A	A/C	T	C	T	A	S
Camboja	PV060	A	A	A/C	T	C	T	A	S
SriLanka	SL16	IND	IND	IND	C	T	T	A	N
SriLanka	SL27	IND	IND	C	C	T	T	A	N
SriLanka	SL29	A	A	C	C	T	T	A	N
SriLanka	SL32	A	A	C	C	T	T	A	N
SriLanka	SL33	A	A	C	C	T	T	A	N
SriLanka	SL35	A	A	C	C	T	T	A	N
SriLanka	SL36	A	A	C	C	T	C	A	N
SriLanka	SL38	IND	A	C	C	C	C	A	N
SriLanka	SL41	A	A	C	C	C	T	A	N
SriLanka	SL44	A	IND	C	C	IND	T	A	N
CDC	13CDC	A	A	A	C	T	T	A	N
CDC	29CDC	A	A	C	T	T	IND	A	N
CDC	30CDC	A	IND	A/C	T/C	T	T	A	S
CDC	49CDC	A	A	C	C	T	T	A	N
CDC	56CDC	A	A	C	C	C	T	A	N

Nota: IND= alelo não determinado pela tipagem. S= presença de infecção mista. N= ausência de infecção mista.

4.6 Análise de tipagem molecular por microssatélites

Foram utilizadas para esta análise o mesmo total de 48 amostras utilizadas na tipagem de polimorfismos: 17 amostras do Brasil, 16 amostras do Camboja, 10 amostras do Sri Lanka e 5 amostras do CDC. Resultado semelhante ao encontrado com a tipagem de SNPs com a metodologia TaqMan foi observado quando foram analisados os dados previamente obtidos com a tipagem de MSAT destas mesmas amostras, no entanto, observou-se maior

diversidade genética com os MSAT, por exemplo, nenhuma amostra teve o mesmo haplótipo de microssatélite compartilhado. Estes resultados estão ilustrados na Tabela 15.

Outros resultados também puderam ser inferidos desta análise de MSAT, o número de infecções mistas, aquelas nas quais pelo menos um dos marcadores testados (todos eles presentes em cópia única no genoma nuclear de *P. vivax*) apresenta mais de um alelo, indicando a co-existência de dois ou mais clones geneticamente distintos na amostra, foi maior na tipagem de MSAT do que na tipagem pela metodologia TaqMan.

Neste estudo, a tipagem de MSAT revelou 41% de infecções mistas nas amostras brasileiras, 93% de infecções mistas nas amostras do Camboja, nenhuma infecção mista nas amostras do SriLanka e 20% de infecções mistas nas amostras de diversas localidades do CDC. Na tipagem de SNPs com a metodologia TaqMan® este número foi de 33% de infecções nos isolados brasileiros, 50% de infecções mistas nos isolados do Camboja, nenhuma infecção mista nos isolados do SriLanka e 20% de infecções mistas nos isolados de diversas localidades do CDC. O motivo para este resultado discrepante pode ser analisado pelo número de marcadores utilizados em cada metodologia de tipagem. Foram sete polimorfismos tipados pela metodologia Taqman e 13 marcadores de DNA tipados por microssatélites.

Tabela 16- Resultado da tipagem molecular com 13 microssatélites em isolados de *P. vivax*.

(continua)

		Microssatélites <i>Plasmodium vivax</i>														
Origem	Amostra	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M12	M15	M20	I. M.	
Brasil	4V	240	198	188	192	178	240	145	282	173	192	224	233	210	S	
Brasil	12V	240	198	188	198	226	240	145	333	158	192	215	248	213	N	
Brasil	20V	243	198	185	198	226	210	145	261	158	192	215	233	210	S	
Brasil	21V	237	198	191	195	211	246	142	261	158	192	215	251	207	N	
Brasil	22V	243	202	182	198	205	210	145	294	161	240	242	245	204	S	
Brasil	23V	237	194	185	216	175	210	145	216	173	240	215	251	198	S	
Brasil	24V	240	202	191	210	205	240	142	294	161	192	212	233	213	N	
Brasil	25V	240	202	188	216	169	210	145	294	161	192	224	233	201	N	
Brasil	26V	237	198	185	198	205	210	145	294	161	240	215	251	201	N	
Brasil	27V	240	202	188	216	166	210	145	294	161	192	224	236	201	S	

Tabela 16- Resultado da tipagem molecular com 13 microsatélites em isolados de *P. vivax*.
(conclusão)

		Microsatélites <i>Plasmodium vivax</i>													
Origem	Amostra	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M12	M15	M20	I. M.
Brasil	28V	240	198	188	195	178	240	145	276	170	192	224	233	210	S
Brasil	42V	237	198	185	192	175	210	145	276	161	192	215	251	225	S
Brasil	35V	234	186	191	198	187	255	145	213	155	198	218	248	213	N
Brasil	53V	234	170	188	198	154	258	166	297	164	168	197	203	171	N
Brasil	56V	237	210	191	213	175	249	142	237	167	177	197	254	207	N
Brasil	58V	234	222	188	204	202	261	196	258	155	216	210	296	213	N
Brasil	77V	228	198	191	198	175	249	148	258	155	201	212	245	198	N
Camboja	PV002	234	214	185	204	172	237	139	246	152	171	218	260	177	S
Camboja	PV006	237	214	188	195	166	252	151	231	140	201	191	236	195	S
Camboja	PV007	237	198	191	198	160	258	148	255	146	213	215	275	213	S
Camboja	PV016	234	182	188	222	166	210	154	234	152	201	212	242	213	S
Camboja	PV018	231	182	185	192	181	249	145	297	158	204	203	242	216	S
Camboja	PV020	234	182	188	198	172	246	157	285	152	198	210	242	243	S
Camboja	PV024	237	174	185	204	172	246	178	288	164	171	206	239	210	S
Camboja	PV029	237	222	185	207	172	210	160	279	158	168	206	242	204	S
Camboja	PV030	237	198	197	198	175	246	157	210	176	195	218	254	216	S
Camboja	PV034	237	186	191	198	169	252	151	252	167	168	215	242	210	S
Camboja	PV037	225	226	191	198	175	258	157	273	152	201	212	239	210	S
Camboja	PV040	231	198	191	198	178	249	148	273	164	180	236	245	237	S
Camboja	PV045	237	202	191	186	163	240	148	249	167	186	206	242	189	S
Camboja	PV048	234	198	191	201	166	249	157	264	155	213	227	242	213	S
Camboja	PV050	234	238	191	207	184	249	178	249	167	180	218	245	216	N
Camboja	PV060	234	246	194	198	190	249	181	273	152	186	212	245	216	N
SriLanka	SL16	243	214	188	195	169	252	145	249	164	291	224	242	231	N
SriLanka	SL27	231	214	188	192	163	237	139	297	146	222	221	242	174	N
SriLanka	SL29	231	214	182	192	163	237	139	297	146	222	221	245	174	N
SriLanka	SL32	228	222	182	189	166	234	145	324	164	177	206	254	171	N
SriLanka	SL33	222	214	182	192	163	237	139	297	146	222	221	248	174	N
SriLanka	SL35	234	218	182	189	166	234	139	324	164	177	206	254	171	N
SriLanka	SL36	228	210	176	198	163	240	133	306	161	195	212	242	204	N
SriLanka	SL38	222	206	176	189	166	240	145	318	158	177	203	245	264	N
SriLanka	SL41	234	222	182	189	166	234	151	324	161	177	200	260	171	N
SriLanka	SL44	234	214	182	189	166	234	151	324	164	177	200	254	171	N
CDC	13CDC	240	186	185	198	175	246	145	276	170	234	200	260	222	N
CDC	29CDC	240	214	185	192	172	249	145	201	155	207	182	263	198	N
CDC	30CDC	237	206	185	195	175	234	142	294	161	195	218	251	246	N
CDC	49CDC	231	218	194	210	175	252	130	246	155	210	236	254	219	N
CDC	56CDC	234	210	194	192	178	231	145	201	161	210	218	248	183	S

Nota: I. M. = infecção mista. S= sim (presença); N= não (ausência)

5.1 Marcadores moleculares – microarranjos

No início deste estudo, somente um genoma de *P. vivax* havia sido publicado (Carlton et al., 2008) e um único conjunto limitado de SNPs ao longo de um segmento de 200 kb do cromossomo 8 dessa espécie havia sido caracterizado (Feng et al., 2003) o que dificultava muito estudos populacionais através da caracterização e tipagem de SNPs.

Por este motivo adotamos inicialmente a estratégia de identificar polimorfismos em isolados de *P. vivax* de amostras de campo brasileiras que haviam sido submetidos à hibridação com microarranjos contendo o genoma de Sal-I realizada por Westenberger e colaboradores (2010). No mesmo momento que os primeiros resultados de identificação de polimorfismos nos isolados Acre5, Acre9 e Acre10 eram analisados, Dharia e colaboradores publicaram em 2010 o sequenciamento genômico completo de um isolado de campo de *P. vivax* (IQ7) proveniente de Iquitos/Peru. Como o DNA deste isolado havia sido submetido à hibridação com o mesmo microarranjo contendo o genoma de Sal-I, foi possível verificar que os possíveis polimorfismos identificados pela estratégia de hibridação não puderam ser confirmados pelo sequenciamento.

Ainda segundo Dharia, nos genes mais polimórficos, incluindo aqueles com regiões repetitivas de sequenciamento mais difícil, com a estratégia de hibridação foram identificados um número *muito superior* de SNPs do que com o sequenciamento de DNA. Por exemplo, dos 51 polimorfismos identificados por hibridação no gene *MSP-3* (PVX_097685), somente sete foram confirmados por sequenciamento. Os números referentes a outros genes polimórficos são 107 identificados e quatro confirmados para *Reticulocyte Binding Protein 2* (PVX_090325) e 41 identificados e quatro confirmados para *Reticulocyte Binding Protein 1* (PVX_098585). No entanto, nas demais regiões genômicas a estratégia de hibridação identificou *menor número* de polimorfismos por gene do que o sequenciamento genômico, sugerindo a ocorrência de grande número de falso-negativos.

Portanto, esta estratégia não fornece resultados úteis para a seleção de SNPs para posterior padronização e tipagem de isolados. Este motivo fez a nossa primeira estratégia de identificação de polimorfismos ser descartada.

5.2 Sequências de DNA de genes de cópia única disponíveis no GenBank e SNPs previamente descritos na literatura

Para a segunda estratégia de identificação e caracterização de SNPs, adotou-se então duas abordagens alternativas, a análise de sequências de DNA de genes de cópia única disponíveis no GenBank e a análise de SNPs previamente descritos na literatura.

Destas análises foram selecionados 14 polimorfismos para serem tipados em amostras de diferentes localidades endêmicas de malária vivax. Foram identificadas substituições sinônimas com boa distribuição geográfica nos seguintes genes: (a) três no *MSP-1* - cromossomo 7; (b) uma no *AMA-1* - cromossomo 9; (c) uma no *MSP* 5-cromossomo 4; (d) uma substituição não sinônima no *DBP* - cromossomo 6; Foram também selecionados os SNPs previamente descritos na literatura: (e) uma substituição sinônima em uma proteína hipotética (PVX-119300) duas substituições em uma região intergênica - cromossomo 8 (f) uma substituição sinônima no gene *MDR-1*, cromossomo 10; (g) duas substituições sinônimas identificadas por Dharia et al. (2010) em genes que codificam proteínas hipotéticas (PVX_083040, cromossomo 12, e PVX_123750, cromossomo 14); (h) duas substituições sinônimas identificadas em isolados indianos por Gupta et al. (2010) – Introns, cromossomo 13.

5.3 Tipagem molecular de SNPs

Para iniciar a seguinte etapa foram desenvolvidos 14 ensaios do tipo Taqman® com base no estudo de Daniels e colaboradores (2008) no qual foi descrito o primeiro ensaio de tipagem

molecular para *P. falciparum* utilizando sondas MGB- Taqman (Minor Groove Binder). Dos 112.000 SNPs descritos, foram selecionados 2.100 SNPs os quais eram comuns nos isolados analisados. Foram inicialmente desenvolvidos 88 ensaios com sondas TaqMan®, dos quais apenas 24 foram escolhidos para serem utilizados na identificação e localização geográfica de parasitos de *P. falciparum*.

Neste estudo foram utilizadas duas metodologias distintas para a tipagem dos SNPs selecionados: (a) análise da genotipagem de SNPs na plataforma Mastercycler Realplex gradient thermal cyler; (b) análise dos ensaios de genotipagem de SNPs pela na plataforma StepOnePlus v.2.1. As principais diferenças entre as metodologias estão na forma de obtenção dos resultados e na validação dos mesmos. Na primeira metodologia a hibridação das sondas era medida pelo ponto no qual a curva passa a linha de *treshold* (limiar), quanto menor era este número, maior era considerada a parasitemia das amostras. As amostras que hibridizaram com as duas sondas poderiam conter dois clones distintos de *P. vivax* (infecção mista); ou um resultado falso-positivo para uma das sondas, o que significa que uma sonda pode hibridizar com maior facilidade que a outra. Esta metodologia também não contou com controles positivos.

Com a análise dos ensaios de genotipagem de SNPs pela na plataforma StepOnePlus v.2.1 (segunda metodologia) foram novamente tipados os ensaios que se mostraram mais polimórficos com a primeira metodologia (TaqMan1, TaqMan3, TaqMan7), foram adicionados os demais ensaios e ainda foram incluídos os controles positivos oriundos de genes sintéticos controle. O resultado da tipagem das amostras do estudo pelos ensaios TaqMan1, TaqMan3 e TaqMan7 foi semelhante em ambas metodologias, mas a segunda metodologia mostrou-se superior à primeira, especialmente pela forma de expor os resultados por discriminação alélica, que não deixa dúvida sobre qual sonda foi hibridizada.

Dos 14 ensaios selecionados, sete puderam ser padronizados: TaqMan1-Proteína Hipotética PVX_119300, TaqMan3-PvMDR-1, TaqMan7-MSP-1, TaqMan9-Região Intergênica, TaqMan11-MSP-1, TaqMan13-Intron e TaqMan14- Intron. Com estes resultados foi possível

formar uma rede de haplótipos que caracteriza e diferencia cada amostra entre si e agrupa amostras de mesmo haplótipo. Os ensaios TaqMan1, Taqman2, mostraram-se pouco polimórficos e os demais ensaios foram mais polimórficos, no entanto não foram capazes de agrupar as amostras de acordo com sua localização geográfica. Um resultado diferente é apresentado no estudo de Conway et al. (2000) no qual foi desenvolvida uma técnica simples de genotipagem com apenas quatro SNPs mitocondriais, para agrupar geograficamente 104 isolados de *P. falciparum* de origem brasileira, africana e tailandesa. Como resultado da análise destes dados foram identificados a presença de cinco haplótipos distintos nestas populações e este com este número de marcadores altamente informativos (do ponto de vista genético) foi possível diferenciar geneticamente os isolados de diferentes regiões geográficas.

5.4 Análise de SNPs comparada à análise de microssatélites

Apesar do diferente número de marcadores (sete para SNPs e 13 para MSAT), obteve-se um resultado semelhante na tipagem de ambos marcadores moleculares na mesma população de amostras analisadas neste estudo: grande diversidade genética. Porém, esta diversidade encontrada foi maior com os MSATs, devido ao maior número de marcadores. O número de infecções mistas também foi maior na tipagem de MSAT do que na tipagem pela mitologia TaqMan.

Em nossa área de estudo do Brasil, cerca de 40% das infecções por *P. vivax* contêm mais de um clone geneticamente distinto, identificado por tipagem de MSATs (Ferreira et al., 2007; Orjuela-Sánchez et al., 2009). Neste estudo, a tipagem SNPs com a metodologia TaqMan® obteve um valor semelhante de 33% de infecções mistas nos isolados brasileiros, o que valida bem a comparação entre os dois marcadores moleculares.

Com os resultados obtidos neste estudo, pode-se concluir que:

Unindo os dados obtidos neste estudo àqueles publicados por Dharia *et al.* (2010), sugere-se que a hibridação com o microarranjo contendo o genoma completo de Sal-I não produz resultados confiáveis para a identificação de possíveis SNPs em isolados de campo de *P. vivax*.

O conjunto inicial de sete SNPs nucleares que foram tipados com sucesso pode não ser suficientemente informativo para a realização de estudos populacionais em nossa área de estudo na Amazônia Ocidental brasileira, bem como em outras áreas de transmissão de *P. vivax* na Ásia. Por esse motivo, nosso laboratório vem buscando dados adicionais que possam resultar na descoberta de novos SNPs e na padronização subsequente de métodos de tipagem.

Além da escassa variabilidade genética entre os marcadores utilizados, outros achados relevantes desse estudo são: (a) a alta proporção de infecções geneticamente mistas observada no Camboja, dificultando a definição de haplótipos entre parasitos desse país e (b) a alta proporção de isolados incompletamente tipados do Sri Lanka, possivelmente em função das baixas parasitemias observadas nessas infecções.

Em agosto de 2012, 10 novos genomas de *P. vivax* tiveram sua sequência publicada (Chan et al., 2012; Neafsey et al., 2012), aumentando substancialmente a disponibilidade de informação para a descoberta e caracterização de SNPs utilizáveis em estudos populacionais.

Além disso, iniciamos um projeto extenso de caracterização de genomas mitocondriais de *P. vivax*, que vem resultando em um amplo repertório de novos polimorfismos com uso potencial como marcadores genéticos.

Alves FP. Epidemiologia da Malária em Populações Nativas da Amazônia Brasileira: a Ocorrência de Infecções Assintomáticas. [tese (Doutorado em Ciências)]. São Paulo: Instituto de Ciências Biomédicas da Universidade de São Paulo; 2003.

Anderson TJC, Haubold B, Williams JT, Estrada-Franco JG, Chardson L, Mollinedo R, Bockarie M, Mokili J, Mharakurna S, French N, Whiworth J, Velez ID, Brochman AH, Nosten F, Ferreira MU, Day KP. Microsatellite markers reveal a spectrum of population structures in the malaria parasite *Plasmodium falciparum*. *Mol. Biol. Evol.* 2000; v. 17, p. 1467-1482.

Anderson TJC, Nair S, Sudimack D, Williams JT, Mayxay M, Newton PN, Guthmann JP, Smithuis FM, Tran TH, van den Broek IV, White NJ, Nosten F. Geographical distribution of selected and putatively neutral SNPs in Southeast Asian malaria parasites. *Mol Biol Evol* 2005; 22: 2362-2374.

Andrade BB, Reis-Filho A, Souza-Neto SM, Clarencio J, Camargo LM, Barral A, Barral-Netto M. Severe *Plasmodium vivax* malaria exhibits marked inflammatory imbalance. *Malar J.* 2010; 9:13-17.

Antinori S, Milazzo L, Ridolfo AL, Galimberti L, Corbellino M. Severe *Plasmodium vivax*: fact or fiction? *Clin Infect Dis.* 2012; v. 43: 1256-64. doi: 10.1093/cid/cis709.

Bartoloni A, Zammarchi L. Clinical aspects of uncomplicated and severe malaria. Review Articles. *Open Journal System. Mediterr J Hematol Infect Dis.* 2012; 4 (1).

Brega S, Meslin B, de Monbrison F, Severini C, Gradoni L, Udonsangpetch R, Sutano I, Peyron F, Picot S. Identification of the *Plasmodium vivax* mdr-like gene (*pvm-dr-1*) and analysis of single nucleotide polymorphisms among isolates from different areas of endemicity. *J Infect Dis.* 2005; 191:272-277.

Camargo AA, Fisher K, Lanzer M, del Portillo HA. Construction and characterization of a *Plasmodium vivax* genomic library in yeast artificial chromosomes. *Genomics.* 1997;42:467-473.

* De acordo com: International Committee of Medical Journal Editors. Uniform requirements for manuscripts submitted to Biomedical Journal: sample references. [Cited 2011 Jul 15]. Available from: <http://www.icmje.org>

Carlton JM, Adams JH, Silva JC, Bidwell SL, Lorenzi H, Caler E, Crabtree J, Angiuoli SV, Merino EF, Amedeo P, Cheng Q, Coulson RM, Crabb BS, del Portillo HA, Essien K, Feldblyum TV, Fernandez-Becerra C, Gilson PR, Gueye AH, Guo X, Kang'a S, Kooij TW, Korsinczky M, Meyer EV, Nene V, Paulsen I, White O, Ralph SA, Ren Q, Sargeant TJ, Salzberg SL, Stoeckert CJ, Sullivan SA, Yamamoto MM, Hoffman SL, Wortman JR, Gardner MJ, Galinski MR, Barnwell JW, Fraser-Liggett CM. Comparative genomics of the neglected human malaria parasite *Plasmodium vivax*. *Nature*. 2008; 455:757-763.

Carter R. Speculations on the origins of *Plasmodium vivax* malaria. *Trends in Parasitol*. 2003; v.19 p. 214-219.

Chan ER, Menard D, David PH, Ratsimbaoa A, Kim S, Chim P, Do C, Witkowski B, Mercereau-Puijalon O, Zimmerman PA, Serre D. Whole genome sequencing of field isolates provides robust characterization of genetic diversity in *Plasmodium vivax*. *PLoS Negl Trop Dis*. 2012; 6(9):e1811.

Chaudry JZ, Ahmed S, Qureshi TZ, Ali N. Real time polymerase chain reaction for the detection of malarial parasite. *J Coll Physicians Surg Pak*. 2012; Feb;22(2):98-100. doi: 02.2012/JCPSP.90100.

Coatney, RG, Collins WE, McWilson W, Contaco PG. The primate malarias. Atlanta: CDC, 2003; Ga [original book CD-ROM published 1971, U.S. Government Printing Office, Washington, D.C.].

Conway DJ, Fanello C, Lloyd JM, Al-Joubori BMAS, Baloch AH, Somanath SD, Roper C, Oduola AMJ, Mulder B, Pova MM, Singh B, Thomas AW. Origin of *Plasmodium falciparum* malaria is traced by mitochondrial DNA. *Mol. Biochem. Parasitol*. 2000; v. 111: p. 163-171.

Conway DJ. Molecular Epidemiology of Malaria. *Clinical Microbiology Reviews*. 2007; 188–204.

Cormier LA. The historical ecology of human and wild primate malarias in the new world. *Diversity*. 2010; v. 2:p. 256-280.

Cornejo OE, Escalante AA. The origin and age of *Plasmodium vivax*. *Trends in Parasitology*. 2006; 22: 558-563.

Cui L, Escalante AA, Imwong M, Snounou G. The genetic diversity of *Plasmodium vivax* populations. *Trends Parasitol*. 2003; 19:220-226.

Daniels R, Volkman SK, Milner DA, Mahesh N, Neafsey CE, Park D, Rose D, Angelino E, Sabeti PC, Wirth DF, Wiegand RC. A general SNP-based molecular barcode for *Plasmodium falciparum* identification and tracking. *Malaria Journal*. 2008; 7:223.

da Silva-Nunes M, Codeço CT, Malafronte RS, da Silva NS, Juncansen C, Muniz PT, Ferreira MU. Malaria on the Amazonian frontier: transmission dynamics, risk factors, spatial distribution, and prospects for control. *Am J Trop Med Hyg*. 2008; 79: 624-635.

de Brito CFA, Ferreira MU. Molecular markers and genetic diversity of *Plasmodium vivax*. *Mem Instituto Oswaldo Cruz* 2011, 106:12-26.

de Castro MC, Singer BH. Was malaria present in the Amazon before the European Conquest? Available evidence and future research agenda. *J. Archaeol Sci*. 2005; v. 32, p. 334–340.

de Santana Filho FS, Arcanjo AR, Chehuan YM, Costa MR, Martinez-Espinosa FE, Vieira JL, Barbosa MG, Alecrim WD, Alecrim MG. Chloroquine-resistant *Plasmodium vivax*, Brazilian Amazon. *Emerg Infect Dis*. 2007; 13, 1125-1126.

Dharia NV, Bright AT, Westenberger SJ, Barnes SW, Batalov S, Kuhen K, Borboa R, Federe GC, Mc Clean CM, Vinetz JM, Neyra V, Llanos-Cuentas A, Barnwell JW, Walker JR, Winzeler EA. Whole genome sequencing and microarray analysis of ex vivo *Plasmodium vivax* reveal selective pressure on putative drug resistance genes. *Proc Natl Acad Sci USA*. 2010; 107: 20045-20050.

Divis PSC, Shokoples SE, Sing B, Yanow SK. A TaqMan real-time PCR assay for the detection and quantitation of *Plasmodium knowlesi*. *Malaria Journal*. 2010; 9:344.

Ellegren H. Microsatellites: simple sequences with complex evolution. *Nature Reviews*. 2004; v: 5- 435-45.

Escalante, AA, Cornejo OE, Rojas A, Udhayakumar V, Lal AA. Assessing the effect of natural selection in malaria parasites. *Trends Parasitol*. 2004; 20: 388-395.

Feng X, Carlton JM, Joy DA, Mu J, Furuya T, Suh BB, Wang Y, Barnwell JW, Su X-S. Single-nucleotide polymorphism and genome diversity in *Plasmodium vivax*. *Proc Natl Acad Sci USA*. 2003; 100, 8502-8507.

Ferreira MU, Karunaweera NK, da Silva-Nunes M, Silva NS, Wirth DF, Hartl DL. Population structure and transmission dynamics of *Plasmodium vivax* in rural Amazonia. *J Infect Dis.* 2007; 195: 1218-1226.

Ferreira MU, Silva-Nunes M. Evidence-based public health and prospects for malaria control in Brazil. *J Infect Dev Ctries.* 2010; 4(9):533-545.

Genbank Statistics. Bethesda: National Center Biotechnology Information, 2008. Disponível em: <http://www.ncbi.nlm.nih.gov/genbank/genbankstats-2008/>. Acesso em: 12 Jan 2011

Gentil F, Bargieri DY, Leite JA, Françoso KS, Patrício MBM, Espíndola NM, Vaz AJ, Palatnik-de-Sousa CB, Rodrigues MM, Costa FTM, Soares IS. A recombinant vaccine based on domain II of *Plasmodium vivax* Apical Membrane Antigen 1 induces high antibody titres in mice. *Vaccine.* 2010; 31;28(38):6183-90.

Gómez JC, McNamara T, Bockarie MJ, Baird JK, Carlton JM, Zimmerman PA. Identification of a polymorphic *Plasmodium vivax* polymorphic marker. *Am J Trop Med Hyg.* 2003; 69, 377-389.

Gonçalves RM, Salmazi KC, Santos BAN, Bastos MS, Rocha SC, Boscardin SB, Silber AM, Kalla's EG, Ferreira MU and Scopel KKG. CD4_ CD25_ Foxp3_ Regulatory T Cells, Dendritic Cells, and Circulating Cytokines in Uncomplicated Malaria: Do Different Parasite Species Elicit Similar Host Responses? *Infect. Immun.* 2010; v. 78, n.11, p. 4763-4772.

Guerra CA, Howes RE, Patil AP, Gething PW, Van Boeckel TP, Temperley WH, Kabaria CW, Tatem AJ, Manh BH, Elyazar IRF, Baird JK, Snow RW, Hay SI. The international limits and population at risk of *Plasmodium vivax* transmission in 2009. *Plos Negl Trop Dis.* 2010; 4(8):e774.

Guichoux E, Lagache L, Wagner S, Chumeil P, Léger P, Lepais O, Lepoittevin C, Malausa T, Revardel E, Salin F, Petit RJ, 2011. Current trends in microsatellite genotyping. *Molecular Ecology Resources.* 2012; 11, 591-611.

Gunawardena S, Karunaweera ND, Ferreira MU, Phone-Kyaw M, Pollack RJ, Alifrangis M, Rajakaruna RS, Konradsen F, Amerasinghe PH, Schousboe ML, Galappaththy GNL, Abeyasinghe RR, Hartl DL, Wirth DF. Geographic structure of *Plasmodium vivax*: microsatellite analysis of parasite populations from Sri Lanka, Myanmar and Ethiopia. *Am J Trop Med Hyg.* 2010; 82: 235-242.

Gupta B, Dash AP, Shrivastava N, Das A. Single nucleotide polymorphisms, putatively neutral DNA markers and population genetic parameters in Indian *Plasmodium vivax* isolates. *Parasitology.* 2010; 137:1721-1730.

Hall N. Genomic insights into the other malaria. *Nature Genetics*. 2012; v: 44- 9.

Haubold B, Hudson, RR. LIAN 3.0: detecting linkage disequilibrium in multilocus data. *Linkage Analysis. Bioinformatics*. 2000; 847-848.

Havryliuk T, Ferreira MU. A closer look at multiple-clone *Plasmodium vivax* infections: detection methods, prevalence and consequences. *Mem Inst Oswaldo Cruz*. 2009; 104: 67-73.

Havryliuk T, Orjuela-Sánchez P, Ferreira MU. *Plasmodium vivax*: microsatellite analysis of multiple-clone infections. *Exp Parasitol*. 2008; 120:330-6.

Hay SI, Guerra CA, Tatem AJ, Noor AM, Snow RW. The global distribution and population at risk of malaria: past, present, and future. *Lancet Infect Dis*. 2004; 4:327-336.

Hedrick PW. Perspective: Highly variable loci and their interpretation in evolution and conservation. *Evolution*. 1999; 53:313-318.

Hwang S-Y, Kim S-H, Kho W-G. Genetic Characteristics of Polymorphic Antigenic Markers among Korean Isolates of *Plasmodium vivax*. *Korean J Parasitol*. 2009; Vol. 47, S: S51-S58.

Imwong M, Nair S, Pukrittayakamee S, Sudimack D, Williams JT, Mayxay M, Newton PN, Kim JR, Nandy A, Osorio L, Carlton JM, White NJ, Day NPJ, Anderson TJ. Contrasting genetic structure in *Plasmodium vivax* populations from Asia and South America. *Int J Parasitol*. 2007; 37:1013–1022.

Imwong M, Sudimack D, Pukrittayakamee S, Osório L, Carlton JM, Day NPJ, White NJ, Anderson TJ. Microsatellite variation, repeat array length and population history of *Plasmodium vivax*. *Mol Biol Evol*. 2006; 23: 1016-8.

Imwong M, Boel ME, Pagornrat W, Pimanpanarak M, Mc Gready R, Day MPJ, F Nosteen F, White NJ. The first *Plasmodium vivax* relapses of life are usually genetically homologous. *JID*. 2012; 205.

Jeffares DC, Pain A, Berry A, Cox AV, Stalker J, Ingle CE, Thomas A, Quail MA, Siebenthall K, Uhlemann AC, Kyes S, Krihsna S, Newbold C, Dermitzakis ET, Berriman M. Genome variation and evolution of the malaria parasite *Plasmodium falciparum*. *Nature Genetics*. 2007; V 39:1.

Joy DA, Gonzalez-Ceron L, Carlton JM, Gueye A, Fay M, McCutchan TF, Su XZ. Local adaptation and vector-mediated population structure in *Plasmodium vivax* malaria. *Mol Biol Evol.* 2008; 25: 1245-1252.

Karunaweera ND, Ferreira MU, Hartl DL, Wirth DF. Fourteen polymorphic microsatellite DNA markers for the human malaria parasite *Plasmodium vivax*. *Molecular Ecology Notes.* 2007; 7:172–175.

Karunaweera ND, Ferreira MU, Munasinghe A, Barnwell JW, Collins WE, King CL, Kawamoto F, Hartl DL, Wirth DF. Extensive microsatellite diversity in the human malaria parasite *Plasmodium vivax*. *Gene.* 2008; 410:105-112.

Kimura M, Kaneko O, Qing L, Mian Z, Kawamoto F, Wataya Y, Otani S, Yamaguchi Y, Tanabe K. Identification of the four species of human malaria parasites by nested PCR that targets variant sequences in the small subunit rRNA gene. *Parasitol. Int.* 1997; v. 46, p. 91-95.

Ladislau JL, Leal MC, Tauil PL. Avaliação do plano de intensificação das ações de controle da malária na região da Amazônia Legal, Brasil, no contexto da descentralização. *Epidemiol Serv Saúde.* 2006; 15:9-20.

Leclerc MC, Durant P, Gauthier C, Patot S, Billote N, Menegon M, Severini C, Ayala FJ, Renaud F. Meager genetic variability of the human malaria agent *Plasmodium vivax*. *Proc Natl Acad Sci USA.* 2003; 101:14455-14460.

Livak KJ, Marmaro, Todd JA. Towards fully automated genome-wide polymorphism screening [Letter]. *Nature Genet.* 1995; 9: 341-342.

Loiola CC, da Silva CJ, Tauil PL. Malaria control in Brazil: 1965 to 2001. *Rev Panam Salud Publica* 2002, 11:235-244.

Ministério da Saúde. DATASUS. Secretaria de Vigilância Sanitária. Sistema de Informação de Vigilância epidemiologia- SIVEP – Malária, 2007 disponível em: http://portaweb04.saude.gov.br/sivep_malaria/default.asp.2007. Acesso em: 23 de Outubro de 2012.

Moll K, Ljungstrom I, Perlmann H, Scherf A, Wahlgren M, editors. *Methods in Malaria research*. Manassas: Malaria research and Reference Reagent Resource Center (MR4);2008.

Mu J, Duan J, Makova KD, Joy DA, Huynh CQ, Branch OH, Li WH, Su X. Chromosome-wide SNPs reveal an ancient origin for *Plasmodium falciparum*. *Nature*. 2002; 418:323-326.

Muhamad P, Chaijaroenkul W, Congpuong K, Na-Bangchang K. SYBER Green I and Taqman Quantitative Real-Time polymerase chain reaction methods for the determination of amplification of *Plasmodium falciparum* multigene resistance-1 gene (PFMDR-1). *J Parasitol*. 2011; 97(5):939-942.

Neafsey DE, Schaffner SF, Volkman SK, Park D, Montgomery P, Milner DA Jr, Lukens A, Rosen D, Daniels R, Houde N, Cortese JF, Tyndall E, Gates C, Stange-Thomann N, Sarr O, Ndiaye D, Ndir O, Mboup S, Ferreira MU, Moraes S do L, Dash AP, Chitnis CE, Wiegand RC, Hartl DL, Birren BW, Lander ES, Sabeti PC, Wirth DF. Genome-wide SNP genotyping highlights the role of natural selection in *Plasmodium falciparum* population divergence. *Genome Biol*. 2008; 9:R171.

Neafsey DE, Galinsky K, Jiang RHY, Young L, Sykes SM, Saif S, Gujja S, Goldberg JM, Young S, Zeng Q, Chapman SB, Dash AP, Anvikar AR, Sutton PL, Birren BW, Escalante AA, Barnwell JW, Carlton JM. The malaria parasite *Plasmodium vivax* exhibits greater genetic diversity than *Plasmodium falciparum*. *Nature genetics*. 2012; doi:10.1038/ng.2373.

Oliveira-Ferreira J, Lacerda MVG, Brasil P, Ladislau JLB, Tauil PL, Daniel-Ribeiro CT. Malaria in Brazil: an overview. *Malaria J*. 2010; 9:115.

Orjuela-Sánchez P, da Silva NS, da Silva-Nunes M, Ferreira MU. Parasitemia recurrences and population dynamics of *Plasmodium vivax* polymorphisms in rural Amazonia. *Am J Trop Med Hyg*. 2009; 81:961-968.

Orjuela-Sánchez P, Karunaweera NK, da Silva-Nunes M, da Silva NS, Scopel KKG, Gonçalves RM, Amaratunga C, Sá JM, Socheat D, Fairhurst RM, Gunawardena S, Thavakodirasah T, Galappaththy GNL, Abeyasinghe RR, Kawamoto F, Wirth DF, Ferreira MU. Single-nucleotide polymorphism, linkage disequilibrium and geographic structure in the malaria parasite *Plasmodium vivax*: prospects for genome-wide association studies. *BMC Genetics*. 2010; 11:65.

Pan American Health Organization (2008). *Malaria: Progress Report 2007*. Washington, DC: Pan American Health Organization.

Price RN, Douglas NM, Anstey NM. New developments in *Plasmodium vivax* malaria: severe disease and the rise of chloroquine resistance. *Curr Opin Infect Dis*. 2009;22:430-43.

Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics*. 2000; 155:945-959.

Putaporntip C, Udomsangpetch R, Pattanawong U, Cui L, Jongwutiwes S. Genetic diversity of the *Plasmodium vivax* merozoite surface protein-5 locus from diverse geographic origins. *Gene*. 2010; 456:24–35.

Rocha HF. Análise e mapeamento dos tipos de assentamentos no Brasil. Presidente Prudente, 2008. [relatório (Iniciação Científica em Geografia)] São Paulo: Faculdade de Ciências e Tecnologia da Universidade Estadual Paulista; 2008.

Sá JM, Nomura T, Neves J, Baird JK, Wellems TE, de Portillo HA. *Plasmodium vivax*: allelevariants of the *mdr-1* gene do not associate with chloroquine resistance to amodiaquine among isolates from Brazil, Papua and monkey adapted strains. *Exp Parasitol*. 2005; 109:256-259.

Sattabongkot J, Tsuboi T, Zollner GE, Sirichaisinthop J, Cui L. *Plasmodium vivax* transmission: chances for control? *Trends Parasitol*. 2004; 20: 192-198.

Schlötterer C. The evolution of molecular markers - just a matter of fashion? *Nat Rev Genet*. 2004; 5:63-69.

Snounou G, Viriyakosol S, Zhu XP, Jarra W, Pinheiro L, do Rosario VE, Thaithong S, Brown KN. High sensitivity method of detection of human malaria parasites by the use of nested polymerase chain reaction. *Mol. Biochem. Parasitol*. 1993; v. 61, p. 315-320.

Su X, Jiang H, Yi M, Mu J, Stephens RM. Large-scale Genotyping and Genetic Mapping in *Plasmodium* Parasites. *Korean J Parasitol*. 2009; Vol. 47, No. 2: 83-91.

Thompson JD, Higgins DG, Gibson TJ. ClustawW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic acid research*. 1994; 22:4673-4680.

Toonem R, Hughes S. Increased throughput for fragment analysis on ABI Prism 377 automated sequencer using a membrane comb and STrand software. *Biotechniques*. 2011; v. 31, p. 1320-1324.

Van den Eede P, Van der Auwera G, Delgado C, Huyse T, Soto-Calle V, Gamboa D, Grande T, Rodriguez H, Llanos A, Anné J, Erhart A, D'Alessandro U. Multilocus genotyping reveals high heterogeneity and strong local population structure of the *Plasmodium vivax* population in the Peruvian Amazon. *Malaria J*. 2010; 9:151.

Volkman SK, Sabeti PC, DeCaprio D, Neafsey DE, Schaffner SF, Milner DA, Jr1, Daily JP, Sarr O, Ndiaye D, Ndir O, Mboup S, Duraisingh MT, Lukens A, Derr A, Stange-Thomann N, Waggoner S, Onofrio R, Ziaugra L, Mauceli E, Gnerre S, Jaffe DB, Zainoun J, Wiegand RC, Birren BW, Hartl DL, Galagan JE, Lander ES, Wirth DF. A genome-wide map of diversity in *Plasmodium falciparum*. *Nat Genet*. 2007; vol: 39: 113-119.

Westenberger, SI, Dharia NV, Bacon DJ, Vinetz JM, Winzeler EA. Whole genomic variability analysis of *P. vivax* using high-density microarrays reveals population genetic structures. Abstract Book, *Vivax Malaria Research III: 2009 and Beyond*. 2010: p. 118.

Wilson PE, Kazadi W, Kamwendo DD, Mwapasa V, Purfield A, Meshnick SR. Prevalence of *pfprt* mutations in Congolese and Malawian *Plasmodium falciparum* isolates as determined by a new TaqMan assay. *Acta Tropica* 93. 2005; 97-106.

Woodberry T, Minigo G, Piera KA, Hanley JC, de Silva HD, Salwati E, Kenangalem E, Tjitra E, Coppel RL, Price RN, Anstey NM, Plebansk M. Antibodies to *Plasmodium falciparum* and *Plasmodium vivax* Merozoite Surface Protein 5 in Indonesia: Species-Specific and Cross-Reactive Responses. *J Inf Diseases*. 2008; 198:134–42.

World Health Organization (WHO). Ten facts on malaria. Geneva, 2011. Available from: <<http://www.who.int/features/factfiles/malaria/en/index.html>>. Acesso em: 20 Set 2012.

World Health Organization (WHO). World malaria report 2010. Geneva, 2010. Available from: <http://www.who.int/malaria/world_malaria_report_2010/en/index.html> Acesso em: 21 Set 2012.