



Ecole Nationale Vétérinaire de Toulouse

Rapport de stage
soutenu publiquement

le 10/06/2020

Par madame **BOULARIAS Ghania**

Titre de rapport de stage

**Développement d'un nouvel outil de séquençage d'amplicons pour
le génotypage de *Plasmodium vivax*.**

Pour obtenir le diplôme de **Master 2** de université Paul Sabatier de Toulouse

Spécialité : Gestion intégrée des maladies animales tropicales

Membres de jury

Dr. Mariette Ducatez	Examineur
Dr. Isabelle Morlais	Rapporteur
Dr. Michael White	Encadrant
M. Jason Rosado Sandoval	Co- encadrant

Laboratoire d'accueil

Institut Pasteur de Paris, France : département parasites et insectes vecteurs, équipe Malaria-
parasites et hôtes.

2019 / 2020

Avant-propos

Au départ de mon stage, le projet proposé était d'évaluer les performances d'AmpSeq (un nouvel outil de séquençage d'amplicons développé par l'unité malaria parasites et hôtes pour le géotypage de *Plasmodium vivax*).

Le déroulement du stage était prévu comme suit :

- 1) Optimiser 8 nouveaux marqueurs moléculaires situés sur 8 chromosomes (01,02, 04, 09, 10, 11,12,14) de *P. vivax* puis les rajouter au protocole AmpSeq déjà développé par un étudiant en thèse de notre équipe **M. Jason Rosado Sandoval** sur 6 chromosomes (03, 05, 06, 07, 08, 13) de *P. vivax*.
- 2) Effectuer le séquençage de nouvelle génération (NGS) sur des échantillons de terrain (Pérou) issus de cohortes longitudinales.
- 3) Analyser la diversité génétique des infections à *P. vivax* suite aux résultats de séquençage.

L'objectif final et de pouvoir développer un outil de séquençage d'amplicons (AmpSeq) ciblant 14 marqueurs moléculaires polymorphes situés sur les 14 chromosomes de *P. vivax* (un marqueur par chromosome) et de le tester sur un grand nombre d'échantillons de terrain provenant des pays endémiques au Paludisme.

En raison de la pandémie du Covid19 et le confinement survenu depuis le 17 Mars 2020 et la non autorisation de la reprise d'activité au laboratoire pour les stagiaires à Institut Pasteur avant le 31 Aout 2020. Notre travail s'est arrêté en première étape d'optimisation des marqueurs moléculaires et nous ne pouvions pas continuer la suite des manipulations au laboratoire à savoir la validation de ces nouveaux marqueurs optimisés sur des échantillons du terrains provenant du Pérou par l'AmpSeq et l'analyse des résultats.

De ce fait, nous allons présenter dans ce manuscrit l'état d'avancement de mes travaux de stage jusqu'au début de confinement et comme exemple d'AmpSeq certains résultats obtenus par le doctorant Jason Rosado Sandoval.

Remerciements

A mes encadrants :

Dr. Michael WHITE.

PhD student Jason ROSADO SANDOVAL.

Pour avoir accepté de diriger ce travail.

De m'avoir aidé à valoriser mon travail.

Pour tous ce qui j'ai appris pendant ce stage grâce à vous.

Veillez trouver ici l'expression de ma reconnaissance et de ma profonde estime.

Je remercie particulièrement

Dr. Ivo MUELLER.

Pour m'avoir accepté et accueilli chaleureusement au sein de unité MPH (Malaria Parasites et hôtes) afin de mener à bien ce travail . Mon passage au sein de votre laboratoire restera pour moi une formidable expérience, tant sur le plan professionnel que personnel.

Aux membres du jury

Dr. Mariette DUCATEZ et Dr. Isabelle MORLAIS,

qui ont bien voulu examiné notre travail. Soyez assurées que vous me faites un grand honneur

en acceptant de prendre part au présent jury.

Je remercie beaucoup **Dr. Timothée VERGNE** et **Dr. Stéphane BERTAGNOLI**, les responsables du Master pour leurs disponibilités et leurs aides ainsi que tous les étudiants de Master **GIMAT** Pour les moments partagés ensemble à Toulouse et à Montpellier.

Je remercie toute l'équipe **MPH** pour la disponibilité qu'ils m'ont manifesté, à tout instant, durant la durée de mon stage . je remercie vivement **Anthony RUBERTO, Mirco Miroslav SANDFORT, Narimane NEKKAB, Soazic GARDAIS, Thomas OBADIA et Yasmine ELGHARABAWY.**

Je tiens à remercier aussi **Maryse BRANDT** et **Nirina ANDRIAKOTO** pour votre aide.

A toutes les personnes avec qui j'ai pu interagir à l'école vétérinaire de Toulouse, au CIRAD de Montpellier et à l'institut Pasteur de Paris.

Un grand merci.

Table des matières

Résumé/Abstract

I- Introduction.....	1
1- Actualité sur le fardeau mondial du paludisme	2
2- La biologie unique du <i>Plasmodium vivax</i>	4
3- Les mécanismes biologiques de l'augmentation de l'aire de répartition de <i>Plasmodium vivax</i>	5
4- Caractérisation moléculaire de <i>Plasmodium vivax</i>	7
II- Matériel et méthodes.....	9
1- Développement d'un nouvel outil de séquençage d'amplicons (AmpSeq).....	9
1-1- Sélection et optimisation des marqueurs.....	9
1-2- Exemple d'AmpSeq développé par notre unité.....	11
1-3- Séquençage de nouvelle génération (NGS).....	13
1-4- Pipeline d'analyse bio-informatique «HaplotypR».....	14
2- Analyses statistiques	14
III- Résultats	15
1- Les marqueurs moléculaires optimisés dans cette étude	15
2- Présentation des résultats d'un exemple d'AmpSeq de six marqueurs déjà optimisés.....	16
2-1- Description des marqueurs moléculaires	16
2-2- Diversité des haplotypes par marqueur et par échantillon.....	17
2-3- Diversité intra-hôte.....	19
2-4- Multiplicité d'infection (MOI)	20
2-4-1- Comparaison de la MOI par AmpSeq et MS (Microsatellites).....	21
2-4-2- Association entre la MOI et de l'âge et le sexe.....	22
2-4-3- Corrélation entre la parasitémie et le nombre de lecture	22
IV- Discussion.....	23
V- Conclusion et perspectives	26
Références bibliographiques	27
Annexes	30

Résumé :

La contrainte majeure de l'élimination de *Plasmodium vivax* est liée principalement aux rechutes d'hypnozoïtes du foie. Pour mieux comprendre la biologie de cette espèce, il est nécessaire de distinguer entre les rechutes et les réinfection ainsi que d'identifier les infections multiclonales, en particulier dans les zones endémiques. Les outils génétiques actuels tels que les marqueurs polymorphes de longueur (les microsatellites) n'offrent pas toujours une résolution suffisante pour répondre à cette question.

De ce fait, nous avons optimisé 4 marqueurs moléculaires polymorphes situés sur 4 chromosomes 02,04,10, 11 du *P. vivax* dans le but de développer un nouvel outil de séquençage d'amplicon courts basés sur les SNPs (AmpSeq) en utilisant le séquençage de nouvelle de génération (NGS) puis le valider sur des échantillons provenant du Pérou.

Un exemple d'AmpSeq mené sur six autres marqueurs déjà optimisés par notre équipe puis validés sur des échantillons provenant du Pérou et de Papouasie Nouvelle Guinée (PNG) a montré une diversité génétique élevée pour 3 marqueurs : *msp1_Chro07*, *STP1_Chro05* et *PIR_Chro06*. De plus, des clones majoritaires et minoritaires ont été détectés à la fois chez un hôte.

Ces résultats antérieurs ont démontré que AmpSeq est un outil puissant et très sensible pour estimer le MOI à *P. vivax* à des taux de parasitémie faible. Pour cela, des études supplémentaires sont nécessaires afin d'explorer les 8 marqueurs optimisés ce qui permettrait d'identifier plus haplotypes avec une diversité génétique plus élevée.

Mots clés : *Plasmodium vivax*, PCR, Séquençage d'amplicons (AmpSeq), séquençage de nouvelle génération (NGS), SNP, Haplotypes, Multiplicité d'infection (MOI), Pérou, Papouasie Nouvelle Guinée (PNG).

Abstract:

The major constraint to eliminate *Plasmodium vivax* infection is mainly link to relapses of hypnozoïtes from the liver. To better understand the biology of this species, it is necessary to distinguish between relapses and reinfection as well as to identify multiclonal infections with minority clones, especially in endemic areas. Current genetic tools such as polymorphic length markers (microsatellites) do not always offer sufficient resolution to answer this question.

Therefore, we have optimized 4 polymorphic molecular markers (with several SNPs) located on chromosomes 02, 04, 10, 11 of *P. vivax* in order to develop a new amplicon deep sequencing tool based on SNPs (AmpSeq) using next generation sequencing (NGS) and then validating it on field samples from Peru.

An example of AmpSeq carried out with six other markers already optimized by our team then validated on samples from Peru and Papua New Guinea (PNG) showed a high genetic diversity for a 3 markers: *msp1_Chro07*, *STP1_Chro05* and *PIR_Chro06*. AmpSeq detected both majority and minority clones within host. These previous results demonstrated that AmpSeq is a powerful and very sensitive tool for estimating the MOI of *P. vivax* at low parasitaemia density. For this, additional studies are necessary in order to explore the eight optimized markers which would allow us to identify more haplotypes with higher genetic diversity.

Keywords: *Plasmodium vivax*, PCR, Amplicon deep sequencing (AmpSeq), Next generation sequencing (NGS), SNP, Haplotypes, Multiplicity of infection (MOI), Peru, Papua New Guinea (PNG).

Liste des tableaux

Tableau 01 :	Les données recueillies sur les échantillons testés.....	12
Tableau 02 :	La fréquence des haplotypes (Diversité intra-hôte).....	19
Tableau 03 :	MOI par individu de la population du Pérou.....	21
Tableau 04 :	MOI par individu de la population de PNG.....	21
Tableau 05 :	Comparaison entre MOI entre MS et AmpSeq.....	22

Liste des figures

Figure 01 :	Carte mondiale du paludisme selon le statut épidémiologique des pays	3
Figure 02 :	Cycle de vie du parasite du paludisme humain à <i>Plasmodium vivax</i>	5
Figure 03 :	Représentation schématique de la PCR primaire et secondaire.....	10
Figure 04 :	Les marqueurs candidats sur chaque chromosome de <i>P. vivax</i>	11
Figure 05 :	Carte géographique des localités d'échantillons analysés dans cette étude	13
Figure 06 :	Représentation schématique des amorces Library (Library PCR).....	13
Figure 07 :	Image de migration en électrophorèse (PCR primaire et secondaire).....	15
Figure 08 :	Image de migration en électrophorèse (PCR multiplexes).....	16
Figure 09 :	Nombre d'haplotypes par marqueur dans tous les échantillons	18
Figure 10 :	Nombre d'haplotypes par marqueur dans la population du Pérou et PNG.	18
Figure 11 :	Nombre d'haplotypes par marqueur par échantillon du Pérou.....	18
Figure 12 :	Nombre d'haplotypes par marqueur par échantillon de PNG.....	19
Figure 13 :	Fréquence des Haplotypes au seins de deux cas du Pérou et de PNG....	20
Figure 14 :	Corrélation entre MOI et âge.....	22
Figure 15:	Association entre MOI et le sexe.....	22
Figure 16:	Corrélation entre la parasitémie et le nombre de lecture.....	23

Liste des abréviations

ADN :	Acide désoxyribonucléique.
AMA1 :	Apical membrane antigen 1.
AmpSeq :	Amplicon sequencing.
Chr :	Chromosome.
CPMP:	Conserved <i>Plasmodium</i> membrane protein.
CPP :	Conserved <i>Plasmodium</i> protein.
CSP :	Circumsporozoite protein.
F :	Forward.
GILP :	Glyoxalase I-like protein GILP, putative.
He	Heterozygosity.
Inn :	Inner.
MOI :	Multiplicity of infection.
MS :	Microsatellites sequence.
MSP :	mérozoïtes surface protein .
NGS :	Next generation sequencing.
Out :	Outer.
Pb :	Paire de bases.
PCR :	Polymerase chain reaction.
PEP :	<i>Plasmodium</i> exported protein.
PIR protein :	<i>Plasmodium</i> interspersed repeat protein.
PNG :	Papouasie nouvelle guinée.
PTP2 :	Protein tyrosine phosphatase, putative.
R :	Reverse.
SNP :	Single nucleotide polyporphism.
STP1 protein :	Sugar transport protein 1.
WHO :	World Health Organization.

I- Introduction

Le paludisme ou Malaria est une maladie parasitaire vectorielle. Sa transmission fait appel à des moustiques du genre *Anophele*. Elle est endémique dans les régions tropicales et subtropicales du monde. Le paludisme est causé par des protozoaires du genre *Plasmodium* (*Apicomplexa: Plasmodiidae*), un groupe diversifié qui infecte une variété d'hôtes vertébrés, notamment des primates. Cette diversité a conduit à une maladie impliquant plusieurs parasites et espèces vectrices dans divers écosystèmes du monde (Escalante and Pacheco, 2019).

Environ 228 millions de cas de paludisme ont été estimés dans le monde en 2018 avec un nombre de décès qui s'élève à 405 000, dont 272 000 (67 %) des décès sont des enfants de moins de 5 ans (World Health Organization, 2019). L'âge avancé, la grossesse, la splénectomie sont également des facteurs de risque d'accès grave de la maladie. De plus, les sujets naïfs du paludisme en séjour ou récemment installés en pays d'endémie sont aussi à risque de forme sévère de la maladie (Diallo, 2018).

Il existe plus de 140 espèces de *Plasmodium* touchant diverses espèces animales dont 5 de ces espèces sont retrouvées en pathologie humaine. Il s'agit de *Plasmodium falciparum*, *Plasmodium vivax*, *Plasmodium ovalae*, *Plasmodium malariae* et *Plasmodium knowlesi* (Diallo, 2018). *P. falciparum* et *P. vivax* sont responsables de la majorité des cas de malaria dans le monde. *Plasmodium falciparum*, *P. knowlesi* et *P. malariae* envahissent les érythrocytes matures bien que *P. falciparum* et *P. knowlesi* puissent également pénétrer préférentiellement dans les réticulocytes. En revanche, *P. vivax* et *P. ovalae* envahissent généralement les réticulocytes (Chan et al., 2020). La base biologique de cette exigence apparente, en dehors des ligands parasites spécifiques qui pourraient permettre l'entrée de *P. vivax* dans les réticulocytes n'est pas comprise. Cela pourrait également être une adaptation parasitaire pour limiter les hyperparasitémies et la virulence associée (puisque les réticulocytes ne représentent que 1 à 2% des érythrocytes) ou bien, le réticulocyte pourrait offrir un microenvironnement spécial pour soutenir la croissance de cette espèce (Mueller et al., 2009).

1- Actualité sur le fardeau mondial du paludisme

En 2018, environ 228 millions de cas de paludisme sont survenus dans le monde contre 251 millions de cas en 2010. La plupart des cas de paludisme en 2018 se sont produits dans la

région africaine (213 millions, 93%), suivi par la région de l'Asie du Sud-Est avec 3,4% des cas et le Région de la Méditerranée orientale avec 2,1% (World Health Organization, 2019).

19 pays d'Afrique subsaharienne et l'Inde possèdent près de 85% du paludisme mondial. 6 pays représentaient plus de la moitié de tous les cas de paludisme dans le monde: le Nigéria (25%), République démocratique du Congo (12%), Ouganda (5%) et Côte d'Ivoire, Mozambique et Niger (4% chacun). Le taux d'incidence du paludisme a diminué dans le monde entre 2010 et 2018, passant de 71 à 57 cas par 1000 populations à risque. Cependant, de 2014 à 2018, le taux d'incidence a considérablement ralenti, réduit à 57 en 2014 et restant à des niveaux similaires jusqu'en 2018. La Région de l'Asie du Sud-Est a continué de voir son taux d'incidence chuter de 17 cas pour 1000 personnes à risque en 2010 à 5 cas en 2018. Dans la Région africaine, les niveaux d'incidence des cas sont également passés de 294 en 2010 à 229 en 2018, une réduction de 22% (World Health Organization, 2019). Toutes les autres régions ont enregistré soit peu de progrès, soit une augmentation de taux de l'incidence comme il a été signalé dans certaines régions d'Amériques (World Health Organization, 2019).

En 2018, il y avait environ 405 000 décès dus au paludisme dans le monde, contre 416 000 décès estimés en 2017 et 585 000 en 2010. La Région africaine a été à l'origine de 94% de tous les décès dus au paludisme en 2018. Bien que cette région ait enregistré le plus grand nombre de décès dus au paludisme en 2018, elle a également représenté 85% des 180 000 décès de paludisme dans le monde signalés en 2018 par rapport à 2010. Près de 85% des décès dus au paludisme dans le monde en 2018 étaient concentrés dans 20 pays de la Région africaine et en Inde; Le Nigéria est responsable de près de 24% de tous les décès dus au paludisme dans le monde, suivi de la République démocratique du Congo (11%), de la République-Unie de Tanzanie (5%) et de l'Angola, du Mozambique et du Niger (4% chacun) (World Health Organization, 2019). En 2018, seules la Région africaine et la Région de l'Asie du Sud-Est ont enregistré des réductions des décès dus au paludisme par rapport à 2010. La Région africaine a enregistré la plus forte réduction absolue des décès, passant de 533 000 en 2010 à 380 000 en 2018. Malgré ces gains, le taux de réduction de la mortalité due au paludisme a également ralenti depuis 2016 (World Health Organization, 2019).

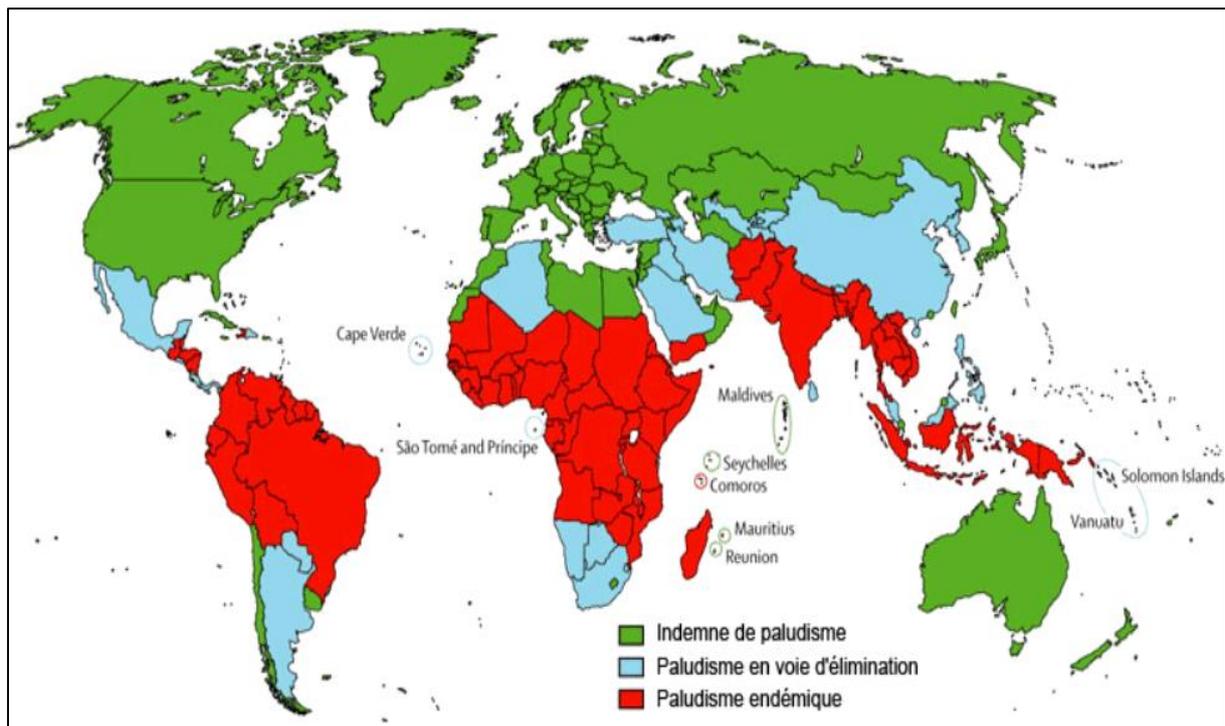


Figure 01 : Carte mondiale du paludisme selon le statut épidémiologiques des pays (WHO, 2010).

L'accent récemment mis sur l'élimination du paludisme et l'accessibilité aux outils de diagnostic plus sensibles ont révélé une plus grandes complexités dans l'épidémiologie du paludisme à travers les pays. *Plasmodium falciparum* est l'espèce la plus répandue dans la région africaine et représente 99,7% des cas estimés de paludisme en 2018 dans ce continent (World Health Organization, 2019). En dehors de l'Afrique, *Plasmodium vivax* est la première cause du paludisme la plus répandue dans le monde. Depuis 2010, le nombre d'études identifiant *P. vivax* endémique a considérablement augmenté, avec 88 nouveaux rapports scientifiques publiés, doublant approximativement les données disponibles et compliquant ainsi davantage la compréhension pratique actuelle de l'épidémiologie du paludisme dans différentes régions (Twohig et al., 2019). En Afrique, la domination écrasante du groupe sanguin Duffy négatif qui signifie que les globules rouges n'ont pas le récepteur Duffy, une molécule importante pour l'invasion du parasite *P. vivax* empêche sa transmission endémique prolongée. Néanmoins, des études récentes ont démontré qu'une infection à *P. vivax* peut toujours survenir dans les globules rouges Duffy-négatifs et donc la Duffy-négativité seule ne peut pas empêcher la transmission endémique (Twohig et al., 2019).

À ce jour, il y a eu moins de recherches sur *P. vivax* par rapport à *P. falciparum*, principalement en raison de l'absence de système de culture in vitro à long terme pour cette

espèce et également parce qu'elle est considérée comme plutôt bénigne par rapport à *P. falciparum*. Néanmoins, *P. vivax* est responsable de symptômes cliniques graves et invalidants ayant des effets importants sur la santé humaine (Adams and Mueller, 2017; Rougeron et al., 2020). Il a été estimé que les infections à *P. vivax* représentent jusqu'à 80 millions de cas de paludisme clinique chaque année avec plus de 70% des infections en Asie et en Amériques (Mendis et al., 2001; Price et al., 2009; World Health Organisation, 2015).

2- La biologie unique du *Plasmodium vivax*

Des stratégies efficaces de lutte contre le paludisme à *Plasmodium vivax* nécessitent une compréhension précise de sa biologie. En effet, *Plasmodium vivax* est doté d'un cycle de vie unique et très complexe, durant lequel cette espèce subit plus de dix étapes de différenciation cellulaire et envahit au moins quatre types de cellules au sein de deux hôtes différents (Figure 02). En effet, le moustique du genre *Anophèles* femelles inocule le parasite sous sa forme sporozoïtes infectieux suite à une piqure de la peau qui atteignent la circulation sanguine et pénètrent dans les hépatocytes amorçant le stade exo-érythrocytaire. Dans le foie, *P. vivax* peut soit se différencier en Schizontes tissulaires, qui après des milliers de réplifications mitotiques dans des hépatocytes individuels libèrent les mérozoïtes dans la circulation sanguine, ou soit se différencient en un stade dormant appelé hypnozoïte qui après des mois ou des années peuvent se réactiver provoquant ainsi une rechute clinique. Pendant les stades érythrocytaires, *P. vivax* sous sa forme merozoïte envahit exclusivement les réticulocytes et ces cellules deviennent grandes et déformées. Ce processus de développement cyclique prend environ 48 heures. De plus, *P. vivax* produit des protéines spécifiques pour créer des complexes (alvéole-vésicule) qui apparaissent comme des mouchetures abondantes sur les frottis sanguins colorés au Giemsa, connus sous le nom de points de Schüffner. La fonction de ces structures intrigantes et leur potentiel en tant que cibles d'intervention restent largement inexplorés.

Certains parasites *P. vivax* peuvent se différencier en gamétocytes matures avant que l'infection clinique se développe, ayant ainsi l'avantage de continuer à se transmettre à l'insecte vecteur avant l'apparition des symptômes cliniques et le traitement ultérieur. Ces gamétocytes circulants sont de forme arrondie et lors de leur absorption dans le repas sanguin des anophèles, les moustiques commencent le cycle sexuel, qui comprend la libération des gamètes mâles et femelles, la fécondation et la formation d'un ookinète mobile qui traverse l'épithélium de l'intestin moyen. puis la différenciation en une nouvelle forme répllicative connue sous le nom d'oocyste, puis la libération de sporozoïtes, la migration et l'invasion des glandes salivaires qui met fin à ce cycle de vie complexe (Mueller et al., 2009).

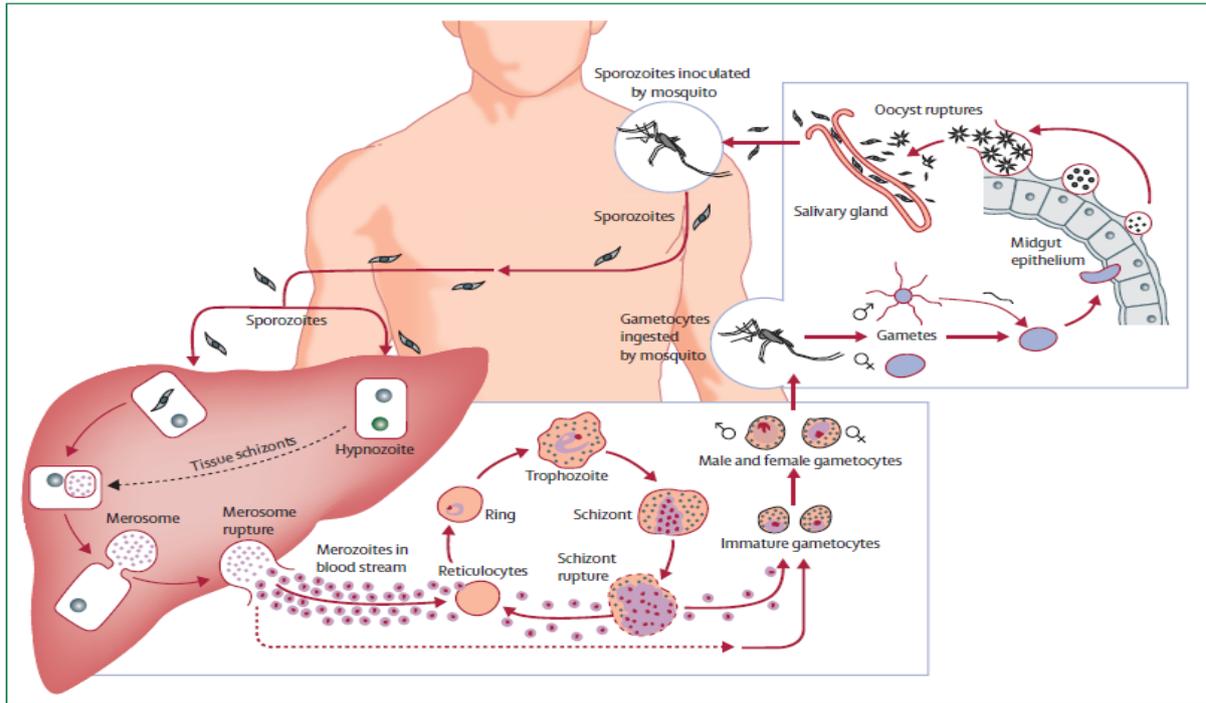


Figure 02: Cycle de vie du parasite du paludisme humain à *Plasmodium vivax* (Mueller et al., 2009).

3- Les mécanismes biologiques de l'augmentation de l'aire de répartition de *Plasmodium vivax*

Arriver à un monde exempt de paludisme est un objectif de santé publique et de développement durable ambitieux et considérable (World Health Organization, 2019). Au cours de cette dernière décennie, de nombreux pays ont enregistré une réduction du nombre de cas cliniques de paludisme (World Health Organisation, 2015). Par conséquent, 34 pays tentent activement d'éliminer le paludisme (Adams and Mueller, 2017). Parallèlement à cette réduction de l'incidence globale, un changement prononcé dans la composition des espèces a été observé avec *P. vivax* désormais le *Plasmodium* spp. prédominant dans la grande majorité des pays hors Afrique et le principal obstacle à l'élimination du paludisme en Asie et en Amériques (Adams and Mueller, 2017).

Les mécanismes biologiques qui expliquent l'augmentation de l'aire de répartition de *P. vivax* sont probablement liés d'une part au potentiel de transmission élevé provoquée par la production précoce et continue de gamétocytes, et un cycle de développement plus court chez le moustique vecteur par rapport à d'autres *Plasmodium* spp. (Mueller et al., 2009). De l'autre part, La résistance de plus en plus répandue à la chloroquine, le principal médicament de première ligne pour traiter le paludisme à *P. vivax* (Ruebush et al., 2003 ;Ketema et al., 2009 ;

Rijken et al., 2011) et à la Primaquine, le seul médicament actuellement autorisé capable d'éliminer les hypnozoïtes. Cela représente un défi supplémentaire pour les programmes existants de contrôle et de prévention du paludisme à *P. vivax* et met en évidence le besoin critique d'efforts supplémentaires pour développer de nouveaux outils ciblant spécifiquement *P. vivax* (Adams and Mueller, 2017).

Enfin, la difficulté majeure de contrôle et d'élimination du paludisme à *P. vivax* est liée principalement à sa particularité de développer un stade hépatique dormant « hypnozoïte » métaboliquement inactif mais qui se réactive provoquant une rechute (infection des globules rouges) après plusieurs semaines voire plusieurs mois ou années de l'infection primaire (Bright et al., 2014). Les rechutes représentent ainsi une parasitémie émergente qui survient malgré une thérapie schizonticide sanguine efficace de la parasitémie primaire et qui échappent totalement aux moyens conventionnels de lutte antipaludique, c'est-à-dire l'utilisation d'insecticides, de moustiquaires, de méthodes de diagnostic et de presque toutes les interventions chimio prophylactiques ou chimio thérapeutiques (Mueller et al., 2009).

Bien que les rechutes soient connues depuis la fin du XIX^e siècle et malgré l'importante de morbidité due à *P. vivax* dans le monde, peu de progrès ont été réalisés pour caractériser ce stade parasitaire (Mueller et al., 2009).

Des études antérieures ont montré que les rechutes survenant chez les patients avec un faible taux d'hypnozoïtes dans leur foie, sont génétiquement homologues aux parasites de l'infection initiale (Imwong et al., 2012). En revanche, dans les régions endémiques où les patients hébergent des hypnozoïtes provenant de multiples piqûres de moustiques infectieuses (infections polyclonales) avec des clones dominants et des clones minoritaires (Lerch et al., 2017). Dans ce cas-là, les rechutes sont causées par des parasites génétiquement hétérologues à l'infection initiale (Imwong et al., 2007 ; De Araujo et al., 2012).

Ajoutant que la dynamique complexe des infections à *P. vivax* empêche de classer les épisodes de paludisme récurrents sur le terrain en tant qu'infections par rechute provoquées par des hypnozoïtes, ou en tant qu'infections recrudescentes causées par un échec d'élimination de l'infection initiale, ou encore en tant que réinfections. L'incapacité de distinguer ces trois causes de l'infection du paludisme dans les zones d'endémie ainsi que l'identification des clones minoritaires des clones dominants empêchent des estimations précises de la prévalence des hypnozoïtes et inhibent la capacité d'étudier directement ce stade parasitaire (Bright et al., 2014). Cela explique aussi les échecs de traitements dans les pays endémiques et entrave le

développement de la prochaine génération de médicaments anti-hypnozoïtes (Snounou and Beck, 1998; Greenhouse et al., 2006 ; Bright et al., 2014).

4- Caractérisation moléculaire de *Plasmodium vivax*

De nombreuses investigations épidémiologiques moléculaires caractérisent les variants génétiques du parasite. En effet, le génotypage s'est initialement concentré sur la prévalence des mutations conférant une résistance aux antipaludiques, ou sur la diversité des gènes codant pour les antigènes considérés comme des candidats vaccinaux (Escalante et al., 2015). Ces approches se sont avérées utiles pour déterminer le nombre de génotypes de parasites différents co-infectant un seul patient appelée la multiplicité d'infection (MOI) qui est définie comme le nombre de souches parasitaires génétiquement distinctes co-infectant un seul hôte . L'infection multiclonale peut être le résultat de piqûres indépendantes de moustiques infectés (également appelés surinfection), ou d'une piqûre de moustique unique transmettant un inoculum de sporozoïtes génétiquement diversifié(Zhong et al., 2018b).

Les parasites du paludisme génétiquement distincts ont des taux extrêmement élevé de recombinaisons génétiques au cours des stades sexuels chez un moustique vecteur, entraînant souvent la transmission simultanée de plusieurs souches. Dans le cas de *Plasmodium vivax*, des infections multiclonaux peuvent également être causées par la rechute d'un hypnozoïte au stade hépatique (Zhong et al., 2018b). Certaines études ont démontrées que la MOI au sein d'un hôte pourrait être un indicateur important de l'intensité de la transmission dans les zones endémiques du paludisme. En effet, une corrélation positive entre le taux d'infections polyclonales et l'incidence annuelle des parasites a été observée dans les populations de parasites d'Indonésie et de Papouasie-Nouvelle Guinée (Zhong et al., 2018b) alors que dans d'autres études, aucune corrélation n'a été trouvée entre la proportion infections multiclonaux et la prévalence de infection (Zhong et al., 2018b).

Les méthodes de génotypage basées sur des marqueurs polymorphes de longueur étaient appliquées depuis des décennies, notamment en ciblant les marqueurs microsatellites (qui sont des séquences de di-, tri- ou tétra- nucléotides répétés en tandem, réparties en plusieurs exemplaires sur l'ensemble du génome d'une espèce et présentent un taux de polymorphisme élevé. Ce polymorphisme repose sur la variation du nombre d'unités de répétition constituant le microsatellite). Ou en ciblant des gènes codants pour des antigènes du parasites tels que les mérozoïtes surface Protein (*m*sp 1 et 3) et circumsporozoïte protein (*csp*) (Anderson et al., 1999).

Malgré la large utilisation des marqueurs microsatellites dans de nombreux laboratoires de recherche sur le paludisme, ces derniers ont des limites importantes principalement dans la détection des clones minoritaires lors d'infection multiclonaux (Lerch et al., 2017). Une autre limitation des marqueurs polymorphes de longueur, consiste dans les différences de taille généralement importantes entre les allèles. Des différences de taille importantes conduisent à un biais d'amplification, préférant les fragments de PCR plus courts dans les échantillons qui contiennent simultanément plusieurs *Plasmodium* (MOI) (Messerli et al., 2017a).

Et pour surmonter certaines de ces limites citées au-dessus, le séquençage d'amplicons polymorphes basés sur les polymorphismes mono nucléotidiques (SNP) a été remplacé et a pris place dans plusieurs études (Lerch et al., 2017). Les SNPs sont devenus l'objet d'un grand nombre d'études conçues pour identifier les différences critiques dans la séquence d'ADN, qui contribuent aux variations génotypiques et à identifications des différents haplotypes.

En effet, deux approches différentes ont été décrites pour le génotypage de *P. falciparum* et *P. vivax* par le séquençage d'amplicon basé sur les SNPs : La première consiste au séquençage des marqueurs de génotypage polymorphes de longueur classique en ciblant des régions non répétitives et qui abritent un vaste polymorphisme mono nucléotidique (SNP), comme pour le gène qui code pour l'antigène de surface *msp1* pour *P. vivax* ou la protéine CSP de *P. falciparum*. La force de ces approches est que tous les SNPs d'un amplicon sont liés par une seule lecture de séquence, ce qui conduit directement à l'identification des haplotypes (Bailey et al., 2012 ; Neafsey et al., 2015 ; Lerch et al., 2017). La deuxième consiste au séquençage de plusieurs loci de taille courte très polymorphiques de distribution à l'échelle du génome, chaque locus comprenant des SNPs (Friedrich et al., 2016).

De ce fait, l'objectif de notre travail serait de développer un nouvel outil de séquençage d'amplicons (AmpSeq) très sensible applicable dans les zones endémiques ciblant 14 régions courtes mais hautement polymorphes (beaucoup de SNPs) sur les 14 chromosomes du *Plasmodium vivax* (un seul marqueur pour chaque chromosome) qui permet la détection des clones minoritaires à faible parasitémie ainsi que distinguer les rechutes dues aux hypnozoïtes des réinfections.

II- Matériel et méthodes

1- Développement d'un nouvel outil de séquençage d'amplicon (AmpSeq)

1-1- Sélection et optimisation des marqueurs

Quatre marqueurs moléculaires avec une forte hétérozygotie (He) (riche en SNPs) ont été sélectionnés sur 4 chromosomes différents de *Plasmodium vivax*. Il s'agit du lysophospholipase_Chr_02 avec 11 SNPs, apical membrane antigen 1 Chr_09 (AMA1_Chr_09) avec 17 SNPs, merozoite surface protein 3_Chr_10 (MSP3_Chr_10) avec 15 SNPs et le marqueur glyoxalase I-like protein GILP_Chr_11, putative (GILP_Chr_11) avec 6 SNPs. La taille des amplicons était limitée à environ 200 (pb) pour chaque marqueur.

Le contrôle de la qualité des amorces a été évalué à l'aide d'outils en ligne pour les structures secondaires et l'interaction dimère en utilisant (<https://www.eurofinsgenomics.eu/en/ecom/tools/oligo-analysis/>). Pour chaque marqueur, nous avons procédé à une PCR nichée avec une PCR primaire et PCR secondaire. Les amorces de la PCR primaire ont été désignées pour chaque marqueur à l'aide de la base de données en ligne PlasmoDB (<https://plasmodb.org/>) et Primers Blast (<https://www.ncbi.nlm.nih.gov/tools/primer-blast/>). Toutes les PCR ont été effectuées par le kit PCR KAPA HiFi HotStart Ready Mix (KAPA Biosystems, USA). Nous avons optimisé les marqueurs en tenant compte de la concentration des amorces, de la température d'hybridation, et de la concentration du tampon. Pour cela, une première PCR de 35 cycles a été effectuée pour optimiser les amorces des PCR primaires. Pour chacun des marqueurs, cette PCR a été répétée à 25 cycles suivie d'une nested PCR spécifique au marqueur à 25 cycles en utilisant le produit de la PCR primaire avec des amorces qui portent la séquence de liaison à l'extrémité 5' (5'F- GTGACCTATGAACTCAGGAGTC), (5'R- CTGAGACTTGACATCGCAGC). (Figure 03). Ces séquences « linker » vont servir de liaison dans les prochaines étapes à savoir la « Library PCR » et le séquençage de nouvelle génération (NGS). Des échantillons positifs à *Plasmodium vivax* ont été utilisés dans chaque réaction.

Les détails des séquences des amorces des marqueurs moléculaires désignés puis optimisés ne sont pas présentés dans ce manuscrit car ils resteront confidentiels jusqu'à la publication de ces résultats dans un article scientifique.

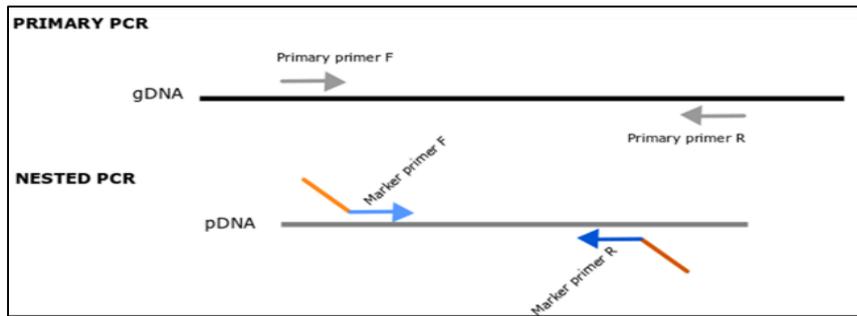


Figure 03 : Représentation schématique de la PCR primaire et secondaire. gDNA: ADN génomique du parasite. pDNA: produit de PCR primaire. Marker primer : amorces spécifiques avec une séquence « Linker » aux extrémités 5'.

Pour chaque marqueur optimisé, un volume de 5 µl de chaque amplicon de la PCR secondaire mélangé avec 2 µl du tampon de charge (6X) ont été résolus sur le gel d'agarose à 1,5% avec du SYBR® Safe (1X) en électrophorèse dans un tampon (TAE 1X) à 90 V pendant 45 min. Les bandes ont été visualisées sous un lecteur ultraviolet. Les 20 µl du volume restants ont été purifiés en utilisant le kit de purification Qiagen (QIAquick PCR Purification Kit, Germany) suivant les instructions du fabricant, puis envoyer au séquençage (séquençage de Sanger à Eurofin genomic). Les séquences obtenues ont été alignées et comparées aux séquences des marqueurs d'intérêt à l'aide de Blast nucléotide (<https://www.ncbi.nlm.nih.gov/tools/nucleotides-blast/>) et le logiciel MEGAX. Le but de cette étape est de vérifier l'identité des séquences amplifiées avec les séquences des marqueurs d'intérêt de *Plasmodium vivax* (Annexes 01). Par la suite, des PCR multiplexes ont été effectuées pour les amorces des PCR primaires. Cela permet de gagner du temps et de minimiser le volume des échantillons d'ADN à utiliser dans les réactions. Pour cela, nous avons séparé les amorces des PCR primaires, en fonction de la taille moléculaire des amplicons (pb) de façon à pouvoir les identifier sur gel d'agarose en fonction de leurs tailles. Enfin, des PCR nichées ont été effectuées de nouveau séparément pour chacun des marqueurs en utilisant le produit des PCR multiplexe comme matrice et les produits de PCR sont analysés sur gel d'agarose comme précédemment.

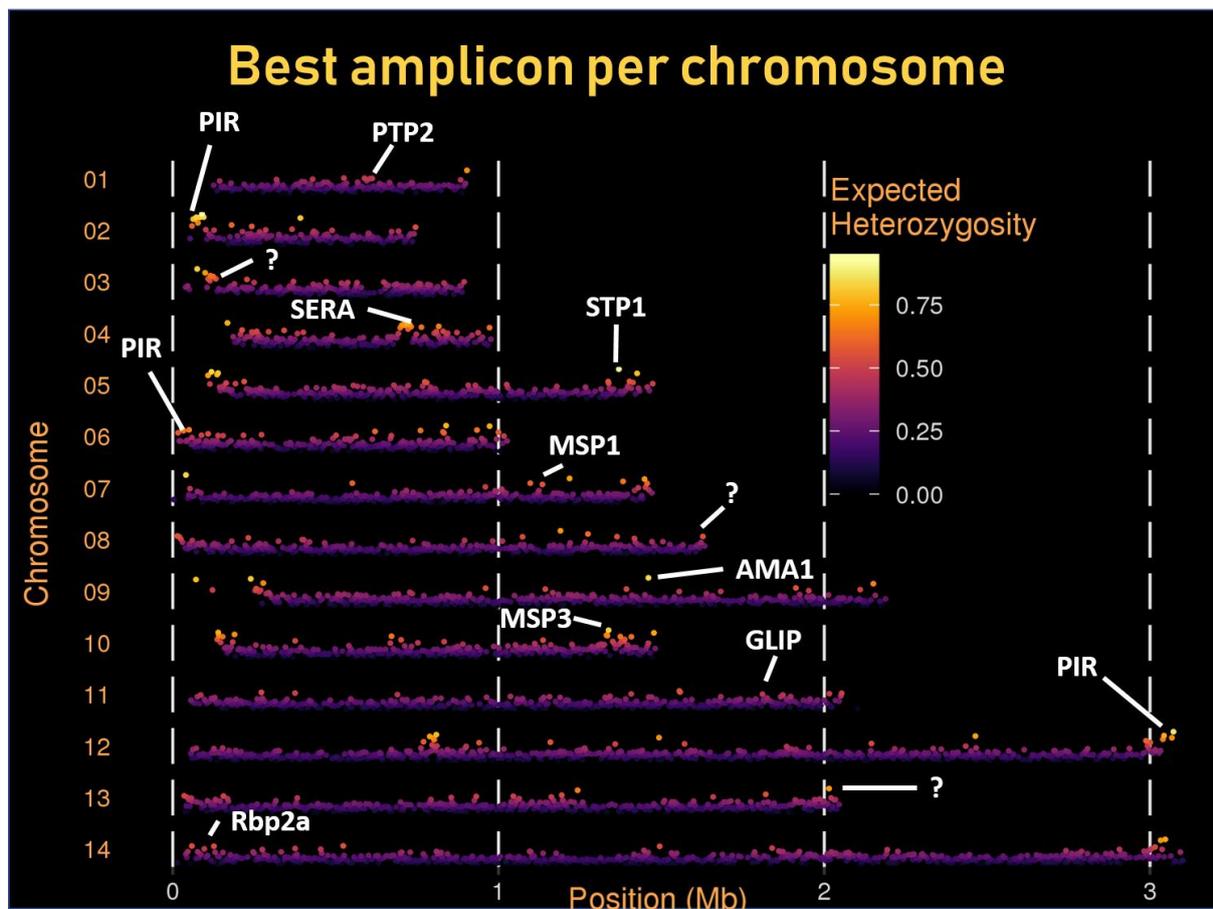


Figure 04 : Les marqueurs candidats sur chaque chromosome de *P. vivax*. Les marqueurs ayant une hétérozygotie (He) élevée ont été sélectionnés: suivant l'échelle à droite de la figure : les marqueurs ayant une He (0.75 et 1) sont indiqués sur les chromosomes en points jaunes. Ceux qui ont He (0.5 et 0.75) sont indiqués en points oranges.

1-2- Exemple d'AmpSeq développé par notre unité

En raison de la pandémie du Covid19 et le confinement survenu depuis le 17 Mars 2020. Nous ne pouvons pas continuer la suite des manipulations au laboratoire à savoir la validation de des nouveaux marqueurs optimisés sur des échantillons provenant du Pérou par le séquençage de nouvelle génération (NGS) et analyse des résultats. De ce fait, Nous allons présenter dans ce manuscrit comme exemple d'AmpSeq les résultats obtenus par un étudiant en thèse de notre équipe.

Son travail consiste à optimiser 6 autres marqueurs moléculaires : le marqueur *msp1* situé sur le chromosome 07 (*msp1*_Chr07) contenant 37 SNPs, marqueur PIR protein situé sur le Chromosome 06 (PIR_Chr 06) avec 11 SNPs. Le marqueur conserved *Plasmodium* protein

CPP sur le Chromosome 03 (CPP_Chr 03) et le marqueur *Plasmodium* exported protein PXP sur le chromosome 08 (PXP_Chr08) avec 9 SNPs chacun. Le marqueur conserved *Plasmodium* protein CPP sur le chromosome 13 (CPP_Chr13) qui présente 6 SNPs et enfin le marqueur STP1 protein situé sur le Chromosome 05 (STP1_Chr05) qui possède 5 SNPs.

Puis cet outil a été appliqué sur 20 échantillons d'ADN de patients provenant du Pérou (n=12) et de Papouasie Nouvelle Guinée(PNG) (n=8), deux pays endémiques au paludisme à *P. vivax*. Des informations sur l'âge, le sexe sont recueillies pour certaines individus. De plus, le taux de parasitémie à *Plasmodium vivax* de chacun des patients analysés a été enregistré ainsi que les valeurs des résultats de la multiplicité d'infection MOI issus de analyse de 28 marqueurs microsatellites pour 4 patients du Pérou (Tableau 01).

Tableau 01 : Les données recueillies sur les échantillons testés.

Nom de l'échantillon	Pays	Sexe*	Age (ans)	Parasitémie Pv /copy /uL	MS_MOI
AAKC041	PNG	1	53	94.31797294	NA
AAKC118	PNG	2	33	42.58300862	NA
AAKE013	PNG	2	40	2.39047211	NA
AAKE038	PNG	1	12	3901.80936	NA
AAKE030	PNG	1	14	9.749210226	NA
AAKC094	PNG	2	8	23.54672707	NA
AAKC033	PNG	2	4	190.5728616	NA
Pv6	Pérou	NA	NA	13167.56782	NA
AAKC027	PNG	1	29	7.849902928	NA
Pv8A	Pérou	NA	NA	29.4168647	NA
Pv6A	Pérou	NA	NA	28.69079049	NA
Pv8	Pérou	NA	NA	44246.16977	NA
AL019832	Pérou	2	17	1862.517632	1
AL020379	Pérou	2	44	75.23273168	1
AL022178	Pérou	1	13	23.13878314	2
AL019574	Pérou	2	12	13.63303591	2
AC005998	Pérou	2	42	630.7326902	NA
AC009316	Pérou	1	42	75.43398713	NA
AL020433	Pérou	1	66	0.958396692	NA
AC006302	Pérou	2	38	5.883443372	NA

Pv : *Plasmodium vivax*, MS : microsatellites analysis , MOI : Multiplicité Of Infection. NA : Non Appliqué. PNG: Papouasie Nouvelle Guinée. 1* : Male, 2* : Femelle.

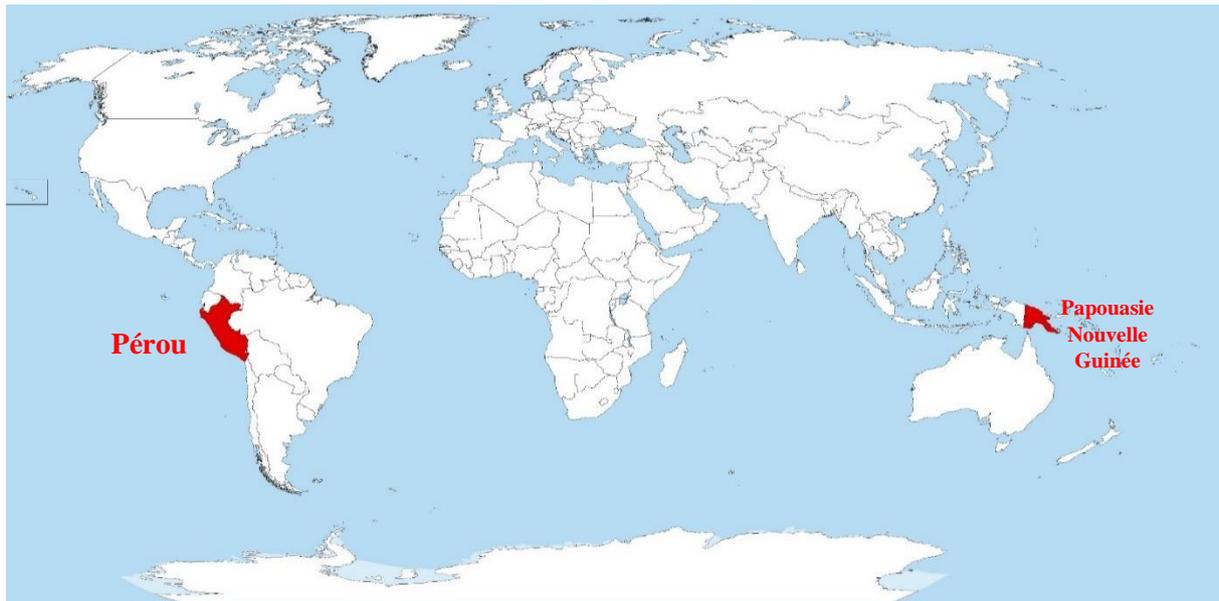


Figure 05 : Carte géographique des localités d'échantillons analysés dans cette étude.

1-3- Séquençage de nouvelle génération (NGS)

Après optimisation des marqueurs moléculaires, le séquençage de nouvelle génération a été effectué en utilisant la plateforme Illumina MiSeq pour les 20 échantillons du Pérou et de PNG . Pour cela, une amplification d'ADN par PCR primaire et secondaire a été effectuée pour tous les échantillons comme indiqué précédemment. Par la suite, une dernière amplification (Library PCR) a été effectuée en utilisant des amorces qui se lient à la séquence de liaison « linker » F et R à l'extrémité 3', qui a introduit une séquence de code-barres moléculaire « barcode Sequence » spécifique à l'échantillon plus l'adaptateur de séquence Illumina à l'extrémité 5' et cela pour permettre le regroupement et le démultiplexage ultérieur des amplicons. Chaque amplicon a été séquençé en double (duplicata) (Les détails du Protocole de séquençage ont été décrits par Lerche et al (2017)).

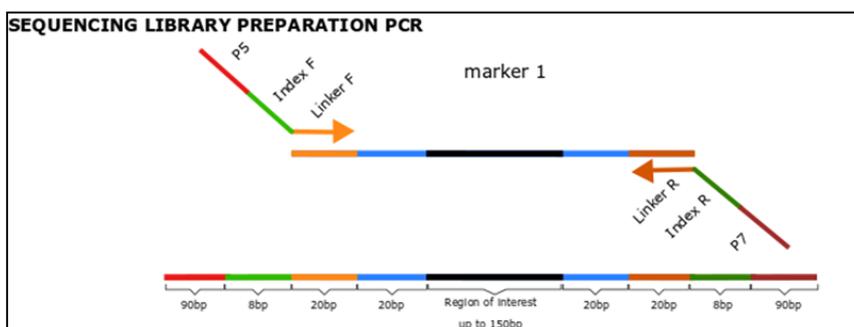


Figure 06 : Représentation schématique des amorces Library (Library PCR).

1-4- Pipeline d'analyse bio-informatique «HaplotypR package »

Les résultats du séquençage ont été obtenus après suppression du bruit de séquençage. Pour cela, HaplotypR a été utilisé et il a permis de démultiplexer les fichiers obtenus du séquenceur. Dans le processus de démultiplexage, « HaplotypR » a fusionné les lectures des fichiers Forward et Reverse afin d'obtenir un contig par échantillon et par marqueur. HaplotypR appelle également les haplotypes en examinant la fréquence des SNPs dans les marqueurs de tous les échantillons. Ensuite, le nettoyage des données a été procédé en supprimant tous les haplotypes avec moins de 25 lectures et ceux classés comme «Indel», «Chimera», «Noise», «Singleton», comme cela a été décrit par Lerche et ses collaborateurs (2017). Lors de l'analyse de la diversité génétique, les haplotypes n'ont été comptés comme véritables haplotypes que si les deux duplicata dépassent le « Cut-off = 25 lectures par échantillon » fixé comme de vrai haplotypes. Ce critère plus strict a été introduit pour éviter une surestimation erronée de la multiplicité d'infection due à de faux haplotypes (Lerche et al., 2017).

2- Analyses statistiques

Les données ont été saisies sur le Microsoft Excel et les analyses statistiques ont été effectuées avec le logiciel statistique R studio et le Microsoft Excel.

La multiplicité des infections (MOI) pour toute la population a été calculée en divisant la moyenne de l'ensemble des haplotypes détectés pour chaque marqueurs par le nombre d'échantillons positifs pour les mêmes marqueurs. Et la multiplicité d'infection par individu et par marqueur est égale à la valeur de l'ensemble des haplotypes.

L'hétérozygotie (H_e) représente la probabilité d'être infecté par deux parasites avec des allèles différents à un moment donné. Elle est comprise entre 0 et 1, et calculée en utilisant la formule suivante: $H_e = [n / (n-1)] [(1-\sum p_i^2)]$, où n est le nombre d'échantillons testés et p_i est la fréquence de l'allèle i (Lerch et al., 2017).

Le test de la régression linéaire a été utilisé pour étudier la corrélation entre la MOI est l'âge des individus. Et entre le taux de parasitémie à *Plasmodium vivax* et le nombre de lectures obtenues par AmpSeq. Pour cela, coefficient de corrélation ($corr$) a été calculé pour ces variables quantitatives.

Le test de student a été utilisé pour vérifier une éventuelle association entre la MOI est le sexe des patients, la différence est estimée significative si la valeur de ($P < 0.05$).

III. Résultats

1- Les marqueurs moléculaires optimisés dans cette étude

Nous avons optimisé 4 marqueurs moléculaires sur 4 chromosomes du *Plasmodium vivax*. Le marqueur_ AMA1_Chr_09 est le plus polymorphique et possède 17 SNPs suivi du MSP3 Chr_10 avec 15 SNPs, le marqueur lysophospholipase_ Chr_02 avec 11 SNPs et enfin GILP_Chr_11 avec 6 SNPs.

Certaines images des résultats d'électrophorèse sur gel d'agarose pour les PCR primaires, les PCR secondaires et les PCR multiplexes sont illustrées sur les figures 07 et 08.

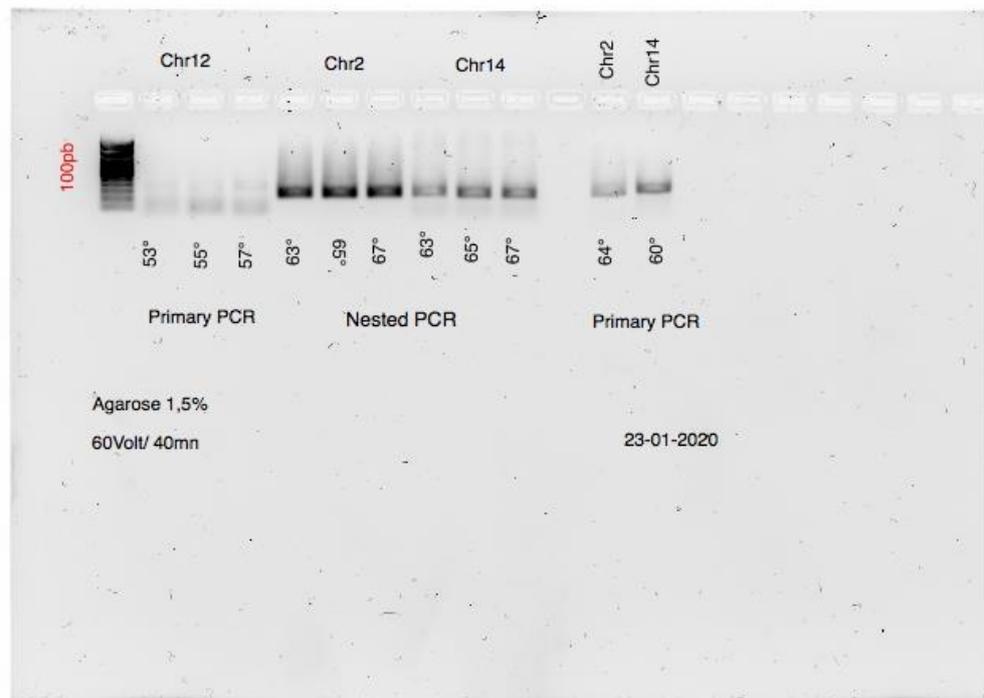


Figure 07 : Image de migration en électrophorèse (PCR primaire et secondaire).

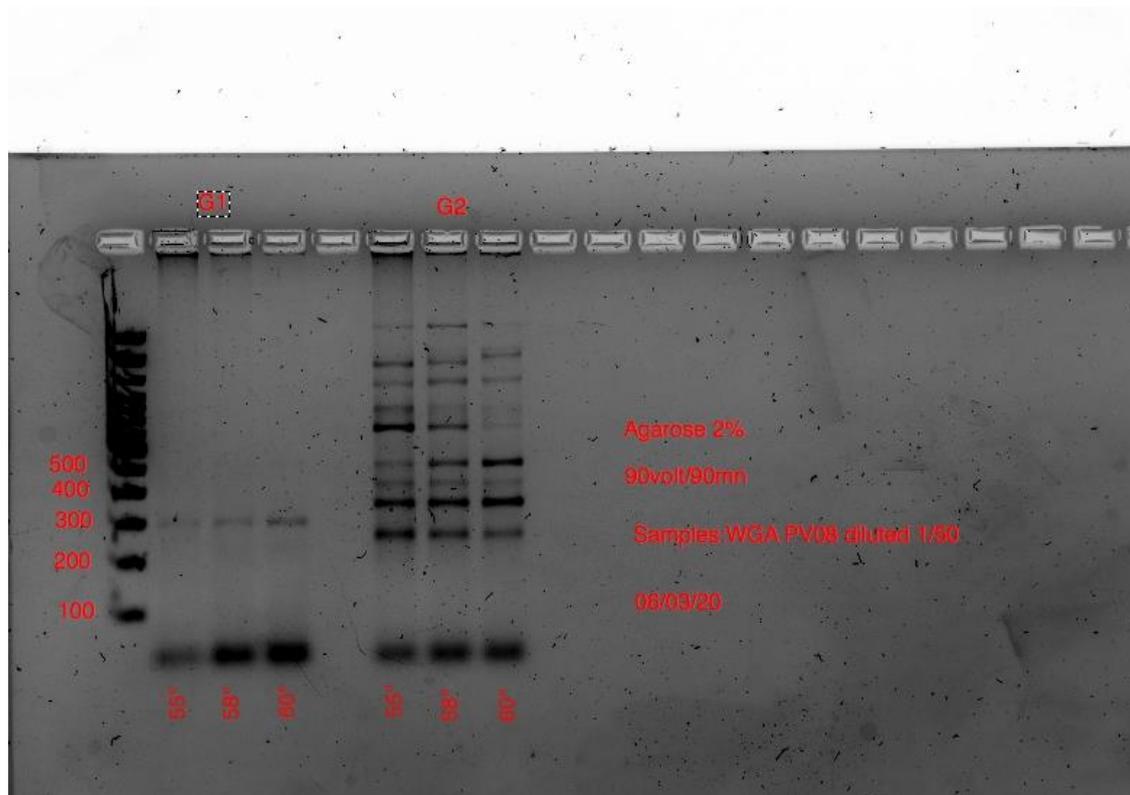


Figure 08 : Image de migration en électrophorèse (PCR multiplexe).

Avant de procéder au séquençage d'amplicon par le séquençage de nouvelle génération (NGS), les résultats du séquençage de Sanger ont confirmé l'identité des séquences amplifiées avec celles des marqueurs d'intérêts désignés pour *Plasmodium vivax* (Alignement des séquences est illustré dans l'annexe 01).

2- Présentation des résultats d'un exemple d'AmpSeq de six marqueurs déjà optimisés

2-1- Description des marqueurs moléculaires

Un protocole de séquençage d'amplicons courts (AmpSeq) en utilisant le séquençage de nouvelle génération (NGS) a été développé par notre équipe sur six marqueurs moléculaires désignés dont la taille est de 200pb environs et qui possèdent un haut polymorphisme. Le marqueur moléculaire *msp1* situé sur le chromosome 07 (*msp1*_Chr07) présente le plus haut polymorphisme avec 37 (SNPs) suivi de marqueur PIR protein situé sur le Chromosome 06 (PIR_Chr 06) avec 11 SNPs. Le marqueur conservé *Plasmodium* protein CPP sur le Chromosome 03 (CPP_Chr03) et le marqueur *Plasmodium* exported protein PXP sur le chromosome 08 (PXP_Chr08) avec 9 SNPs chacun, le marqueur conservé *Plasmodium*

protein CPP sur le chromosome 13 (CPP_Chr13) présente 6 SNPs et enfin le marqueur STP1 protein situé sur le Chromosome 05 (STP1_Chr05) possède 5 SNPs.

2-2- Diversité des haplotypes par marqueur et par échantillons

En utilisant cet outil, 240 réactions de PCR (en duplicata) ont été effectuées et séquencées à la fois sur 20 échantillons positifs à *Plasmodium vivax* provenant de Pérou (n=12) et de Papouasie Nouvelle Guinée (n=08). Les résultats de séquençage et l'analyse des données ont permis de détecter des haplotypes pour chacun des marqueurs moléculaires. La confirmation des haplotypes a été effectuée pour chaque amplicon et son duplicata afin d'éviter des multiplicités d'infections dues aux faux haplotypes.

Au sein des échantillons testés , le marqueur *msp1*_Chro07 a révélé le plus grand nombre d'haplotypes (n= 11), suivi par STP1_Chro05 (n= 9), PIR_Chro06 avec 7 haplotypes, CPP_Chro13 avec 5 haplotypes, CPP_Chro03 avec 3 et enfin le chromosome PXP_Chro08 (n= 2) (Figure 09).

Les génomes originaires de PNG ont révélé un plus grand nombre d'haplotypes pour le marqueur *msp1*_Chro07(n=07) par rapport à ceux originaires du Pérou (n=05). Tandis ce que ces derniers (Pérou) présentent un nombre d'haplotypes supérieurs avec le marqueur PIR_Chro06 et CPP_Chro13 (n= 5, 5) par rapport au génomes originaires de PNG (n= 3, 4) respectivement. Les 3 marqueurs CPP_Chro03, STP1_Chro05 et PXP_Chro08 ont révélé nombre similaire d'haplotypes au seins des deux populations testées (n= 3, 6 et 2) respectivement (Figure 10).

Au sein de la population de PNG, Le Marqueur *msp1*_Chro07 présente la diversité la plus élevée avec une heterozygocité attendue H_e égale à 0.80, suivi par le marqueur STP1_Chro05 ($H_e=0.76$) et PIR_Chro06 ($H_e=0.53$).

Pour les échantillons du Pérou, STP1_Chro05 présente la diversité la plus élevée une heterozygocité ($H_e=0.81$) suivi par le Marqueur *msp1*_Chro07 ($H_e=0.78$) et enfin le marqueur et PIR_Chro06 ($H_e=0.75$).

Nos résultats ont révélé que 25% (3/12) des échantillons péruviens ont présenté deux haplotypes révélés par les marqueurs *msp1*_Chro07 , STP1_Chro05, PIR_Chro06, CPP_Chro03 , et CPP_Chro13 (Figure 11).

Pour les échantillons PNG, la diversité génétique était plus importante, en effet, (2/8) des patients ont présenté 3 haplotypes révélés par le marqueurs *msp1*_Chro07 (Figure 12).

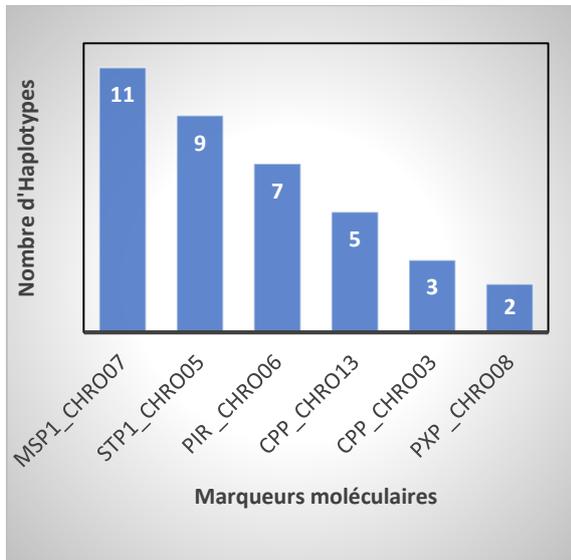


Figure 09 : Nombre d'haplotypes par marqueur dans tous les échantillons

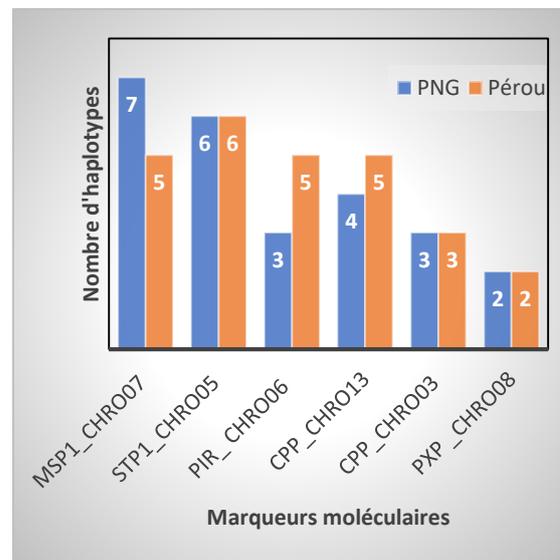


Figure 10 : Nombre d'haplotypes par marqueur au Pérou et au PNG.

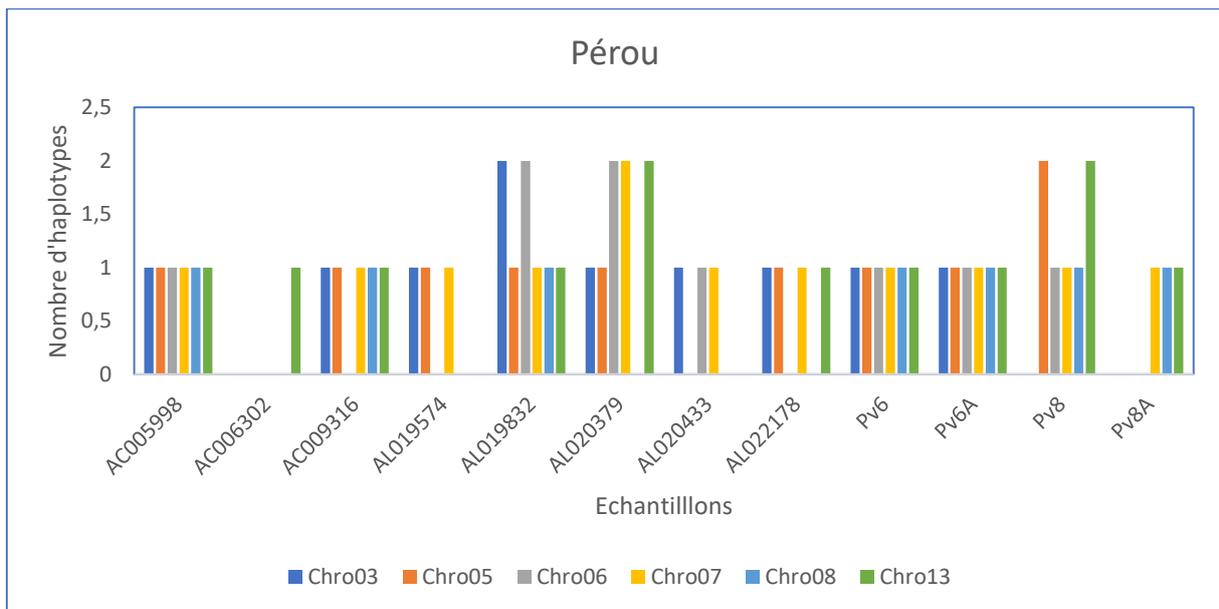


Figure 11: Nombre d'haplotypes par marqueur par échantillon du Pérou.

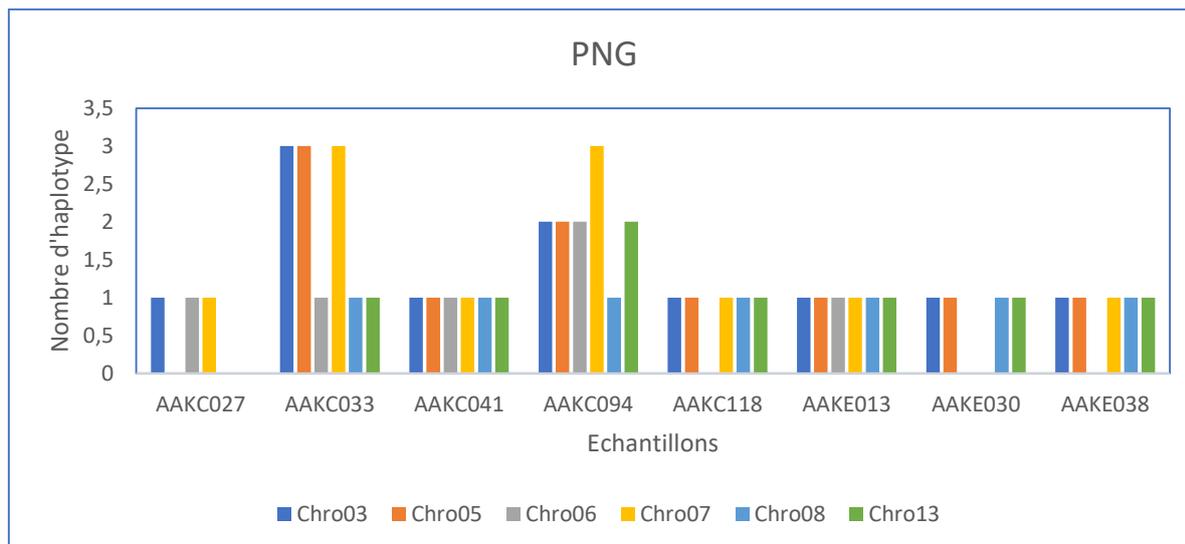


Figure 12 : Nombre d'haplotypes par marqueur par échantillon de PNG.

2-3- Diversité intra-hôte

La diversité intra-hôte représente la fréquence de chaque haplotype par échantillon. Elle a été calculée par le nombre de lectures de chaque haplotype divisé par le nombre total de lectures dans cet échantillon. Nous avons calculé la fréquence des haplotypes par marqueurs pour les 5 échantillons qui ont un nombre d'haplotypes > 1 (Tableau 02). Cela nous permet de distinguer les clones dominants avec une fréquence élevée et les clones minoritaires avec une fréquence faible (La figure 13 montre les clones dominants et les clones minoritaires d'un patient Péruvien et un autre de PNG en fonction des fréquences de leurs haplotypes pour les différents marqueurs). Il est à noter que les clones minoritaires ont été détectés à de très faibles fréquences allant jusqu'à 1% (Tableau 02, Figure13).

Tableau 02 : La fréquence des haplotypes (Diversité intra-hôte).

Pays	Echantillon	Chromosome	Haplotypes	N° lecture	Total lecture	Fréquence des Haplotypes	
Pérou	AL019832	Chr03	Chr03-1	942	4977	0,19	
			Chr03-3	4035		0,81	
		Chr06	Chr06-4	705		4928	0,14
			Chr06-5	4223			0,86
	AL020379	Chr06	Chr06-4	603	4948	0,12	
			Chr06-7	4345		0,88	
		Chr07	Chr07-11	2937	4851	0,6	
			Chr07-4	1914		0,4	
		Chr13	Chr13-1	662	4993	0,13	
			Chr13-2	4331		0,87	

Pv8	Chr05	Chr05-4	4862	4938	0,98	
		Chr05-7	76		0,02	
	Chr13	Chr13-3	4909	4985	0,98	
		Chr13-4	76		0,02	
PNG	AAKC033	Chr03	Chr03-1	767	4910	0,16
			Chr03-2	1007		0,2
			Chr03-3	3136		0,64
	Chr05	Chr05-1	4593	4937	0,93	
		Chr05-5	273		0,06	
		Chr05-6	71		0,01	
	Chr07	Chr07-5	1287	4830	0,27	
		Chr07-6	2346		0,48	
		Chr07-7	1197		0,25	
	AAKC094	Chr03	Chr03-1	4470	4991	0,9
			Chr03-2	521		0,1
		Chr05	Chr05-1	4197	4991	0,84
			Chr05-2	794		0,16
		Chr06	Chr06-1	645	2366	0,27
			Chr06-6	1721		0,73
		Chr07	Chr07-3	2833	4912	0,58
Chr07-4			675	0,14		
Chr07-7			1404	0,28		
Chr13		Chr13-3	1969	4994	0,39	
		Chr13-4	3025		0,61	

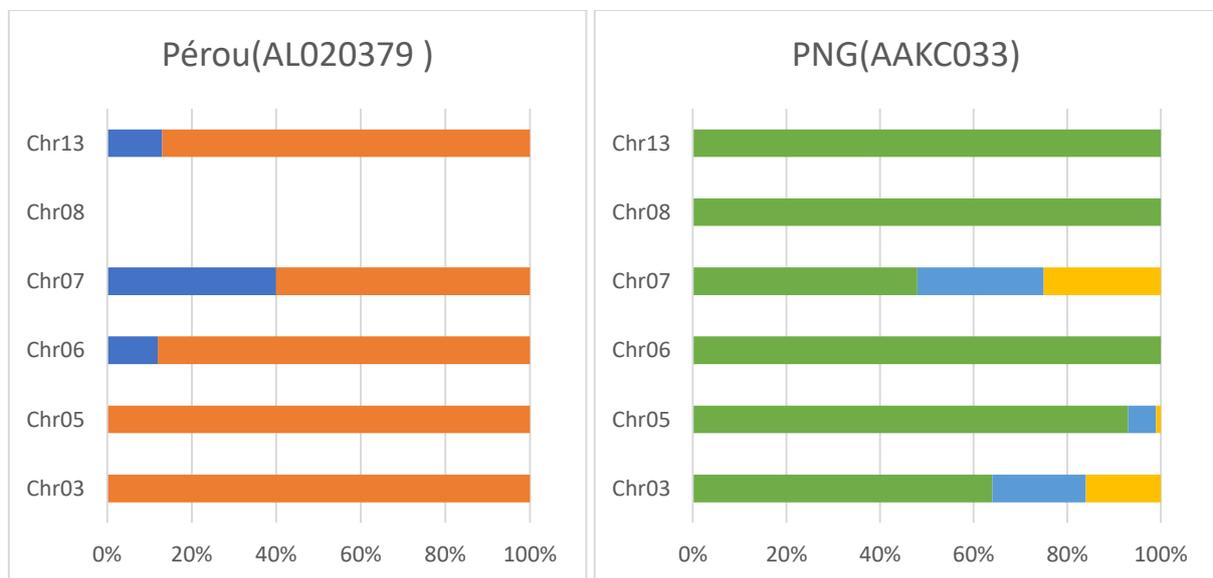


Figure 13 : Fréquence des Haplotypes au seins de deux cas du Pérou et de PNG.

2-4- Multiplicité d'infection (MOI)

Nous avons calculé la MOI qui est égale au nombre d'haplotypes par marqueur dans un échantillon donné pour les 6 marqueurs optimisés. Les valeurs obtenus pour chaque marqueur ainsi que pour chaque individu de PNG et du Pérou sont illustrées dans les

tableaux 03 et 04). AmpSeq a permis de détecter 2 haplotypes dans certains échantillons péruviens et jusqu'à 3 haplotypes chez certains échantillons de PNG principalement par les 2 marqueurs *mspI*_Chr07 et STP1_Chro05.

Tableau 03: MOI par individu de la population du Pérou.

Pérou	AC0059 98	AC0063 02	AC0093 16	AL0195 74	AL0198 32	AL0203 79	AL0204 33	AL0221 78	Pv 6	Pv6 A	Pv 8	Pv8A
Chr0 3	1	NA	1	1	2	1	1	1	1	1	NA	NA
Chr 05	1	NA	1	1	1	1	NA	1	1	1	2	NA
Chr0 6	1	NA	NA	NA	2	2	1	NA	1	1	1	NA
Chr0 7	1	NA	1	1	1	2	1	1	1	1	1	NA
Chr0 8	1	NA	1	NA	1	NA	NA	NA	1	1	1	1
Chr1 3	1	1	1	NA	1	2	NA	1	1	1	2	1
Max MOI	1	1	1	1	2	2	1	1	1	1	2	1

NA : Non appliqué

Tableau 04 : MOI par individu de la population de PNG.

PNG	AAKC02 7	AAKC03 3	AAKC04 1	AAKC09 4	AAKC11 8	AAKE01 3	AAKE03 0	AAKE03 8
Chro0 3	1	3	1	2	1	1	1	1
Chro0 5	NA	3	1	2	1	1	1	1
Chro0 6	1	1	1	2	NA	2	NA	NA
Chro0 7	1	3	1	3	1	1	NA	1
Chro0 8	NA	1	1	1	1	1	1	1
Chro1 3	NA	1	1	2	1	1	1	1
Max MOI	1	3	1	3	1	2	1	1

NA : Non appliqué

2-4-1- Comparaison de la MOI par AmpSeq et les données de MS (microsatellites)

Par la suite, nous avons comparé la MOI pour le marqueur le plus polymorphique (*mspI*_Chr07) à celle rapportée pour 4 patients péruviens par l'analyse basée sur de 28 marqueurs microsatellites. Nous avons constaté que la MOI moyenne obtenue par Miseq (MOI=1.25) est légèrement inférieure à celle obtenue par l'analyses basées sur microsatellites (MOI= 1.5).

Tableau 05 : Comparaison entre MOI entre MS et AmpSeq.

Echantillons	MS	Miseq (<i>msp1_Chr07</i>)
AL019832	1	1
AL020379	1	2
AL022178	2	1
AL019574	2	1
Mean(MOI)	1.5	1.25

2-4-2- Association entre la MOI l'âge et le sexe

D'après nos résultats , Il n'y a pas de corrélation entre multiplicité d' infections et l'âge des individus testés ($corr=0.4$, $P\ value=0.12$). De même pour le sexe, le test de student a montré qu'il n'y a pas différence significative entre multiplicité d'infection et les deux sexes ($P =0.57$).

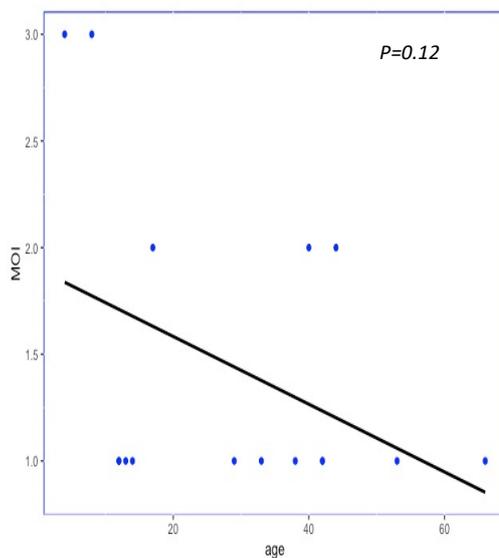


Figure 14 : Corrélation entre MOI et âge

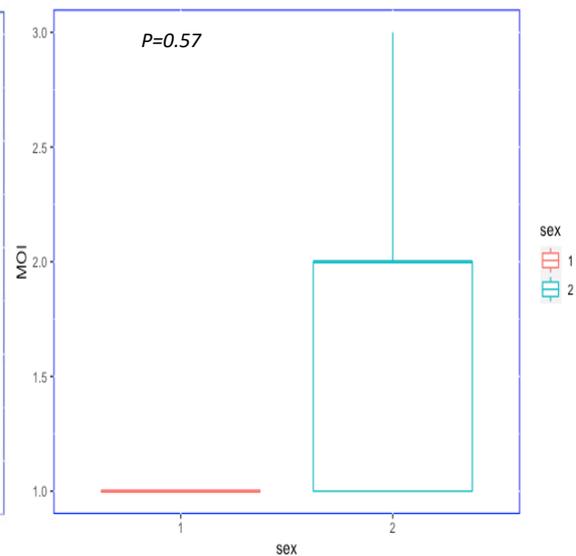


Figure 15: Association entre MOI et le sexe.

2-4-3- Corrélation entre la parasitémie et le nombre de lecture

Nous avons recherché une éventuelle existence de corrélation entre le nombre de lectures obtenues suite aux analyses des résultats de NGS pour chacun des échantillons séquencés et leurs taux de parasitémie à *P. vivax*. En utilisant la régression linéaire, le calcul du coefficient de corrélation a confirmé que y'a pas de corrélation entre la densité parasitaire et le nombre de

lecture pour les échantillons de PNG (Corr=0.01 , P=0.76) et les échantillons provenant du Pérou (Corr=0.16, P=0.10). Nous avons obtenu des lectures quel que soit le taux de parasitémie des échantillons analysés. En effet, des échantillons à faible densité parasitaire ont été également séquencés.

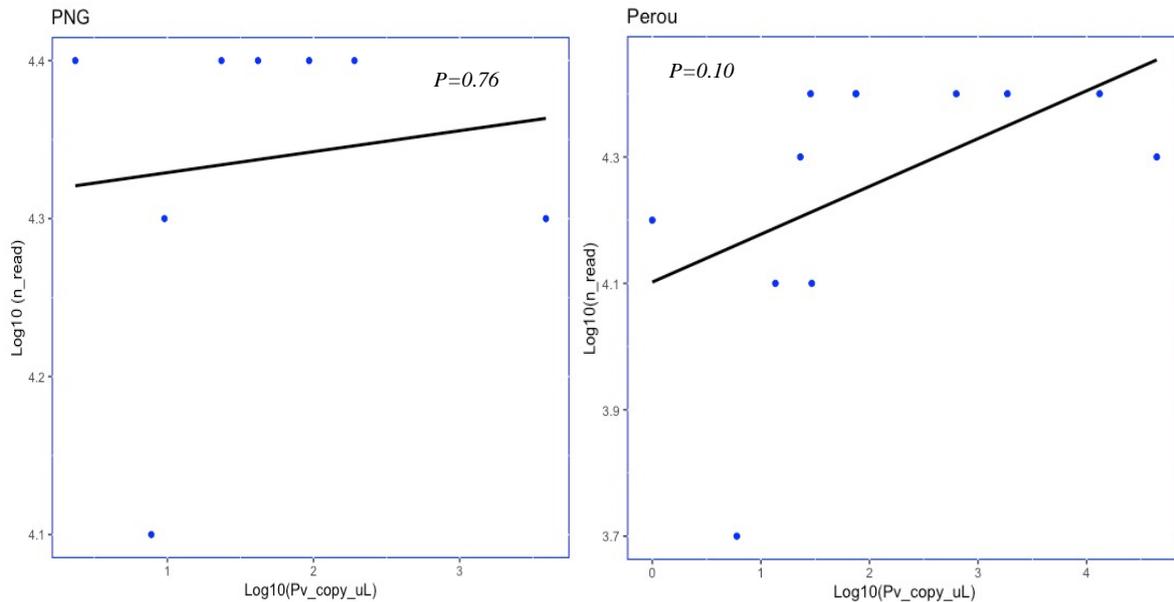


Figure 16: Corrélation entre la parasitémie et le nombre de lecture.

VI- Discussion

Ce présent travail a pour but de développer un nouvel outil de génotypage pour *Plasmodium vivax* basé sur le séquençage des amplicons courts. La recherche et l'optimisation de nouveaux marqueurs a été procédées pour remédier à certaines limites des marqueurs polymorphes de longueur « microsatellites », utilisés habituellement pour le génotypage des parasites du genre *Plasmodium*. Cela est due principalement à une amplification préférentielle des fragments plus courts dans les infections multiclonales, De ce fait, les clones minoritaires ne sont pas détectés (Messerli et al., 2017a). Pour cela, de nouveaux protocoles de séquençage des amplicons sont nécessaires pour le génotypage dans les études épidémiologiques et les essais d'efficacité médicamenteuse de *P. vivax* surtout dans les zones où le paludisme est endémique. Dans ce contexte, Une approche alternative a été développée, elle est basée sur plusieurs SNPs plutôt que des marqueurs microsatellites qui consiste au séquençage d'amplicons courts et qui permet une détection sensible des clones minoritaires et améliore le pouvoir discriminant pour le génotypage des infections multiclonales à *Plasmodium vivax*.

Cette approche basée sur les SNPs a déjà été développée pour *Plasmodium falciparum* ciblant des marqueurs moléculaires courts mais très polymorphiques *csp* et *cpmp*. Ce nouvel outil a montré une plus grande sensibilité dans la détection des clones minoritaires comparé aux marqueurs microsatellites et il permet de détecter de manière fiable les haplotypes à de très basses fréquences (Lerch et al., 2017). Le séquençage d'amplicons est aussi largement appliqué dans d'autres domaines par exemple dans les études de la diversité génétique d'autres agents pathogènes notamment les virus (Beerenwinkel et al., 2012).

Pour *Plasmodium vivax*, l'AmpSeq a déjà été développé ciblant une seule région très polymorphiques de 117 pb du gène *mssl* et validé sur 118 isolats de *P. vivax* provenant de Cambodge (Lin et al., 2015). Les résultats présentés dans ce manuscrit décrivent un nouvel outil d'AmpSeq basé sur les SNPs en ciblant différentes régions polymorphes au sein de différents chromosomes de *Plasmodium vivax*.

Au cours de cette étude, nous avons réussi à optimiser 4 nouveaux marqueurs moléculaires situés sur 4 différents chromosomes du *Plasmodium vivax*. Il est intéressant de signaler la difficulté d'amplifier le génome de *Plasmodium vivax*, cela est due au faible taux de parasitémie. En effet, *P. vivax* colonise généralement les réticulocytes qui représentent uniquement 1 à 2% des érythrocytes (Chan et al., 2020).

D'après les résultats antérieurs d'AmpSeq obtenus par notre équipe, le marqueur moléculaire *mssl* situé sur le chromosome 7 contenant 37 SNPs a montré la plus haute diversité avec 11 haplotypes détectés sur les 20 échantillons du terrain testés. Nous suggérons qu'avec un nombre croissant d'échantillons analysés, d'autres haplotypes sont susceptibles d'être trouvés. *mssl* exprimé à la surface des mérozoïtes a montré une très haute diversité dans d'autres études antérieures menées sur *Plasmodium vivax* et *P. falciparum* (Zhong et al., 2018a ; Mwingira et al., 2011) suggérant que *mssl* est le premier candidat.

De nouveaux marqueurs optimisés ont montré aussi une haute diversité au sein des échantillons testés. Il s'agit de STP1_Chro05 avec 9 haplotypes, PIR_Chro06 avec 7 haplotypes, CPP_Chro13 avec 5 haplotypes. Tant dis qu'une diversité génétique faible a été observée pour le marqueur CPP_Chro03 et le PXP_Chro08 avec uniquement 2 et 3 haplotypes respectivement dans les 20 échantillons testés. Ainsi, ces deux derniers ne sont pas adaptés pour servir de marqueurs de génotypage unique au PNG et au Pérou.

L'AmpSeq a montré aussi une sensibilité très élevée. En effet, il a le potentiel de détecter des génotypes de *P. vivax* dans des échantillons à faible parasitémie de l'ordre de 5 copie de *P.*

vivax / μ l (Tableau 01) ainsi que de détecter des infections multiclonales avec de clones minoritaires de l'ordre de fréquence des 1% (Figure13).

Aucune corrélation n'a été rapportée entre la densité parasitaire et le nombre de lecture (Figure 16). En effet, un Protocole de quantification en utilisant le kit PicoGreen® dsDNA kit (Thermo Fisher Scientific) puis la normalisation qui consiste à diluer les produits Nested PCR à 15 ng / μ l et enfin la mise en commun en quantité équimolaire d'amplicons des produits de PCR nichées a été effectué avant le séquençage. Il est à noter que les infections à *Plasmodium vivax* sont caractérisées par un faible taux de parasitémie ce qui est très difficile de travailler avec ce genre d'échantillon. Grâce à ce protocole de normalisation, nous pouvons séquencer des échantillons de faible densité parasitaire sans perdre de lectures.

La multiplicité d'infection (MOI) est définie comme le nombre de différentes souches parasitaires co-infectant un seul hôte, elle peut être un indicateur utile pour estimer le niveau de transmission dans les pays endémiques (Zhong et al., 2018a). D'après les résultats obtenus, les infections multiclonales sont plus fréquentes en PNG par rapport à Pérou. Ce qui indique que la transmission de *P. vivax* est plus importante au PNG qu'au Pérou. Ces résultats sont en accord avec une autre étude menée au PNG où ils ont rapporté une corrélation positive entre le taux d'infections polyclonales et l'incidence annuelle de l'infection (Zhong et al., 2018b).

Traditionnellement, la MOI était évaluée par le génotypage des marqueurs microsatellites, qui étaient considérés comme l'or standard en raison de leur polymorphisme élevé (Lalremruata et al., 2017). Cependant, cette méthode n'a pas pu distinguer les variations de séquences parmi les souches parasitaires et détecter les clones minoritaires au sein d'un hôte (Messerli et al., 2017b). Dans cette étude, malgré que la MOI moyenne obtenue pour 4 individus par AmpSeq est légèrement inférieure (MOI = 1.25) à celle obtenue par MS (MOI=1.5) chez ces mêmes individus (Tableau 05), cela ne veut pas discriminer la sensibilité de AmpSeq dans la détection des infections multiclonales. En effet, elle pourrait être expliquée par le fait que la MOI (AmpSeq) a été calculée uniquement pour un seul marqueur moléculaire *mssl_Chr07*, par contre la MOI(MS) a été calculée pour 28 différents marqueurs microsatellites. Ajoutant que la MOI a été estimée uniquement pour 4 individus et cet effectif est très réduit et non représentatif de la population pour évaluer la sensibilité de AmpSeq. Pour cela, nous suggérons d'autres études afin d'évaluer AmpSeq sur un plus grand nombre d'échantillons en utilisant plus de marqueurs moléculaires. D'ailleurs, les résultats d'une étude menée en 2015 sur 118 isolats de *P. vivax* issus 78 cambodgiens qui ont subi des rechutes à *P. vivax* ont démontré que la MOI moyenne était égale 3.6 par le séquençage d'amplicon de 117pb du gène *mssl*

comparée à MOI= 2,1 obtenue par une combinaison de 3 marqueurs microsatellites(Lin et al., 2015). De plus, une étude comparative entre la MOI par le génotypage basé sur d'amplicon court du marqueur *cpmp* du *P. falciparum* et le séquençage de longueur du marqueur *msh2* a montré que la MOI moyenne par *cpmp* est légèrement supérieure (2.41) à celle rapportée par le marqueur *msh2* (2.19)(Lerch et al., 2017). Ce qui confirme que AmpSeq est très puissant dans la detection des infections polyclonales dans les zones d'endémicité élevée du paludisme.

Dans cette étude, aucune corrélation n'a été trouvée entre la MOI et l'âge des patients, ainsi qu'entre les groupes masculins et féminins, aucune différence significative n'a été trouvée. Des résultats similaires ont été rapportés dans d'autres études (Zhong et al., 2018a ; Pacheco et al., 2016). Cependant, dans les zones de transmission élevée, la complexité de l'infection a été suggérée comme étant associée à l'âge (Waltmann et al., 2015). Ça était expliquée par le fait que le immunité anti- *Plasmodium vivax* a tendance à être acquise avec l'âge. D'ailleurs les enfants de moins de 5 ans constitue la catégorie la plus sensible et par conséquent la plus touchée par ce parasite.

V- Conclusion et perspectives

Cette étude a démontré que le séquençage d'amplicons courts basés sur les SNPs est un outil puissant et très sensible pour estimer la MOI à *Plasmodium vivax* à des taux de parasitémie faible et étudier la diversité génétique intra-hôte. Ce travail représente une première étape vers un objectif important, Pour cela, il serait utile d'explorer les autres marqueurs moléculaires optimisés et d'étendre ces investigations à d'autres régions endémiques et à des cohortes plus grandes afin de mieux comprendre la dynamique des infections à *Plasmodium vivax*.

Références bibliographiques

1. Adams, J.H., Mueller, I., 2017. The biology of *Plasmodium vivax*. Cold Spring Harb. Perspect. Med. 7.
2. Anderson, T.J.C., Su, X.Z., Bockarie, M., Lagog, M., Day, K.P., 1999. Twelve microsatellite markers for characterization of *Plasmodium falciparum* from finger-prick blood samples. Parasitology 119, 113–125.
3. Bailey, J.A., Mvalo, T., Aragam, N., Weiser, M., Congdon, S., Kamwendo, D., Martinson, F., Hoffman, I., Meshnick, S.R., Juliano, J.J., 2012. Use of massively parallel pyrosequencing to evaluate the diversity of and selection on *Plasmodium falciparum* csp T-cell epitopes in Lilongwe, Malawi. J. Infect. Dis. 206, 580–587.
4. Beerenwinkel, N., Günthard, H.F., Roth, V., Metzner, K.J., 2012. Challenges and opportunities in estimating viral genetic diversity from next-generation sequencing data. Front. Microbiol. 3, 1–16.
5. Bright, A.T., Manary, M.J., Tewhey, R., Arango, E.M., Wang, T., Schork, N.J., Yanow, S.K., Winzeler, E.A., 2014. A High Resolution Case Study of a Patient with Recurrent *Plasmodium vivax* Infections Shows That Relapses Were Caused by Meiotic Siblings. PLoS Negl. Trop. Dis. 8.
6. Chan, L.J., Dietrich, M.H., Nguiragool, W., Tham, W.H., 2020. *Plasmodium vivax* Reticulocyte Binding Proteins for invasion into reticulocytes. Cell. Microbiol. 22, 1–11.
7. De Araujo, F.C.F., de Rezende, A.M., Fontes, C.J.F., Carvalho, L.H., Alves de Brito, C.F., 2012. Multiple-Clone Activation of Hypnozoites Is the Leading Cause of Relapse in *Plasmodium vivax* Infection. PLoS One 7, 1–9.
8. Diallo ,M.L. P., 2018. Etudiants subsahariens à Toulouse et prévention du paludisme d'importation. Thèse d'université Paul Sabatier Toulouse.1-78.
9. Escalante, A.A., Ferreira, M.U., Vinetz, J.M., Volkman, S.K., Cui, L., Gamboa, D., Krogstad, D.J., Barry, A.E., Carlton, J.M., van Eijk, A.M. ari., Pradhan, K., Mueller, I., Greenhouse, B., Pacheco, M.A., Vallejo, A.F., Herrera, S., Felger, I., 2015. Malaria Molecular Epidemiology: Lessons from the International Centers of Excellence for Malaria Research Network. Am. J. Trop. Med. Hyg. 93, 79–86.
10. Escalante, A.A., Pacheco, M.A., 2019. Malaria Molecular Epidemiology: An Evolutionary Genetics Perspective *. Microbiol. Spectr. 7, 1–26. <https://doi.org/10.1128/microbiolspec.ame-0010-2019>.
11. Friedrich, L.R., Popovici, J., Kim, S., Dysoley, L., Zimmerman, P.A., Menard, D., Serre, D., 2016. Complexity of Infection and Genetic Diversity in Cambodian *Plasmodium vivax*. PLoS Negl. Trop. Dis. 10, 1–18.
12. Greenhouse, B., Dokomajilar, C., Hubbard, A.E., Rosenthal, P.J., Dorsey, G., Gatton, L.M., Cheng, Q., Stepniewska, K., White, N.J., Collins, W.J., Greenhouse, B., Rosenthal, P.J., Dorsey, G., Myrick, A., Woo, J.M., Carlson, E.J., WHO, Snounou, G., Beck, H.P., Juliano, J.J., Gadalla, N., Sutherland, C.J., Meshnick, S.R., Dennis, M.K., Field, A.S., Burai, R., Ramesh, C., Whitney, K., Bologa, C.G., Oprea, T.I., Yamaguchi, Y., Hayashi, S., Sklar, L. a, Hathaway, H.J., Arterburn, J.B., Prossnitz, E.R., Cattamanchi, A., Kyabayinze, D., Hubbard, A.E., Rosenthal, P.J., Dorsey, G., Irion, A., Felger, I., Abdulla, S., Smith, T., Mull, R., Tanner, M., Hatz, C., Beck, H.P., Ranford-Cartwright, L., J, T., T, U., Lh, T., Ha, B., B, L., Jr, S.-O., Lg, L., D, W., Pg, K., Ohrt, C., Mirabelli-Primdahl, L., Looareesuwan, S., Wilairatana, P., Walsh, D., Kain, K.C., Brockman, A., Paul, R.E.L., Anderson, T.J.C., Hackford, I., Phaiphun, L., Looareesuwan, S., Nosten, F., Day, K.P., Viriyakosol, S., Siripoon, N., Petcharapirat, C., Petcharapirat, P., Jarra, W., Thaitong, S., Brown, K.N., Snounou, G., Gupta, V., Dorsey, G., Hubbard, A.E., Rosenthal, P.J., Greenhouse, B., Gatton, M.L., Cheng, Q., Kain, K.C., Craig, A. a., Ohrt, C., Slater, M., Kiggundu, M., Dokomajilar, C., Kamya, M.R., Bakayita, N., Talisuna, A., Rosenthal, P.J., Dorsey, G., Rouse, P., Mkulama, M.A., Thuma, P.E., Mharakurwa, S., 2006. Methods and techniques for clinical trials on antimalarial drug efficacy: genotyping to identify parasite populations. Malar. J. 5, 68–73.

13. Imwong, M., Boel, M.E., Pagornrat, W., Pimanpanarak, M., McGready, R., Day, N.P.J., Nosten, F., White, N.J., 2012. The first *Plasmodium vivax* relapses of life are usually genetically homologous. *J. Infect. Dis.* 205, 680–683.
14. Imwong, M., Snounou, G., Pukrittayakamee, S., Tanomsing, N., Kim, J.R., Nandy, A., Guthmann, J., Nosten, F., Carlton, J., Looareesuwan, S., Nair, S., Sudimack, D., Day, N.P.J., Anderson, T.J.C., White, N.J., 2007. Relapses of *Plasmodium vivax* Infection Usually Result from Activation of Heterologous Hypnozoites. *J. Infect. Dis.* 195, 927–933.
15. Ketema, T., Bacha, K., Birhanu, T., Petros, B., 2009. Chloroquine-resistant *Plasmodium vivax* malaria in serbo town, Jimma zone, south-west Ethiopia. *Malar. J.* 8, 1–8.
16. Lalremruata, A., Jeyaraj, S., Engleitner, T., Joanny, F., Lang, A., Bélard, S., Mombo-Ngoma, G., Ramharter, M., Kremsner, P.G., Mordmüller, B., Held, J., 2017. Species and genotype diversity of *Plasmodium* in malaria patients from Gabon analysed by next generation sequencing. *Malar. J.* 16, 1–11.
17. Lerch, A., Koepfli, C., Hofmann, N.E., Messerli, C., Wilcox, S., Kattenberg, J.H., Betuela, I., O'Connor, L., Mueller, I., Felger, I., 2017. Development of amplicon deep sequencing markers and data analysis pipeline for genotyping multi-clonal malaria infections. *BMC Genomics* 18, 1–13.
18. Lin, J.T., Hathaway, N.J., Saunders, D.L., Lon, C., Balasubramanian, S., Kharabora, O., Gosi, P., Sriwichai, S., Kartchner, L., Chuor, C.M., Satharath, P., Lanteri, C., Bailey, J.A., Juliano, J.J., 2015. Using Amplicon Deep Sequencing to Detect Genetic Signatures of *Plasmodium vivax* Relapse. *J. Infect. Dis.* 212, 999–1008.
19. Mendis, K., Sina, B.J., Marchesini, P., Carter, R., 2001. The neglected burden of *Plasmodium vivax* malaria. *Am. J. Trop. Med. Hyg.* 64, 97–106.
20. Messerli, C., Hofmann, N.E., Beck, H.P., Felger, I., 2017a. Critical evaluation of molecular monitoring in malaria drug efficacy trials and pitfalls of length-polymorphic markers. *Antimicrob. Agents Chemother.* 61, 1–11.
21. Messerli, C., Hofmann, N.E., Beck, H.P., Felger, I., 2017b. Critical evaluation of molecular monitoring in malaria drug efficacy trials and pitfalls of length-polymorphic markers. *Antimicrob. Agents Chemother.* 61, 1–11.
22. Mueller, I., Galinski, M.R., Baird, J.K., Carlton, J.M., Kochar, D.K., Alonso, P.L., del Portillo, H.A., 2009. Key gaps in the knowledge of *Plasmodium vivax*, a neglected human malaria parasite. *Lancet Infect. Dis.* 9, 555–566.
23. Mwingira, F., Nkwengulila, G., Schoepflin, S., Sumari, D., Beck, H.P., Snounou, G., Felger, I., Olliaro, P., Mugittu, K., 2011. *Plasmodium falciparum* msp1, msp2 and glurp allele frequency and diversity in sub-Saharan Africa. *Malar. J.* 10, 1–10.
24. Neafsey, D.E., Juraska, M., Bedford, T., Benkeser, D., Valim, C., Griggs, A., Lievens, M., Abdulla, S., Adjei, S., Agbenyega, T., Agnandji, S.T., Aide, P., Anderson, S., Ansong, D., Aponte, J.J., Asante, K.P., Bejon, P., Birkett, A.J., Bruls, M., Connolly, K.M., D'Alessandro, U., Dobaño, C., Gesase, S., Greenwood, B., Grimsby, J., Tinto, H., Hamel, M.J., Hoffman, I., Kamthunzi, P., Kariuki, S., Kremsner, P.G., Leach, A., Lell, B., Lennon, N.J., Lusingu, J., Marsh, K., Martinson, F., Molel, J.T., Moss, E.L., Njuguna, P., Ockenhouse, C.F., Ragama Ogutu, B., Otieno, W., Otieno, L., Otieno, K., Owusu-Agyei, S., Park, D.J., Pelle, K., Robbins, D., Russ, C., Ryan, E.M., Sacarlal, J., Sogoloff, B., Sorgho, H., Tanner, M., Theander, T., Valea, I., Volkman, S.K., Yu, Q., Lapierre, D., Birren, B.W., Gilbert, P.B., Wirth, D.F., 2015. Genetic diversity and protective efficacy of the RTS,S/AS01 malaria vaccine. *N. Engl. J. Med.* 373, 2025–2037.
25. Pacheco, M.A., Lopez-Perez, M., Vallejo, A.F., Herrera, S., Arévalo-Herrera, M., Escalante, A.A., 2016. Multiplicity of Infection and Disease Severity in *Plasmodium vivax*. *PLoS Negl. Trop. Dis.* 10, 1–18.
26. Price, R.N., Tjitra, E., Guerra, C. a, Yeung, S., White, N.J., Anstey, N.M., 2009. UKPMC Funders Group *Vivax* malaria : neglected and not benign 77, 79–87.
27. Rijken, M.J., Boel, M.E., Russell, B., Imwong, M., Leimanis, M.L., Phyto, A.P., Muehlenbachs, A., Lindegardh, N., McGready, R., Rénia, L., Snounou, G., Singhasivanon, P., Nosten, F., 2011. Chloroquine

- resistant *vivax* malaria in a pregnant woman on the western border of Thailand. *Malar. J.* 10, 1–5.
28. Rougeron, V., Elguero, E., Arnathau, C., Hidalgo, B.A., Durand, P., Houze, S., Berry, A., Zakeri, S., Haque, R., Alam, M.S., Nosten, F., Severini, C., Woldearegai, T.G., Mordmüller, B., Kremsner, P.G., González-Cerón, L., Fontecha, G., Gamboa, D., Musset, L., Legrand, E., Noya, O., Pumpaibool, T., Harnyuttanakorn, P., Lekweiry, K.M., Albsheer, M.M., Hamid, M.M.A., Boukary, A.O.M.S., Trape, J.F., Renaud, F., Prugnolle, F., 2020. Human *Plasmodium vivax* diversity, population structure and evolutionary origin. *PLoS Negl. Trop. Dis.* 14, 1–17.
 29. Ruebush, T.K., Zegarra, J., Cairo, J., Andersen, E.M., Green, M., Pillai, D.R., Marquiño, W., Huilca, M., Arévalo, E., Garcia, C., Solary, L., Kain, K.C., 2003. Chloroquine-resistant *Plasmodium vivax* malaria in Peru. *Am. J. Trop. Med. Hyg.* 69, 548–552.
 30. Snounou, G., Beck, H.P., 1998. The use of PCR genotyping in the assessment of recrudescence or reinfection after antimalarial drug treatment. *Parasitol. Today* 14, 462–467.
 31. Twohig, K.A., Pfeffer, D.A., Baird, J.K., Price, R.N., Zimmerman, P.A., Hay, S.I., Gething, P.W., Battle, K.E., Howes, R.E., 2019. Growing evidence of *Plasmodium vivax* across malaria-endemic Africa. *PLoS Negl. Trop. Dis.* 13.
 32. Waltmann, A., Darcy, A.W., Harris, I., Koepfli, C., Lodo, J., Vahi, V., Piziki, D., Shanks, G.D., Barry, A.E., Whittaker, M., Kazura, J.W., Mueller, I., 2015. High Rates of Asymptomatic, Sub-microscopic *Plasmodium vivax* Infection and Disappearing *Plasmodium falciparum* Malaria in an Area of Low Transmission in Solomon Islands. *PLoS Negl. Trop. Dis.* 9, 1–18.
 33. World Health Organization, 2015. World Malaria Report . Geneva.
 34. World Health Organization, 2019. World Malaria Report 2019. Geneva.
 35. Zhong, D., Lo, E., Wang, X., Yewhalaw, D., Zhou, G., Atieli, H.E., Githeko, A., Hemming-Schroeder, E., Lee, M.C., Afrane, Y., Yan, G., 2018a. Multiplicity and molecular epidemiology of *Plasmodium vivax* and *Plasmodium falciparum* infections in East Africa. *Malar. J.* 17, 1–14.
 36. Zhong, D., Koepfli, C., Cui, L., Yan, G., 2018b. Molecular approaches to determine the multiplicity of *Plasmodium* infections. *Malar. J.* 17, 1–9.

Annexes

Annexe 01 : Alignement des séquences : les séquences des amorces Forward et Reverse sont indiquées par la couleur jaune. Les SNPs sont indiqués en remplissage de couleur rouge pour la base « T », Bleu « C », Violet « G » et Vert « A ».

- Les séquences des primers forward et reverse
- * SNPs

Lysophospholipase Chr 02

DNA Sequences	Translated Protein Sequences
Species/Abbrv	
1. Chr02.resultat CCTAACGATTTTCATCCTTTGAATCTTCCATCATTTGTAAATTTCTCAATGCTAC*GTTCCCC*CCCA*GTGA*TACCCAAGAAATA*ATGGGAAGGT
2. chr 02.bnest	TATGCACCCCTTCAATTAATACTCCTAACGATTTTCATCCTTTGAATCTTCCATCATTTGTAAATTTCTCAATGCTAC*GTTCCCC*CCCA*GTGA*TACCCAAGAAATA*ATGGGAAGGT

Amorce Forward

DNA Sequences	Translated Protein Sequences
Species/Abbrv	
1. Chr02.resultat	AATTCTCAATGCTAC*GTTCCCC*CCCA*GTGA*TACCCAAGAAATA*ATGGGAAGGTTTTTGCATTTGGGGAG*TTTCCTTGTTAGGCAGCGTACCTCTTCTTTA*GCTGCGATGTGC
2. chr 02.bnest	AATTCTCAATGCTAC*GTTCCCC*CCCA*GTGA*TACCCAAGAAATA*ATGGGAAGGTTTTTGCATTTGGGGAG*TTTCCTTGTTAGGCAGCGTACCTCTTCTTTA*.....

Amorce reverse

AMA1 Chr 09

DNA Sequences	Translated Protein Sequences
Species/Abbrv	
1. Chr09.resultat TCATAGA-GAATTCCTGCTGTTAGCTTCTTAACCCCTGTGGCTACAGGAGATCAGAGGCTGAAGGATGGAAGTTTCGCCCTTCCCAA*GCGGATG
2. PVP01_09_v1 Plasmi	TCCCTGTTTTGGAAAGGGTATCGTCATAGA-GAATTCCTGCTGTTAGCTTCTTAACCCCTGTGGCTACAGGAGATCAGAGGCTGAAGGATGGAAGTTTCGCCCTTCCCAA*GCGGATG

DNA Sequences	Translated Protein Sequences
Species/Abbrv	
1. chr09 result	CCC A A T G C G G A T G A C C A T A T C T C C C C A T G A C A T T A G C G A A C C T T A A G G C A A G G T A T A A A G A C A A T G T A G A G A T G A T G C T G C G A T G T G C A A G T C T C A G A A G C A G A C C C A C G C A G C T A A
2. PIP01_09_v1 PIP	CCC A A T G C G G A T G A C C A T A T C T C C C C A T G A C A T T A G C G A A C C T T A A G G A A A G G T A T A A A G A C A A T G T A G A G A T G A T

MSP3 Chr 10

DNA Sequences	Translated Protein Sequences
Species/Abbrv	
1. chr10 08 T A A T A G C A C T T C A T T C C T A G G C G T T A T A T G C C A C A A T T G C T T A A A A C G A A T G A G A T A A C C T T T A C A G A T C G C T A A A T C C T T C G A T G T G T C C T C T
2. chr10 07 T A A G T A G C A C T T C A T T C C T A G G C G T T A T A T G C C A C A A T T G C T T A A A A C G A A T G A G A T A A C C T T T A C A G A T C G C T A A A T C C T T C G A T G T G T C C T C T
3. sequence nested 10	A A C A G A C T G A A C G C A A A A A C A G A T A G C A C T T C A T T C C T A G - C G T T A T A T G C C A C A A T T G C T T A A A A C G A A T G A G A T A A C C T T T A C A G A T C G C T A A A T C C T T C G A T G T G T C C T C T

DNA Sequences	Translated Protein Sequences
Species/Abbrv	
1. chr10 08	G G C G T T A T A T G C C A C A A T T G C T T A A A A C G A A T G A G A T A A C C T T T A C A G A T C G C T A A A T C C T T C G A T G T G T C C T C T T C C C A C T - G T T T G T G C T G C G A T - G T G C A A G T C T C A G A A
2. chr10 07	G G C G T T A T A T G C C A C A A T T G C T T A A A A C G A A T G A G A T A A C C T T T A C A G A T C G C T A A A T C C T T C G A T G T G T C C T C T T C C C A C T - G T T T G T G C T G C G A T - G T G C A A G T C T C A G A A
3. sequence nested 10	G - C G T T A T A T G C C A C A A T T G C T T A A A A C G A A T G A G A T A C A C C T T T A C A G A T C G C T A A A T C C T T C G A T G T G T C C T C T T C C C A C T - G T T T G T G

GILP Chr 11

DNA Sequences	Translated Protein Sequences
Species/Abbrv	
1. chr11 result	AAATCAATGGATGGCCGC--GAGGTGTGGCGGGCC*CCCGCTGAATGG-AAGCGGCTTTGCTCGTCCACTTTGCTGCACCATCGCTGCATTGTGGCTGCCATTATCGCCG
2. PVP01_11_v1 Plasm	CTGTAGGAGAGGTATCTTCGCCAAGAGGTGTGGCGGGCC*CCCGCTGAATGG-AAGCGGCTTTGCTCGTCCACTTTGCTGCACCATCGCTGCATTGTGGCTGCCATTATCGCCG

DNA Sequences	Translated Protein Sequences
Species/Abbrv	
1. chr11 result	CTTTGCTGCACCATCGCTGCATTGTGGCTGCCATTATCGCCGCCCTTTCCACCATTAT*GAACCGCCGAAACAGCCCGCAAACCCTTTTCCCCACCTGCTGCGATGTGCAAGTCTC
2. PVP01_11_v1 Plasm	CTTTGCTGCACCATCGCTGCATTGTGGCTGCCATTATCGCCGCCCTTTCCACCATTATGAACCGCCGAAACAGCCCGCAAACCCTTTTCCCCACCT-----

