# CROSS-SECTIONAL GENETIC ANALYSIS OF *Plasmodium falciparum* Rh5 INTERACTING PROTEIN (PfRipr) AND CYSTEINE-RICH PROTECTIVE ANTIGEN (CyRPA) GENES IN KILIFI, KENYA

**By**

**Kevin Ochwedo Omondi**

**I56/82769/2015**

**BSc. Microbiology and Biotechnology (University of Nairobi)**

**A thesis submitted in partial fulfillment of the requirements for the award of degree of Master of Science in Biotechnology, University of Nairobi**

## DECLARATION

I declare that this research is my own work and has not been submitted for examination in any other university.

**Kevin Omondi Ochwedo**

Registration Number: I56/82769/2015

Signature……..... Date……..

## SUPERVISORS' APPROVAL

We confirm that this thesis has been submitted with our approval as university supervisors;

**Dr. Isabella Oyier,**

Kenya Medical Research Institute (KEMRI), Wellcome Trust,

Centre for Geographic Medicine Research Coast,

Kilifi.

Signature…….....……. Date…….....…….

**Dr. George Obiero,**

Centre for Biotechnology and Bioinformatics (CEBIB),

University of Nairobi.

Signature……. Date……..

**ACKNOWLEDGEMENTS**

I express my gratitude to my supervisors for their enduring support, mentorship and professional guidance. Working under your supervision was a privileged and has molded me to be what I am today. I sincerely thank you Dr. Isabella Oyier for identifying my interest in malaria research, subsequent support and supervision that has granted me an opportunity to add value in this field. The continuous support and supervision role of Dr. George Obiero was magnificent, through his constructive comments, I found the courage and knowledge to counter the challenges that I faced throughout this study.

I extend my gratitude to CEBIB staffs for the technical support they offered me. I appreciate Mrs. Anne Owiti for her exemplary support throughout my study. This also extends to Mr. Edwin Rono, Enock and Collins. Special thanks go to my fellow students, Kevin Wamae, Abneel Kumar, Leonard Marigi, Sammuel Mwafurilwa and Martin Inyimili for their encouragement.

I acknowledge KEMRI-Wellcome Trust Collaborative Research Programme, Centre for Medicine Research-Coast, Kilifi for supporting this research through Dr. Oyier and thanks to Victor Osoti for assistance in sequencing. Finally, I thank my family for the support throughout my years of study, research and writing this thesis.

**DEDICATION**

I dedicate this Thesis to my late parents, Charles Ochwedo and Naomi Akeyo Ochwedo.
Your sacrifice towards my education was never in vain.

**TABLE OF CONTENTS**

**LIST OF FIGURES**

**LIST OF TABLES**

**ABBREVIATIONS**

**ACTs**….....Artemisinin-based combination therapies

**AMA1**….....Apical membrane antigen 1

**Basigin**…….....Basic immunoglobulin

**$Ca^{2+}$**…….……Calcium ions

**CyRPA**……..……Cysteine-rich protective antigen

**DNA**…….……Deoxyribonucleic acid

**EBLs**……Erythrocyte binding ligands

**EBAs**……..Erythrocyte binding antigens

**EDTA**.....Ethylenediaminetetraacetic acid

**EGF**….....Epidermal growth factor

**FMP1/AS02**……...…Falciparum Merozoite Protein-1/a proprietary Adjuvant System

**GIA**……Growth inhibition assay

**GLURP**….....Glutamate-rich *Plasmodium falciparum* antigen

**GPI**……....Glycophosphatidylinositol

**IMC**….....Inner membrane complexes

**ITNs**.....Insecticide-treated bednets

**KDa**……...Kilo Dalton

**MSP**……Merozoite surface protein

**MSPs**…….Merozoite surface proteins

**Nm**……..Nanometre

***P. falciparum***…….……*Plasmodium falciparum*

**PfEBLs**……...…….*Plasmodium falciparum* Erythrocyte binding ligands

**PfRhs**….....*Plasmodium falciparum* Reticulocyte binding-protein homologues

**PTRAMP**……...……Thrombospondin-related apical membrane protein

**PV**……....……Parasitophorous vacuole

**RALP1**…..... Rhoptry-associated,Leucine zipper-like protein-1

**RBCs**……...Red blood cells

**Rhs**……Reticulocyte binding-protein Homologues

**Ripr**……...……..Rh5 interacting protein

**ROM**……...……Rhomboid proteases

**RON**……...……..Rhoptry neck proteins

**SERA**……..Serine repeat antigen

**SNPs**……...……Single nucleotide polymorphisms

**SUB1**…..Subtilisin protease 1

**SUB2**…..Subtilisin protease 2

**µm**……Micrometre

**LD**……...…….……linkage disequilibrium

**ABSTRACT**

The discovery of antibodies against the Reticulocyte binding-protein homologue (Rh) 5 interacting protein (Ripr) and Cysteine-rich protective antigen (CyRPA), which are crucial in the invasion process across all parasite strains has brought new hope to the vaccine development field. Determining whether the Ripr and CyRPA genes in *P. falciparum* isolates from malaria-endemic population in Kilifi-Kenya are polymorphic provide data that is important in vaccine development targeting the two antigens. These results aid in preventing the development of a malaria blood-stage vaccine that would not proceed beyond the clinical trial stages due to the presence of multiple antigen variants. The genomic DNA of *P. falciparum* extracted from blood samples collected in 2013 and 2014 from 162 children aged below 8 years were used in the study. These children suffered from uncomplicated malaria and had been admitted to Kilifi County Hospital. From the extracted genomic DNA, Exon 1 and Exon 2 of CyRPA gene were separately amplified by different primer sets whereas Ripr gene was amplified using two different primer sets. Good quality amplicons were sequenced and analysed using CLC Genomics Workbench 7, MEGA 6.0 and DnaSP 5.10.01 software. Sequence assembly was done using CLC bio and subsequent analysis conducted using DnaSp software and MEGA 6.0. A total of three mutations were detected in sequences of exon 2 of CyRPA gene at positions 193, 1005 and 1086. SNPs at position 193 and 1005 resulted in non-synonymous mutations, whereas position 1086 was a synonymous mutation. The identified SNPs were under purifying selection, suggesting a possible stabilization of the CyRPA gene. The parasites tend to ensure that mutations that may interfere with the CyRPA antigen functionality are eliminated. Such a result reaffirms CyRPA antigen as a possible candidate in developing a blood-stage malaria vaccine in the future. Similar to Exon1 of CyRPA gene, the Ripr gene lacked polymorphisms. The result was obtained from analysis of 39 samples which accounted for 24.1% of the total samples analysed. The lack of polymorphism in Ripr, Exon1 of CyRPA sequences and polymorphisms in Exon 2 of CyRPA show that both these genes appear to be highly conserved with hardly any polymorphisms making them good vaccine candidates with no possible limitation of allele-specific immunity

# CHAPTER ONE

## 1.1 INTRODUCTION

*Plasmodium falciparum* is the most virulent malaria parasite among the five human *Plasmodium* species. Each year, it accounts for millions of deaths globally with a high mortality in children below 5 years and pregnant women (Elliott and Beeson, 2008). 85-90% of the fatalities have been reported in sub-Saharan Africa especially in Kenya (Geels *et al.*, 2011). By 2015, over 200 million new malaria cases had been recorded in more than 90 countries globally (World Health Organization, 2015). Eighty percent of death as a result of malaria is concentrated in 15 countries which are mainly within Africa. Among the most susceptible individuals are children under the age of 5 years out of whom, 292,000 have been estimated to have succumbed to malaria infection in 2015 within Africa (WHO, 2015). In Kenya, over 20 million individuals are estimated to be at risk of malaria infection while approximately 6 million new clinical cases were noticed each year (Halliday *et al.,* 2014). High transmission rates are recorded in endemic regions such as the western part of Kenya and Kilifi (Mogeni *et al.,* 2016).

Despite the existence of various vector control measures such as the use of insecticide sprays, repellents and Insecticide-treated bednets (ITNs), malaria prevalence has gradually increased across Africa. A similar observation is noted in Kilifi region (von Seidlein and Knudsen, 2016). The recent increase in *P. falciparum* transmission rates may be attributed to parasite strains that are resistant to antimalarials such as quinine and artemisinin-based combination therapies (Fairhurst and Dondorp, 2016). The clinical manifestation of malaria occurs due to the subsequent invasion of host erythrocytes by invasive *P. falciparum* merozoites (Miller *et al.,* 2002). The persistence of the malaria infection in the host is maintained through the invasion of fresh erythrocytes every 48 hours (Chen *et al.,* 2000).

Invasion of host erythrocytes by *P. falciparum* merozoites is a complex process involving the expression of various antigenic proteins (Gilson and Crabb, 2009). Key *P. falciparum* antigens that ensure invasion process is successful are those aiding the formation of the tight junction which initiates the process of invagination. Cysteine-rich protective antigen (CyRPA) and Reticulate homolog 5 interacting protein (Ripr) are among the proteins that form a complex with other antigens facilitating invagination of erythrocyte by merozites.

Based on the crucial role displayed by CyRPA and Ripr, the two antigens are under consideration as potential candidates for a target in the development of malaria bloodstage vaccines.

### 1.2 Problem statement and Justification

The extensive polymorphisms in immunodominant *P. falciparum* antigens present challenges in the development of effective blood-stage malaria vaccines (Takala and Plowe, 2009; Dzikowski and Deitsch, 2009). The Ripr and CyRPA antigens have been found capable of inducing growth inhibition assay (GIA)-active antibodies singly and in combination with the Rh5 antigen (Chen *et al.,* 2011; Reddy *et al.,* 2015). This qualifies them as potential blood stage malaria vaccines; however extensive polymorphisms may have a significant negative effect on the efficacy of the vaccines. No study has been conducted so far to determine if there are polymorphisms in Ripr and CyRPA genes obtained from *P. falciparum* isolates from a malaria endemic population in Kilifi, Kenya.

This research aimed to inform if a blood stage malarial vaccine based on the Ripr and CyRPA antigens are under selection pressure or may be rendered ineffective due to the polymorphisms. The data supports the research effort towards developing a *P. falciparum* malaria vaccine.

### 1.3 RESEARCH QUESTION

Are there polymorphisms in Ripr and CyRPA genes obtained from *P. falciparum* isolates from a malaria endemic population in Kilifi, Kenya?

### 1.4 HYPOTHESIS

There are no polymorphisms in the *P. falciparum* Ripr and CyRPA genes.

### 1.5 Objectives

#### 1.5.1 Main Objective

To determine polymorphisms in Ripr and CyRPA genes in *P. falciparum* isolates obtained in a malaria endemic population in Kilifi, Kenya.

#### 1.5.2 Specific Objectives

1. To genotype the *P. falciparum* Ripr and CyRPA genes.

2. To determine if polymorphisms are present in these genes.
3. To identify evidence of selection on Ripr and CyRPA genes if polymorphisms are present.

# CHAPTER TWO

# LITERATURE REVIEW

## 2.1 The life cycle of *P. falciparum*

*P. falciparum* undergoes both sexual and asexual stage in its life cycle. The sexual phase occurs in the mosquito (vector), while the asexual phase that is responsible for the clinical manifestations of the disease occurs in humans, the definitive host (Miller *et al.,* 2002). The complex parasite life cycle begins when the *P. falciparum* sporozoites are injected into the human peripheral circulation, through a blood meal by the female *Anopheles* mosquito as illustrated in Fig 1. The injected sporozoites then migrate through the blood stream to the liver and invade hepatocytes, where they undergo hepatic schizogony and differentiate into thousands of schizonts (Fig 1). After 7 days, the schizonts rupture releasing merozoites that egress from the liver and enter the blood stream. Within the blood stream, they invade erythrocytes and reside in a parasitophorous vacuole (PV) as ring stage parasites (Das *et al.,* 2015). The ring stage parasites in the PV transform into early trophozoites then to mature trophozoites, before undergoing erythrocytic schizogony that results in multinucleated schizonts (Beeson *et al.,* 2016).

These multinucleated schizonts develop into late schizonts whose PV membrane rapture and releases merozoites into the erythrocyte cytosol. A few seconds following the release of merozoites, the erythrocyte membrane raptures releasing new merozoites into the bloodstream which invade other uninfected erythrocytes (Josling and Llinás, 2015). Some intra-erythrocytic stage merozoites leave the cycle and develop into the sexual stages. They differentiate into male and female gametes that are taken up by mosquitoes during a blood meal and marking the end of asexual life cycle (Fig 1). On reaching the mosquito midgut, the sexual life cycle begins with the male gametes fertilizing the female gametes, resulting in diploid zygotes that develop into infective elongated motile ookinetes. The ookinetes migrate via the wall of the mosquito midgut and develop into extracellular oocysts. The developed oocyst releases sporozoites into the salivary glands of the vector. These released sporozoites are ready for transmission to the next host during a mosquito bloodmeal (Josling and Llinás, 2015). For a successful asexual cycle, *P. falciparum* has developed a structure that facilitates a series of molecular interactions that result in the infection of erythrocytes.

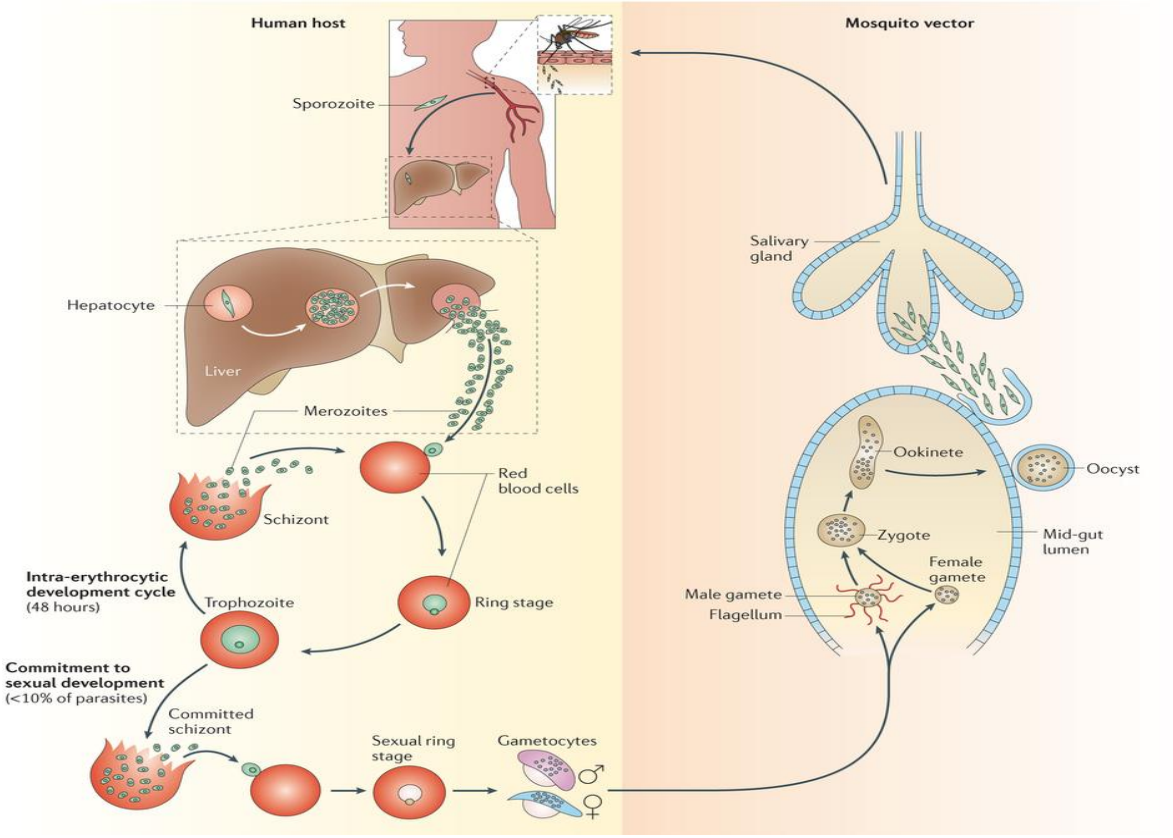



**Figure 1:** The *Plasmodium falciparum* life cycle. The figure shows both the asexual and sexual lifecycle of *P. falciparum* in the human host and mosquito respectively. Adapted from Josling and Llinás, (2015).

## 2.2 The structure of the merozoite

Erythrocyte invasion by *P. falciparum* merozoites initiates the asexual blood stage process which is crucial for parasite development. The *P. falciparum* merozoites have undergone evolution and developed organelles needed for rapid and efficient invasion (Wright *et al.*, 2014; Garcia *et al.*, 2008). The merozoite cell is ovoid with a length of ~1.2μm and has a polarised structural morphology containing both an apical prominence and a posterior end (Wright *et al.*, 2014). The cell surface is covered by a 15 nm thick adhesive coat, which has bristles (Fig 2) that initiate contact with the erythrocyte during invasion. The plasma membrane and the underlying inner membrane complex (IMC) forms a merozoite pellicle that covers the entire cell surface except for the apical end.

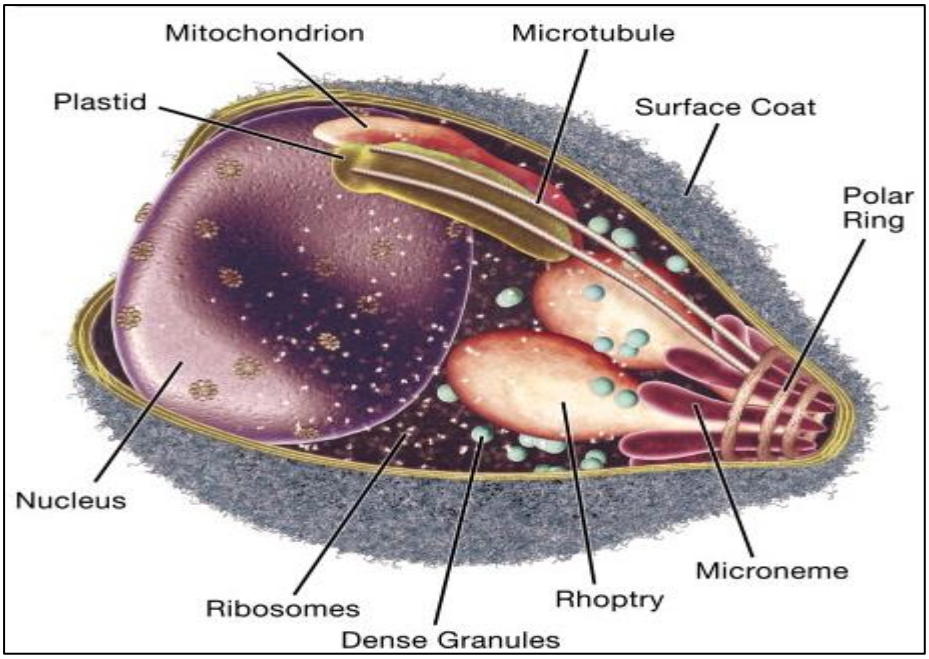



**Figure 2:** The cellular structure of *P. falciparum* merozoite cell**.** The figure shows major organelles and cellular structures of the *Plasmodium falciparum* merozoite. Adapted from Cowman and Crabb, (2006).

The IMC also anchors the actomyosin motor through proteins; the motor has a role in junction formation during invasion (Wright *et al.,* 2014). Microtubules target the apical organelles during the assembly of merozoites and also provide mechanical support (Garcia *et al.,* 2008). The merozoite cytoskeleton has three polar rings (Fig 2) that provide additional mechanical support to the rhoptry and the micronemes. The nuclear, mitochondria and plastid (apicoplast) are located at the posterior end of the merozoite cell where they control genetic and metabolic processes. Other organelles such as the Golgi complex and endoplasmic reticulum are either residual or absent in a mature merozoite cell (Fig 2). The microneme vesicles are densely stained and elongated as seen in Fig 2; they sequentially release different proteins at distinct phases of invasion (Garcia *et al.,* 2008).

The apical half of the merozoite cell contains scattered dense granule vesicles (Fig 2), which discharge their contents that decorate the PV membrane. The mononeme and exoneme which are not indicated in Fig 2 are thought to contain proteases that are crucial for invasion (Riglar *et al.,* 2011; Garcia *et al.,* 2008). The merozoite prominence centre contains two pear-shaped rhoptries, which are subdivided into head and bulb functional domains and extracellularly secrete distinct proteins at different times of invasion (Zuccala *et al.,* 2012; Hanssen *et al.,* 2013).

## 2.3 Erythrocyte invasion

The invasion of erythrocytes by *P. falciparum* merozoites involves a series of complex molecular interactions that takes an average time of less than two minutes to be completed (Gilson and Crabb, 2009).

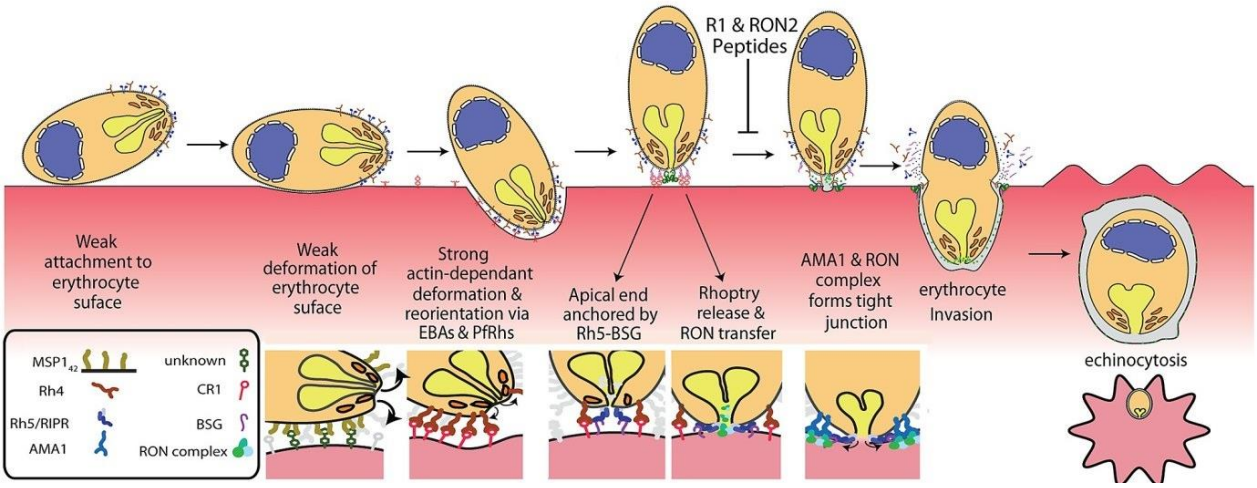



**Figure 3:** The *P. falciparum* merozoite invasion process**.** The model illustrates the initial stage of merozoite attachment on erythrocyte surface, deformation of erythrocyte surface, reorientation, anchorage, the formation of a tight junction, invasion and internalization of the merozoite into the erythrocytes. Merozoite and host proteins are involved. Adapted from Weiss *et al.* (2015).

The merozoite initiates invasion through the exoneme released calcium-dependent redox switch subtilisin protease (SUB1), which produces the active cysteine protease forms from the processing of the serine repeat antigen 6 (SERA6) and other SERA family proteins (Withers-Martinez *et al.*, 2014; Beeson *et al.*, 2016). It also processes merozoite surface protein 1 (MSP1), MSP6, and MSP7 leading to their maturation and gain of invasive functions (Das *et al.*, 2015; Child *et al.*, 2010). The different forms of protease and processed MSP1 that interact with spectrin are thought to mediate the egress of merozoites from schizonts in the PV and the erythrocyte cytosol (Chen *et al.*, 2000; Ruecker *et al.*, 2012; Favuzza *et al.*, 2016). The free merozoite in the blood establishes an initial contact with the erythrocyte through their MSPs, SERA4, SERA5 and the 42 kilo Dalton of MSP1 ( $MSP1_{42}$), binding to the surface receptor band 3, glycophorin A and possibly heparin-like molecules on the erythrocyte surface (Boyle *et al.*, 2010; Wright *et al.*, 2014; Baldwin *et al.*, 2015; Weiss *et al.*, 2015).

Merozoite surface protein 1 (MSP1) aggregates and causes a weak deformation on the erythrocyte through its membrane fluid movement activity (Weiss *et al.*, 2015). This is

followed by a strong erythrocyte surface deformation that is likely caused by the interaction between the Erythrocyte binding antigen (EBA) family including EBA175, the *P. falciparum* Reticulocyte binding-protein homologue (PfRh) family (Rh4) and the action of the acto-myosin motor (Fig 3). The Rh1 contained in the rhoptry duct interacts with the erythrocyte surface and triggers intracellular calcium ion ($Ca^{2+}$) signaling that results in the release of EBA175 by the micronemes (Gao *et al.,* 2013; Wright *et al.,* 2014; Weiss *et al.,* 2015). The release of EBAs, Rhs and their interactions with the acto-myosin motor causes the reorientation of the merozoite (Weiss *et al.,* 2015).

The reoriented merozoite through mediation by the microneme secreted apical membrane antigen 1 (AMA1), Rh1, Rh2a, Rh2b, Rh5, Rhoptry-associated, leucine zipper-like protein-1 (RALP1) and the rhoptry neck proteins (RON2) forms a tight junction (Haase *et al.,* 2008; Riglar *et al.,* 2011; Riglar *et al.,* 2016).The junction is hypothesised to act as a point of traction for leveraging action of the acto-myosin motor (Riglar *et al.,* 2016; Angrisano *et al.,* 2012), which propels the merozoite into the erythrocyte and causes the tight junction to move from the apical end to the posterior end of the merozoite (Fig 3). The microneme released SUB2 cleaves the surface antigens such as Thrombospondin-related apical membrane protein (PTRAMP), MSP1, 33 kilo Dalton of MSP1( $MSP1_{33}$), MSP3, MSP7 and SERA4 and sheds them off (Fig 3) together with the tight junction before complete internalization (Boyle *et al.,* 2010; Riglar *et al.,* 2011; Withers-Martinez *et al.,* 2012; Boyle *et al* 2014). The MSP4 and $MSP1_{19,}$ which is processed from the $MSP1_{42}$ by the SUB2 are left on the merozoites surface and are echinocytosed (Boyle *et al.,* 2014).

The Rhomboid proteases ROM1 and ROM4, cleave and shed off the EBA family proteins (such as EBA175), Rhs, and AMA1 and facilitate the resealing of the erythrocyte membrane at the entrance point (Santos *et al.,* 2012). Upon completion of internalisation, MSP2 which was internalised while on the merozoite surface is degraded within 10 minutes while the 19 kilo Dalton of MSP1 ($MSP1_{19}$) and MSP4 are retained (Boyle *et al.,* 2014). The retained $MSP1_{19}$ is hypothesized to help in food vacuole formation during the post invasion process (Moss *et al.,* 2012). Among the most important proteins during the invasion process, are the Reticulocyte-binding protein homologue 5 (Rh5), Rh5-interacting protein (Ripr), Cysteine-rich protective antigen (CyRPA), and surface protein (P113) which forms a complex that is indispensable during the invasion process.

## 2.4 Rh5/Ripr/CyRPA/P113 complex is indispensable in all *P. falciparum* merozoite invasion process

Reticulocyte-binding protein homologue 5 (Rh5) undergoes processing which results in fragments that interact directly with P113 and CyRPA to form Rh5/ CyRPA /P113/Basigin multiprotein complex, which mediates junction formation (Chen *et al.,* 2011; Reddy *et al.,* 2015). The recruitment of Ripr to the complex following binding of CyRPA to Rh5 is believed to result in the termination or cleavage of CyRPA/Rh5/Ripr complex from the initial Rh5/ CyRPA / P113/Basigin complex (Galaway *et al.,* 2017). Within the complex, the Rh5 N terminus (Rh5Nt) interacts with P113 providing anchorage on the merozoite surface whereas the C-terminus binds to the erythrocyte surface protein basigin or CD147 in the final stages of invasion in all strains (Wright *et al.,* 2014; Galaway *et al.,* 2017). Rh5 is refractory to genetic deletion and it is smaller (~ 60 kDa) protein compared to other Rhs (Hayton *et al.,* 2013; Baum *et al.,* 2009). Only 12 non-synonymous single nucleotide polymorphisms (SNPs) have been found so far in the Rh5 sequence of 227 field isolates (Wright *et al.,* 2014).

The monoclonal and polyclonal antibodies generated against Rh5 have been found to prevent merozoite development in all strains by blocking the interaction of Rh5 with its host receptor basigin *in vitro* (Douglas *et al.,* 2011; Bustamante *et al.,* 2013; Douglas *et al.,* 2014; Reddy *et al.,* 2014; Wright *et al.,* 2014; Galaway *et al.,* 2017). These anti-Rh5 antibodies have been found in relatively low prevalence in humans and are associated with a role in host protection from malaria episodes in Mali (Patel *et al.,* 2013; Tran *et al.,* 2014). Rh5 has therefore been considered as a next generation blood-stage malaria vaccine candidate even though it has been reported to have low immunogenicity.

Key among the proteins that interacts with Rh5 is P113, which is glycosylphosphatidylinositol (GPI)-linked (Galaway *et al.,* 2017). The P113 interacts with the N-terminus of unprocessed Rh5 (comprising of linear sequence of 19 amino acids) hence providing the mechanism upon which the Rh5 complex is tethered on the merozoite surface. The complex formed as a result of interaction between Rh5/P113/Basigin is believed to trigger the secretion of parasitic ligands such as AMA1/RON complex leading to the formation of moving junction. The P113 is expressed in early and late schizont stages and has also been detected in the gametocyte as well as sporozoite stages within the life cycle of *P. falciparum* (Silvestrini *et al.,* 2010). The inability of selection of merozoites that are deficient of P113 has confirmed its indispensability in merozoite growth (Galaway *et al.,* 2017).

Antibodies against P113 in African and Papua New Guinea population have largely been associated with offering protection against malaria (Osier *et al.,* 2014; Richards *et al.,* 2013).

Cysteine-rich protective antigen (CyRPA) is released by the micronemes and is usually expressed as a ~35 kDa protein and plays a role in recruiting Ripr to the complex (Galaway *et al.,* 2017; Reddy *et al.,* 2015; Favuzza *et al.,* 2016). The CyRPA binds to Rh5 in the Rh5/P113/Basigin complex prior to recruitment of Rh5-interacting protein (Ripr) to the complex. The *Cyrpa* gene (PF3D7_0423800) cannot be knocked out and it is located in the subtelomeric region of chromosome 4 (Baum *et al.,* 2009; Richards *et al.,* 2013; Douglas *et al.,* 2014). Six *P. falciparum* field isolates from Tanzania did not have any polymorphisms in the CyRPA gene, however a nonsynonymous SNP has been detected at 1116bp across the K1 (Thailand), FCR3 (Gambia), ITG2F6 (Brazil), and FVO (Vietnam) strains (Dreyer *et al.,* 2012). Antibodies that target full-length CyRPA have been reported to abrogate its interaction with the Rh5/Ripr complex and blocked the multiprotein complex formation leading to inhibition of invasion (Reddy *et al.,* 2015). It has been suggested that CyRPA is not a major target of naturally acquired immunity since its natural immunogenicity is very low (Dreyer *et al.,* 2012).

The microneme released Ripr has a molecular weight of 123 kDa (Chen *et al.,* 2011). It lacks a transmembrane domain and is found localised at the apical end of the merozoite, where it peripherally interacts with the merozoite surface and Rh5. Prior to the invasion, Ripr undergoes processing into two 65 kDa fragments which are almost similar and are purported to act as a scaffold anchoring Rh5 on Rh5/Ripr/CyRPA multiprotein complex (Volz *et al.,* 2016; Chen *et al.,* 2011). The polyclonal antibodies that have been developed against the Ripr antigen inhibit erythrocyte invasion and the gene encoding it is considered refractory to genetic disruption (Chen *et al.,* 2011). Surface protein (P113), Rh5, Ripr and CyRPA have therefore been considered as the next generation blood stage malaria vaccine candidates.

### 2.5 Blood stage malaria vaccine development and challenges

The malaria blood stage vaccines based on MSP1, MSP2, MSP3, EBA175, glutamate rich *P. falciparum* antigen (GLURP) and SERA5 have undergone human vaccine clinical trials and the presence of multiple antigen variants have proven to be one of the greatest challenges, as most of these vaccines are allele specific (Sirima *et al.,* 2011; Ogwang *et al.,* 2013; Tamminga *et al.,* 2013). Vaccines such as the $MSP1_{42}$-3D7 (FMP1/AS02) and the AMA1-3D7 (FMP2.1/ASO2A) had no significant effect on clinical malaria and were allele-specific

during the phase II clinical trial in Kenyan and Mali children (Ogutu *et al.*, 2009; Thera *et al.*, 2011). The vaccines were based on alleles from the 3D7 laboratory generated isolates and thus, the vaccines protected children in the trial with *P. falciparum* alleles expressing the 3D7 target antigen.

Focus has shifted to the Rh5/Ripr/CyRPA/P113 complex that is important in *P. falciparum* merozoite invasion. However, Reddy *et al* (2015) showed that antibodies raised against CyRPA partial constructs had poor invasion inhibitory activity and they could not immunoprecipitate the *P. falciparum* Rh5/Ripr/CyRPA complex. Extensive polymorphisms in the CyRPA and Ripr genes could make a vaccine based on these antigens to be allele specific, reducing their efficacy.

# CHAPTER THREE

# MATERIALS AND METHODS

## 3.1 Study site and population

One hundred and sixty two blood samples from children below 8 years with uncomplicated malaria admitted to Kilifi County Hospital were used. The blood sample were collected in the year 2013 and 2014. The children had variable parasitemia with a median of 190,000 parasites/µl. The study was carried out at the University of Nairobi's Centre for Biotechnology and Bioinformatics (CEBIB) molecular laboratories.

## 3.2 DNA extraction

The DNA from 162 blood samples were previously extracted using the QIAamp DNA extraction Blood mini kit (QIAGEN, Inc.). The kit allows the generation of DNA that is ready for use following a series of steps where 20µl of proteinase K was added into a microcentrifuge tube before the addition of 200µl of the anticoagulated blood sample. It was followed by the addition of 200µl of buffer AL; the sample was vortexed to mix. Incubation of the mixture was done at 56°C for 10 minutes and 200µl of ethanol added before vortexing. It was then transferred to a Mini spin column and centrifuged at 6000 x g for 1 minute, the flow-through and collection tube was discarded. The spin column was then placed in a new 2ml collection tube before addition of 500µl of buffer AW1 and centrifuged for 1 minute at 6000 x g. The flow-through and collection tube were discarded before placing the spin column in another 2ml collection tube. Five hundred (500µl) of buffer AW2 was added and the mixture centrifuged for 3 minutes at 20,000 x g to dry the membrane. The flow-through together with the collection tubes were discarded and the spin column placed in a clean 2ml microcentrifuge tube. Two hundred (200µl) of buffer AE was then pipetted directly onto the membrane followed by incubation at room temperature for 1 minute before centrifuging for 1 minute at 6000 x g to elute. This step was repeated to obtain a maximum DNA yield. The eluted DNA was stored in a freezer at -20°C after making a 1:10 dilution.

### 3.3 Primer design and Re-suspension

Designed primers were used in the optimization process to select the right PCR conditions (Table 1).

**Table 1:** List of designed primers.

| Oligo name | Sequence |
|---|---|
| **Ripr_F1** | ATGTTCAGAATTTTTTTTACCCTTC |
| **Ripr_F2** | ACAATTGTAAATTCGATTTTG |
| **Ripr_F3** | ATGTGCTTATGGAAATACA |
| **Ripr_F4** | ACATGCTATGGGAACAG |
| **Ripr_F5** | ACAAGGTCATGTAGCTGTCA |
| **Ripr_F6** | ACAAATACGTTTATACTATATTAAATCGT |
| **Ripr_F7** | AAAAATGTTTGTGAAGAAAATTATAGATGTAC |
| **Ripr_R1** | CTAATTCTGATTACTATAATAAAATAC |
| **Ripr_R2** | TAGCACATATACACACACCATTCT |
| **Ripr_R3** | TGATGGTATATTAGAATGA |
| **Ripr_R4** | TGCATGTTGCTTTATGTGTATCA |
| **CyprA_F1** | TGAAAATAACATTATGATTATCCCT |
| **CyprA_F2** | CAGTGAGATAACTATAAGT |
| **CyprA_F3** | TGAGTGTACCCATGAAAAGGA |
| **CyprA_R1** | TCCTTGCAGTAACCCCTTTTGTCTAC |
| **CyprA_R2** | TGTCATCCTTCTTATTGTCATCCT |

Before optimization, the designed Ripr and CyRPA primers were re-suspended by addition of PCR water on each primer to make 100µM of stock solutions as per oligonucleotide synthesis report (Eurofins Genomics). Each primer was vortexed for 30 sec and then centrifuged at 14000 rpm for 20 sec. The re-suspended primers were left overnight in the freezer (-20˚C) before 10µl of each were transferred to a sterile, labeled Eppendorf and 90µl of PCR water

were added to make a final volume of 100µl working solution. The primers were stored in the freezer (-20˚C) before PCR optimization.

### 3.4 Primer mapping PCR optimization

Each primer sequence was mapped onto the respective CyRPA (1188 bp) and Ripr (3260 bp) reference sequence (3D7). It was done to predetermine the length of expected PCR products. Both the elongation temperature and time to be used in the PCR process are shown in Table 3. Optimization of both Ripr and CyRPA primers pairs (Table 2) were performed starting with specific forward primers (F) and reverse primers (R). 0.5µl of 3D7 DNA was added to a mix containing; 0.2µl of dNTP mix (10mM), 0.3µl of the forward primers (10×), 0.3µl of the reverse primers (10×), 1µl of Expand High Fidelity Buffer (10×) with 15mM Magnesium chloride ($MgCl_2$), 1µl of 1 × $MgCl_2$ (1.5mM), 0.14µl of Expand ™ High Fidelity enzyme mix (Roche) and 6.56µl of double distilled water in a final volume of 10µl. The Veriti™ Thermal Cycler (Applied Biosystems™) was used in the optimization process, in which 6 different annealing temperatures were tested in all the cycles by a gradient PCR method. The technique allows for the determination of the optimum PCR condition using a minimum number of steps possible. Through gradient PCR, the optimum condition is obtainable in a single step since its gradient function allows for testing six different annealing temperatures at once. Optimization was done as per the conditions in Table 3.

**Table 2:** List of primer pairs, predetermined target length and their respective elongation temperature that were tested during the optimization process.

| **Forward primer** | **Reverse primer** | **Base pairs** | **Elongation temperature (℃)** | **Elongation time** |
|---|---|---|---|---|
| CyRPA_F1 | CyRPA R1 | 1180 | 72 | 1 min |
| CyRPA_F1 | CyRPA R2 | 586 | 72 | 45 sec |
| CyRPA_F2 | CyRPA R1 | 750 | 72 | 45 sec |
| CyRPA_F2 | CyRPA R2 | 154 | 72 | 45 sec |
| CyRPA_F3 | CyRPA R1 | 313 | 72 | 45 sec |
| Ripr_F1 | Ripr_R1 | 3261 | 68 | 2 min |
| Ripr_F1 | Ripr_R2 | 2266 | 72 | 1 min |
| Ripr_F1 | Ripr_R3 | 1485 | 72 | 1 min |
| Ripr_F1 | Ripr_R4 | 805 | 72 | 1 min |
| Ripr_F2 | Ripr_R1 | 2617 | 72 | 1 min |
| Ripr_F2 | Ripr_R2 | 1622 | 72 | 1 min |
| Ripr_F2 | Ripr_R3 | 841 | 72 | 1 min |
| Ripr_F2 | Ripr_R4 | 161 | 72 | 1 min |
| Ripr_F3 | Ripr_R1 | 1935 | 72 | 45 sec |
| Ripr_F3 | Ripr_R2 | 940 | 72 | 1 min |
| Ripr_F3 | Ripr_R3 | 159 | 72 | 45 sec |
| Ripr_F4 | Ripr_R1 | 1316 | 72 | 1 min |
| Ripr_F4 | Ripr_R2 | 321 | 72 | 45 sec |
| Ripr_F5 | Ripr_R1 | 603 | 72 | 45 sec |
| Ripr_F6 | Ripr_R1 | 3064 | 68 | 2 min |
| Ripr_F6 | Ripr_R2 | 2069 | 72 | 1 min |
| Ripr_F6 | Ripr_R3 | 1288 | 72 | 1 min |
| Ripr_F6 | Ripr_R4 | 608 | 72 | 45 sec |
| Ripr_F7 | Ripr_R1 | 2426 | 72 | 1 min |
| Ripr_F7 | Ripr_R2 | 1431 | 72 | 1 min |
| Ripr_F7 | Ripr_R3 | 650 | 72 | 45 min |

**Table 3:** PCR conditions for optimization process.

| | Temperature (℃) | Time | Cycle |
|---|---|---|---|
| **Initial Denaturation** | 94℃ | 2 min | 1× |
| **Denaturation** | 94℃ | 15 sec | |
| **Annealing** | 42℃ – 59℃ | 30 sec | 10× |
| **Elongation** | 68℃ or 72℃ | 45 sec, 1 min or 2min | |
| **Denaturation** | 94℃ | 15 sec | |
| **Annealing** | 42℃ – 59℃ | 30 sec | |
| **Elongation** | 68℃ or 72℃ | 45 sec, 1 min or 2min + 5 sec cycle elongation for each successive cycle | 25× |
| **Final Elongation** | 72℃ | 7 min | 1× |
| **Cooling** | 4℃ | 4 Min | |

### 3.5 Amplification of the CyRPA gene

Among the CyRPA primer sets, only the CyRPA_F1, CyRPA_R2 and CyRPA_ F3 and CyRPA_R1 pairs were optimized successfully yielding the expected PCR products. The PCR conditions for CyRPA F1/R2 that generated the first CyRPA gene fragments (586 bp) were as described in Table 4.

**Table 4:** PCR conditions for amplification of CyRPA F1/R2 fragment.

| | Temperature (℃) | Time | Cycle |
|---|---|---|---|
| **Initial Denaturation** | 94℃ | 2 min | 1× |
| **Denaturation** | 94℃ | 15 sec | |
| **Annealing** | 44℃ | 30 sec | 10× |
| **Elongation** | 72℃ | 45 sec | |
| **Denaturation** | 94℃ | 15 sec | |
| **Annealing** | 44℃ | 30 sec | |
| **Elongation** | 72℃ | 45 sec + 5 sec cycle elongation for each successive cycle | 25× |
| **Final Elongation** | 72℃ | 7 min | 1× |
| **Cooling** | 4℃ | 4 min | |

The PCR conditions for CyRPA F3/ R1 that yielded the second fragment of the CyRPA gene (313 bp) were as described in Table 5.

**Table 5:** PCR conditions for amplification of CyRPA F3/R1 fragment.

| | Temperature (℃) | Time | Cycle |
|---|---|---|---|
| **Initial Denaturation** | 94 ℃ | 2 min | 1× |
| **Denaturation** | 94 ℃ | 15 sec | |
| **Annealing** | 51 ℃ | 30 sec | 10× |
| **Elongation** | 72 ℃ | 45 sec | |
| **Denaturation** | 94 ℃ | 15 sec | |
| **Annealing** | 51 ℃ | 30 sec | |
| **Elongation** | 72 ℃ | 45 sec + 5 sec cycle elongation for each successive cycle | 25× |
| **Final Elongation** | 72 ℃ | 7 min | 1× |
| **Cooling** | 4 ℃ | 4 Min | |

### 3.6 Amplification of the Ripr gene

Among the Ripr primer sets, only the Ripr F3/R1 and Ripr F6/R1 sets were optimized successfully yielding the expected PCR products. The obtained fragments did not cover the entire length of the gene as expected. The PCR conditions for Ripr F6/R1 that generated the first Ripr gene fragments (3064 bp) were as described in Table 6.

**Table 6:** PCR conditions for amplification of Ripr F6/R1 fragment.

| | Temperature (℃) | Time | Cycle |
|---|---|---|---|
| **Initial Denaturation** | 94 ℃ | 2 min | 1× |
| **Denaturation** | 94 ℃ | 15 sec | |
| **Annealing** | 49 ℃ | 30 sec | 10× |
| **Elongation** | 68 ℃ | 2 min | |
| **Denaturation** | 94 ℃ | 15 sec | |
| **Annealing** | 49 ℃ | 30 sec | |
| **Elongation** | 68 ℃ | 2 min + 5 sec cycle elongation for each successive cycle | 25× |
| **Final Elongation** | 72 ℃ | 7 min | 1× |
| **Cooling** | 4 ℃ | 4 min | |

The PCR conditions for Ripr F3/R1 that yielded the second fragment of the Ripr gene (1935 bp) were set as in Table 7 below.

**Table 7:** PCR conditions for amplification of Ripr F3/R1 fragment.

| | Temperature (℃) | Time | Cycle |
|---|---|---|---|
| **Initial Denaturation** | 94℃ | 2 min | 1× |
| **Denaturation** | 94℃ | 15 sec | |
| **Annealing** | 55℃ | 30 sec | 10× |
| **Elongation** | 72℃ | 2 min | |
| **Denaturation** | 94℃ | 15 sec | |
| **Annealing** | 55℃ | 30 sec | |
| **Elongation** | 72℃ | 2 min + 5 sec cycle elongation for each successive cycle | 25× |
| **Final Elongation** | 72℃ | 7 min | 1× |
| **Cooling** | 4℃ | 2 Min | |

### 3.7 PCR product detection by Agarose gel electrophoresis

After PCR, the amplicons were assessed to determine their quality through detection of specific bands upon electrophoresis in 1% w/v agarose gel. The gel used was prepared from previously made 1X Tris-Borate (TBE-Ethylenediaminetetraacetic acid (EDTA)) buffer at pH 8.0. The 10X TBE buffer was reconstituted as follows, 108g of Tris base, 55g of Boric acid and 75g of EDTA was dissolved in 800ml of double distilled water. It was topped up with distilled water to make 1L of 10X TBE buffer that was sterilized and stored at room temperature. 1X TBE buffer solution was regularly prepared by dissolving 50ml of 10X TBE buffer in 450 ml of double distilled water.

Half a gram (0.5g) of agar was dissolved in 50ml of 1X TBE buffer to make 1% w/v gel solution; this was boiled, cooled and 1µl of Ethidium bromide added before dispensing in the gel tank. The gel comb was placed in the gel tank to form the loading wells. It was then allowed to set for 25 minutes before addition of 1µl of loading dye 2µl of amplicons on the wells. 1µl of loading dye was added to 1µl of negative control (PCR water) before loading to the last well. A similar amount of loading dye was added to an equal amount of positive control which had the expected band length. Electrophoresis was done at a constant voltage of 80V for 40 minutes. The amplicons were then visualized under a UV transilluminator (BIORAD). Samples with specific, clear bands of required length as the positive control were considered positive while those with faint bands, non-specific bands, smears or no bands were deemed to be negative. The amplification process for negative samples was repeated by

increasing the sample DNA volume from 0.5µl to 0.7µl. In cases where negative results were obtained after repeats, the sample was declared negative.

### 3.8 Amplicon purification

All positive PCR products were purified using Exonuclease 1-Shrimp Alkaline Phosphatase (ExoSAP) clean-up before sequencing. 1µl of ExoSAP was loaded into each of the 96 well plate before addition of 8µl of the amplicons. The ExoSAP conditions were set as indicated in Table 8.

**Table 8:** ExoSAP conditions for cleaning the amplicons.

| Temperature (℃) | Time | Cycle |
|---|---|---|
| 37℃ | 15 sec | 1× |
| 80℃ | 15 sec | |

### 3.9 Sequencing

Sequencing was done using amplification with PCR and internal primers (Table 1) and the 3700/3730 BigDye® Terminator v3.1 Sequencing Standard kit (ABI PRISM® 3700 DNA Analyzer). Briefly, 4µl of purified PCR product was mixed with 0.3µl BigDye Terminator v3.1, 1.75µl 5X sequencing buffer, 0.32µl of each primer and 3.63µl of nuclease-free water. The sequencing reactions were performed in a thermal cycler under the following conditions; one cycle of polymerase activation at 96°C for 60 sec, 25 cycles of denaturation at 96°C for 30 sec, annealing at 50°C for 15 sec, extension at 60°C for 4 minutes and a final holding temperature of 15°C for 10 minutes. It was followed by ethanol and Sodium Acetate precipitation of 10µl sequencing reactions in 96-well reaction plates where 90µl of the ethanol and Sodium premix was added to each well. The plates were then sealed with micro-seals and incubated at -20°C for 30 minutes after which they were spun at 3000 × g for 30 minutes at 4°C on a centrifuge (5810R bench centrifuge, Eppendorf). Seals were removed and the plates overlaid with absorbent paper towels and gently inverted to drain them. Fresh paper towels were placed on the inverted plates and spun at 50 × g for 1 minute at 4°C, followed by addition of 150µl of ice cold (-20°C) 70% ethanol in each well before the plates were sealed and spun at 3000 × g for 10 minutes at 4°C. The plates were inverted again over paper towels, and excess fluid drained gently before being overlaid with clean paper towels, inverted and spun at 50 × g for 1 minute at 4°C. They were covered with fresh paper towels

and left on the bench to air dry. The samples were re-suspended in Hi-Di for denaturation (96°C for 1 minutes) reaction following the addition of 10µl of Hi-Di™ Formamide into each well and electrophoresis by ABI 3730 xl capillary sequencer (Applied Biosystems, UK).

### 3.10 Sequence editing, assembling, alignment, allele frequency calculations and estimation of genetic diversity

Sequences were assembled, edited and aligned using CLC Genomics Workbench 7 (CLC bio). The cleaned Ripr and CyRPA sequence were aligned separately to a reference sequence PF3D7_0323400 for Ripr and PF3D7_0423800 for CyRPA using the CLC Genomics Workbench 7 (CLC bio). Mega 6.0 and DnaSP v5.0 softwares were used to analyse the DNA polymorphisms, the number of synonymous and nonsynonymous substitutions computing nucleotide diversity and the Tajima's D statistic to determine if the detected polymorphisms were under selection.

### 3.11 Ethical clearance

Ethical clearance for the samples used in this study was obtained from the KEMRI Scientific and Ethics Review Unit (SERU protocol 3149).

# CHAPTER FOUR

# RESULTS

## 4.1 CyRPA amplification

Out of 162 isolated DNA samples, CyRPA_F1/R2 primer set was successfully used in amplifying 124 samples. From the same samples (162 samples), CyRPA_F3/R1 primer set amplified 132. The samples which failed to be amplified even after adjustments such as an increase in template DNA volume were made, were assumed to have degraded. The inability of the designed primer combinations (CyRPA_F1/R1, CyRPA_F1/R2, CyRPA_F2/R1, CyRPA_F2/R2, CyRPA_F3/R1, and CyRPA_F3/R2) to cover the entire length of targeted CyRPA gene (1180 bp) resulted in amplification of two fragments. CyRPA F1/R2 primer combination amplified a fragment of 586 bp whereas CyRPA F3/R1 primer combinations amplified 313 bp. The fragment sizes were determined using a 100 bp DNA ladder (Bioline, London, UK) as seen in Figure 4.1 and Figure 4.2. For quality assurance, the negative control had no band.

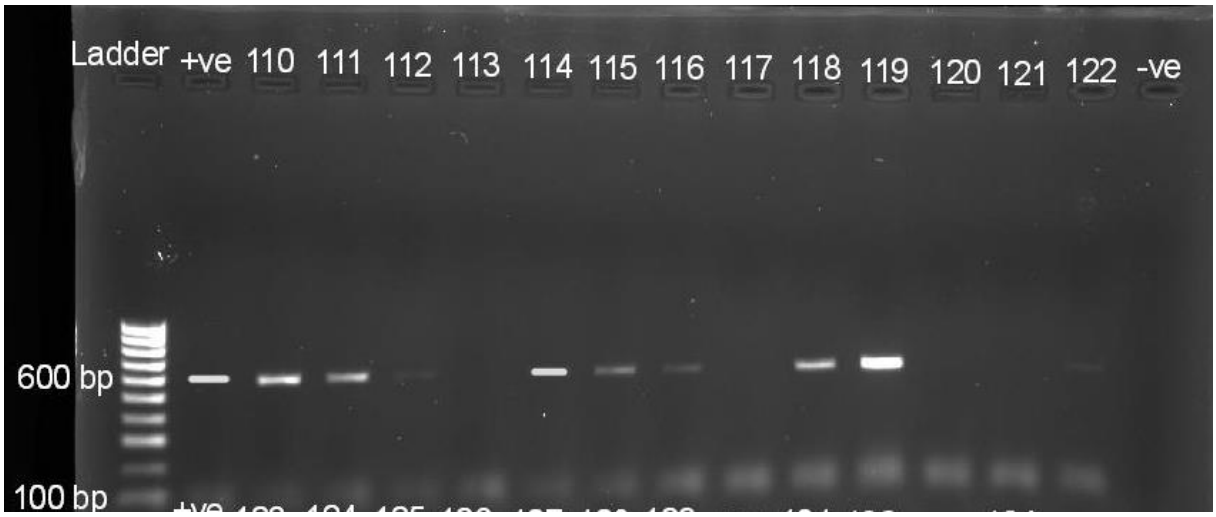



**Figure 4:** Gel image of CyRPA fragment (586 bp) amplified using CyRPA F1/R2 primer set. Bands in lane 110-122 represent amplified CyRPA F1/R2 fragments from different samples. The first lane represents 100 kb ladder, the second lane marked +ve represented positive control and contained fragment amplified from 3D7 DNA. The last lane (-ve) represents negative control (PCR water). Samples in lane 113, 117,120 and 121 had no bands and did not amplify thus regarded as negative.

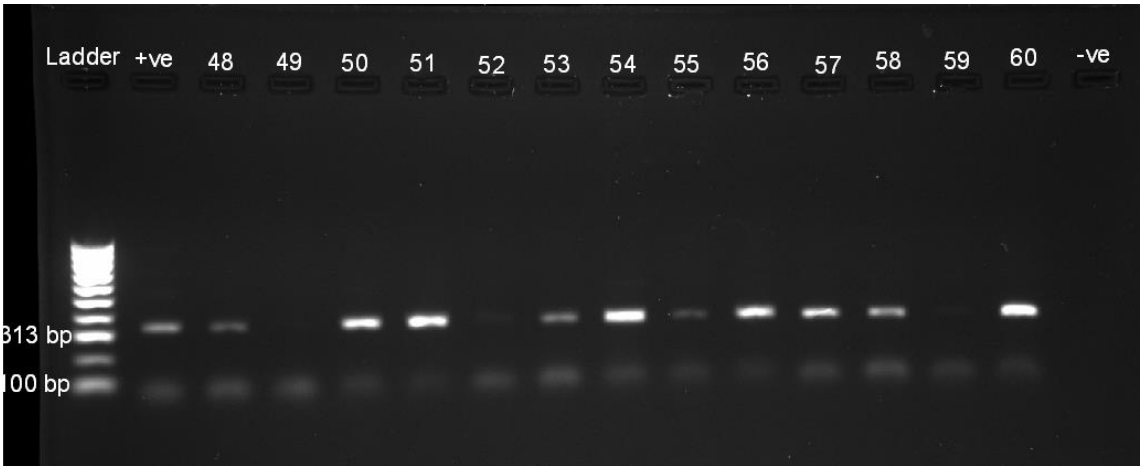



**Figure 5:** Gel image of CyRPA fragment (313 bp) amplified using CyRPA F3/R1 primer set**.** Bands in lane 48, 50, 51, 53, 54, 55, 56, 57, 58 and 60 represents amplified CyRPA fragments from different samples. The first lane represents 100 kb ladder, the second lane marked +ve represented positive control and contained fragment amplified from 3D7 DNA. The last lane (-ve) represents negative control (PCR water). Samples in lane 49, 52 and 59 had no bands and did not amplify thus regarded as negative.

The amplification left a section consisting of 280bp of the CyRPA gene not amplified as indicated in the figure below.

**A.**


**B.**


**Figure 6:** Schematic representation of the full length of CyRPA gene (**A**) and fragments that were successfully amplified by two different primers sets (**B**). The blue line represents the exons of 626bp and 463bp, respectively, whereas the green section represents the intron composed of 99bp. The description of 281 bp not amplified in all samples by the two primers set is in the form of the purple line. Most of the bp not amplified lies within the intron of the CyRPA gene and only 181bp (40bp from upstream exon and 141bp from downstream exon) from exon was not amplified. The red color represents CyRPA F1/R2 and CyRPA F3/R1 base pairs that were successfully amplified.

### 4.2 Ripr amplification

Out of 162 DNA samples, only 64 samples were successfully amplified using Ripr F3/R1 primer set and 84 samples using Ripr F6/R1 primer set. The remaining samples were not amplified even though adjustments such as an increase in template volume were made. The

inability of designed primers combinations to cover the entire length of the targeted Ripr gene (3261 bp) resulted in amplification of two fragments. Ripr F6/R1 primer set amplified a fragment of 3064 bp whereas Ripr F3/R1 primer set amplified 1935 bp. The fragments were within the 1000 bp hyper ladder (Bioline, London, UK) distribution as seen in Figure 4.4 and Figure 4.6. For quality assurance, the negative control had no band.

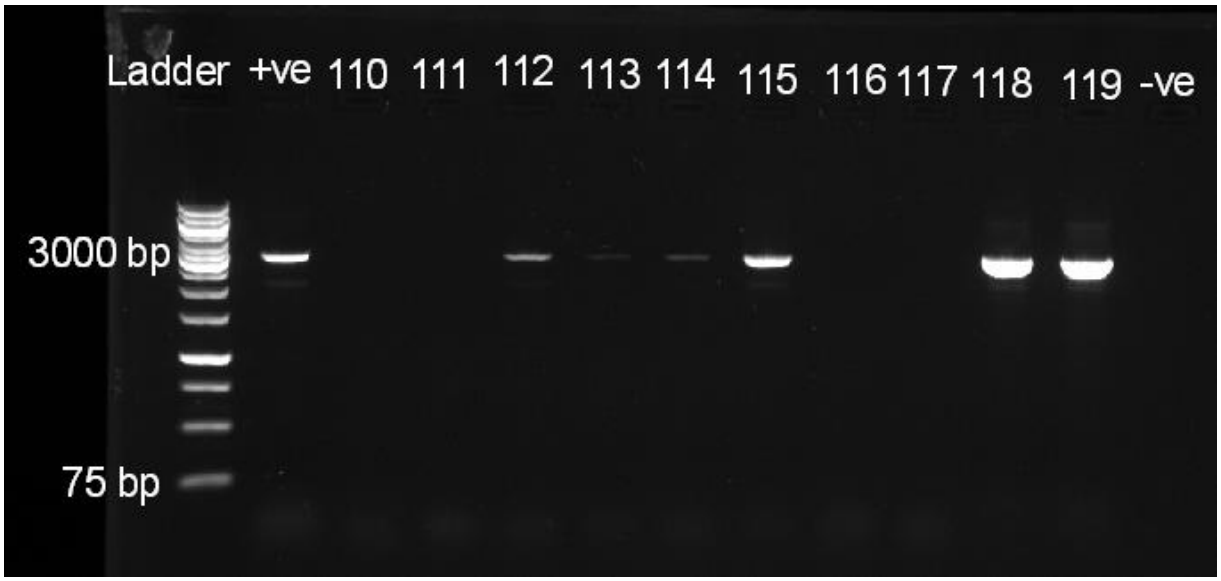



**Figure 7:** Gel image of Ripr fragment (3064 bp) amplified using Ripr F6/R1 primer set. Bands in lane 112, 113, 114, 115, 118 and 119 represent amplified Ripr fragments from different samples. The first lane represents 1 kb hyper ladder, the second lane marked +ve represented positive control and contained fragment amplified from 3D7 DNA. The last lane (-ve) represents negative control (PCR water). Samples in lane 110, 111, 116 and 117 had no bands and did not amplify thus regarded as negative.

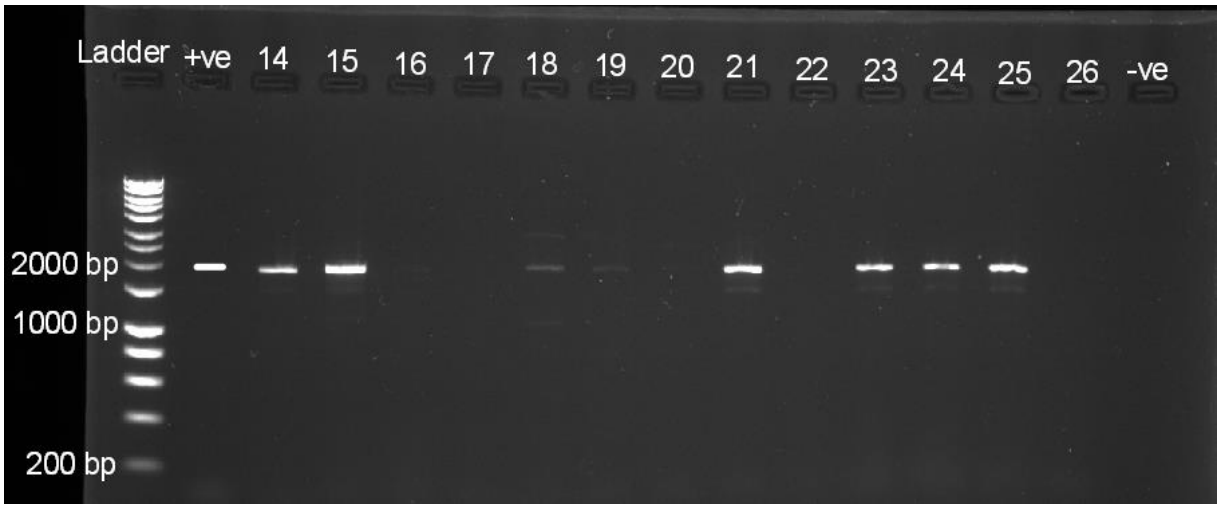



**Figure 8:** Gel image of Ripr fragment (1935 bp) amplified using Ripr F3/R1 primer set. Bands in lane 14, 15, 21, 23, 24 and 25 represent amplified Ripr fragments from different samples. The first lane represents 1 kb hyper ladder, the second lane marked +ve represented positive control and contained fragment amplified from 3D7 DNA. The last lane (-ve) represents negative control (PCR water).

Samples in lane 16, 17, 19, 20, 22 and 26 had no bands and did not amplify thus regarded as negative. The fainted bands below the target bands are amplified multiple bands with no impact on the target band upon sequencing.

The amplified fragment of Ripr gene was short of 196 bp upstream as shown in Figure 4.6 below.

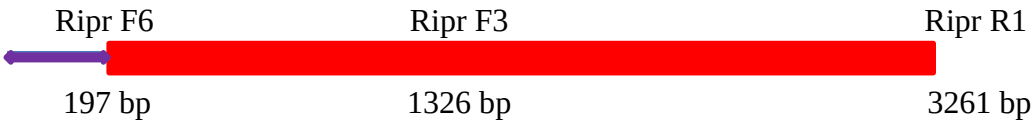



**Figure 9:** Schematic representation of the section of Ripr gene (F6R1 and F3R1 fragment) that was successfully amplified by the four different primer sets**.** The description of 196 bp that were not amplified in all samples by the four primer sets is in the form of the purple line. The red color represents the base pairs that were successfully amplified. The amplified F3/R1 fragment lies within the amplified length of F6/R1.

### 4.3 Sequencing results

Out of 132 CyRPA F3/R1 samples sequenced, a total of 107 samples had read lengths of around 313 bp after assembly against the reference sequence (PF3D7_0423800). 25 samples did not assemble due to short reads. Notably, of the 124 CyRPA F1/R2 amplicons only 52 samples had read lengths of around 586 bp after assembly against reference sequence PF3D7_0423800. Sixty nine samples failed to assemble due to poor peaks and short reads. From the 64 amplified Ripr F3/R1 fragments, only 22 had read lengths of around 1935 bp that correctly assembled against the reference sequence (PF3D7_0323400). Out of 82 F6/R1 amplicons sequenced, only 17 samples had read lengths of around 3064 bp that correctly assembled against the reference sequence (PF3D7_0323400). The samples that failed to assemble against the reference were excluded from the rest of the study.

### 4.4 CyRPA polymorphism analysis

DnaSP software (Librado and Rozas, 2009) was used to identify the segregating sites within the 107 CyRPA F3/R1 sequences. The sites that were segregating differed from the reference sequence and included sites with 3 mutations. A single synonymous change at position 1086 and two non-synonymous changes at position 973 and 1005 in samples 14729, 14343 and 14603 were identified, respectively. The nucleotide base Guanosine was substituted by Thymine in sample 14343, Guanosine was substituted by Cytosine in sample 14603 and Adenine by Thymine in sample 14729. There were no segregating sites within CyRPA F1/R2 sequences.

### 4.5 Genetic diversity of CyRPA polymorphisms

Genetic diversity of CyRPA F3/R1 fragments in the 107 parasites isolated from children with uncomplicated malaria in 2013 to 2014 season was defined by nucleotide diversity and analysis (Librado and Rozas, 2009). Nucleotide diversity ($\pi$) which is the average number of nucleotide differences in the 107 analysed sequences was 0.00038. Table 9 shows the respective codons within CyRPA (F3/R1 sequences) gene that bears both synonymous and non-synonymous mutations. Also displayed is the amino acid codes, which result from nucleotide base substitution.

**Table 9: CyRPA codon differences.**

| Position | Ref. Codon | Ref. Amino Acid Code | Variant Codon | Variant Amino Acid Code |
|---|---|---|---|---|
| **1005** | GAG | E | GAC | D |
| **973** | GTT | V | TTT | F |
| **1086** | GGA | G | GGT | G |

Statistical test of neutrality were conducted on the 107 CyRPA F3/R1 sequences to determine if they were under selection using the Tajim's D, Fu and Li D and F indices. The calculated value of Tajima's D was -1.5990 less than the P-value (0.05) (Table 9.1).

**Table 9.1: Genetic diversity of CyRPA F3/R1.**

| **CyRPA genetic diversity** | | | | | |
|---|---|---|---|---|---|
| **No of sequences** | **Π** | **S** | **Tajima's D** | **Fu and Li D** | **Fu and Li F** |
| **107** | 0.00038 | 3 | -1.5990* | -3.44560* | -3.35861* |

Neutrality test was also computed for the targeted region of CyRPA F3/R1 using a sliding window analysis with a length of 25 sites long and step size of 25 bases as presented in figure 4.7. The result from the output generated the DnaSP graph.

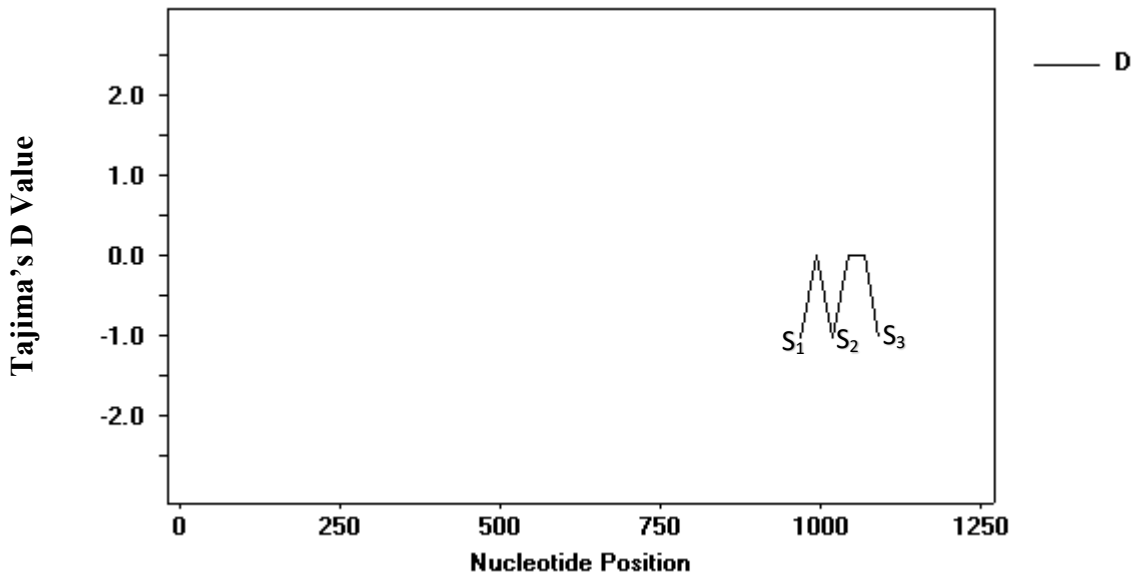



**Figure 10:** Tajima's D graph displayed from sliding window. The X-axis represents shows the nucleotide position whereas the Tajima's D value is represented on the Y-axis. From the graph, the Tajima's value for the segregating sites $S_1$ (973), $S_2$ (1005) and $S_3$ (1086) were not significant ($p<0.01$). The graph was constructed using a sliding window 25 sites long and a step size of 25 sites.

### 4.6 Ripr polymorphism analysis

DnaSP software was used to identify presence of segregating sites within the Ripr F6/R1 and Ripr F3/R1 sequences. There was no site that was segregating within the sequences or differed from the 3D7 laboratory reference sequence.

# CHAPTER FIVE

# DISCUSSION

Malaria pathogenicity is largely dependent on the success of merozoites invading the erythrocytes. The merozoites have therefore evolved multiple pathways, using various antigenic proteins which aid in the invasion process. Among the merozoite's invasive proteins are CyRPA and Ripr, which mediate the tight junction formation during invasion of erythrocytes. The discovery of the two antigens has revamped hope in the search for an effective malaria blood-stage vaccine (Favuzza *et al.,* 2016). Initially, research on malaria blood-stage vaccine focused on polymorphic, immune-dominant antigens, which are involved in the early stages of erythrocyte invasion such as merozoite surface protein (MSP) 1 (Ogutu *et al.,* 2009; Moss *et al.,* 2012). Despite all the efforts, some vaccines based on the polymorphic merozoite antigens have not only failed to elicit clinical protection but have also demonstrated limited efficacy during clinical trials (Halbroth and Draper, 2015). Recent research on the CyRPA gene identified only a single polymorphism (Dreyer *et al.,* 2012). With no substantial polymorphisms in the CyRPA and Ripr genes, the antigens have become potential targets or candidates for developing an effective malaria blood stage vaccine. The present study was undertaken with the aim of establishing whether there were SNPs in both CyRPA and Ripr genes in *P. falciparum* that were circulating in blood samples of children under 8 years in Kilifi county, a malaria endemic area in Kenya.

The study entailed direct capillary sequencing of both Ripr and CyRPA amplicons generated via gradient PCR from DNA samples obtained in 2013 and 2014. Some of the key advantages of using direct sequencing are that it is fast and reliable, since it can generate the entire sequence length. In order to obtain good quality and clean sequence data during sequencing, only PCR products with single strong bands on agarose gel visualization were used. However, in some instances, strong bands were selected over weak bands especially where there was the presence of multiple bands. The amplification process of both CyRPA and Ripr gene was divided into two fragments. The division was due to the inability of the selected primers to amplify the entire length of both genes. The CyRPA sequences that were obtained fell within the exonic region. The CyRPA gene has two exons, the first exon ranges from 1 bp to 626 bp whereas the second exon ranged from 726 bp to 1188bp. The two exons are separated by an intronic region comprising of 100 bp. Only 40 bp from an upstream exon,

141 bp from downstream exon and intronic region (100 bp) was not sequenced. The two regions of exons that were not amplified accounts for 181 bp. The regions (181 bp) not covered in this study may not contain a large number of polymorphic sites, since the majority of the analysed sequences in exon 1 had no SNPs and only three SNPs were identified in exon 2. In exon 2, three segregating sites at base pair position 973, 1005 and 1086 were identified and they were singletons. Two of the mutations were non-synonymous and one was synonymous. In previous findings by Dreyer *et al.* (2012), a single non-synonymous SNP was discovered after sequence analysis of 12 standard *P. falciparum* strains and 6 isolates from Tanzania. The standard strains in the study included 3D7, K1, MAD20, FC27, FCR3, RFCR3, W2met, HB3, RO-33, 7G8, ITG2F6, and FVO while the Tanzania Isolates were IFA4, IFA6, IFA10, IFA12, IFA18, and IFA19. From Dreyer *et al.* (2012) findings, all the Tanzania isolates lacked SNPs within their sequences, the discovered SNP was present in four of the standard strains K1 (Thailand), FCR3 (Gambia), ITG2F6 (Brazil), and FVO (Vietnam) at base pair position 1116 (Dreyer *et al.,* 2012). Contrary, to these findings where the SNP was at position 1116 of the four standard strains sequences, the non-synonymous SNPs in Kilifi samples were located at position 973 and 1005.

Non-synonymous SNPs resulted in changes in amino acids. Within protein sequence, the SNP at position 1005 resulted in a Glutamate/Aspartate change at amino acid position 335 out of total position 396. Nucleotide position 1086 translates to position 325 in the protein sequence, Valine was substituted by Phenylalanine. The positions at which these amino acid changes occurred are in close proximity to position (339) where Dreyer *et al.* (2012) observed the SNP which resulted in Arginine/Serine dimorphism. The low number of SNPs is an indication that CyRPA is not likely to be under immune pressure or a key target of an adaptive immune response. The SNPs are probably not located within the four epitope of growth inhibitory monoclonal antibodies targeting CyRPA antigen (which is made of polypeptide chains from position 26 to 127) (Favuzza *et al.,* 2016). Such suggestions are supported by other research which showed that generated full-length CyRPA antibodies are capable of blocking multiple strains of *P. falciparum* merozoites from invading erythrocytes (Reddy et al., 2014; Favuzza *et al.,* 2016).

The identified SNPs are singletons which are usually rare in populations and extend to include the nucleotide substitution at a position at 1086 that resulted in a silent mutation. The results implies that CyRPA gene does not harbour many mutations more so due to preservation of its antigen function of recruiting Ripr antigen into the Rh5/CyRPA/

P113/Basigin complex during erythrocyte invasion (Galaway *et al.*, 2017). Since this study to our knowledge examined the *cyrpa* gene in a malaria endemic population, it therefore appears to be a potentially good vaccine candidate given its important role in invasion. The presence of the rare variants in the population confirms the reasons behind a marginal number of sequence polymorphisms and reasons as to why the antigen (CyRPA) has limited natural immunogenicity (Favuzza *et al.*, 2016).

Through comparison of the differences between an average number of pairwise differences and number of segregating sites, Tajima's D aided in testing the neutral theory of molecular evolution in exon 2 of CyRPA gene. The neutral theory suggests that the vast majority of differences noted at the molecular level arose through spontaneous mutations which do not influence individual fitness (Tajima, 1989; Kimura 1983). This implies that the process of genetic drift is a key player in the evolution of genomes. Given a population that is constant in size and experiencing genetic drift, the average number of pairwise difference and segregating sites are usually equal. A deviation from the equal value is a key indicator that the sequences under analysis are acted on by different forms of selection. Positive Tajima's D value is an indication of a balancing selection. A negative value confirms the presence of purifying selection and is expected to reveal a high number of singletons. Statistical analysis showed that exon 2 of the CyRPA gene in the Kilifi population is under purifying selection with an excess of rare variants as evidenced by the negative value (-1.5990) of Tajima's D ($p<0.01$) and the Fu and Li's D and F statistics. The result implies that the gene has important function and does not acquire mutation so as to preserve its function.

There were no polymorphisms in the 52 analyzed sequences in Exon 1 of CyRPA (F1R2) circulating in population within Kilifi County. Exon 1 of CyRPA is more conserved when compared to Exon 2. Similar results were obtained in the analysis of the Ripr gene which was fragmented into two sections based on the primer set used in the generation of the amplicons. Out of 64 Ripr F3/R1 amplicons with 1935 bp, only 22 mapped correctly on to the Ripr reference sequence (3D7) and accounted for 13.58% of the amplified samples. Ripr F6/R1 fragment only accounted for 10.49% of the total sample population with sequences covering up to 3064 bp. The low number of amplicons attributes to the failure of the primer sets (F3/R1 and F6/R1) to amplify more samples, probability of some sample DNA being degraded and some of the amplified samples having poor reads upon sequencing. The few sequences of the two fragments, Ripr F3/R1 and Ripr F6/R1 lacked mutations. These results complement the findings by Chen *et al.*, (2011), which showed the Ripr gene as highly

conserved within the *P. falciparum*. The gene usually undergoes cleavage into two fragments of the approximate length of 65kDa each (Chen et al., 2011). Reticulocyte-binding protein homologue 5 (Ripr) F6R1 fragment covered amino acid at position 79 to 1072 (nucleotide position 235 to 3216). Therefore, the fragment analysed covered all the 10 epidermal growth factor (EGF) like domains. Two of the domains are located between amino acid in position 238 to 368 which designates the N-terminal of the Ripr antigen. The other 8 domains covered were between amino acids 791 to 900 designating the C-terminal region (Chen *et al.*, 2011). These domains are rich in cysteine residues and are relatively conserved (Savage *et al.,* 1973). Chen *et al.* (2011) established that prior to Ripr recruitment to the Rh5/CyRPA/P113 complex it undergoes processing which occurs at a position between the second and third domains. Based on the findings by Chen *et al.,* (2011), all sequences of both Ripr fragments generated in this study covered the key domain of Ripr antigen.

The results imply that anti-Ripr antibodies generated by Chen and co-workers to block the growth of *P. falciparum* would successfully block growth in all the analysed samples. Initially, research had established that anti-Ripr1 and anti-Ripr2 antibodies inhibited parasite growth in *P. falciparum* strains W2mef, FCR3, CSL2, T994, E8B, 7G8, HB3, MCAMP, D10, and 3D7 (Chen *et al.,* 2011). These findings provide a strong basis for the argument that a vaccine based on *P. falciparum* Ripr or CyRPA antigens will successfully elicit immune protection against *P. falciparum* circulating among the population of Kenyans residing in Kilifi County since they have no mutations.

## CONCLUSION

Exon 1 of CyRPA genes in *P. falciparum* circulating within the population of Kilifi County in Kenya lacks polymorphisms. However, in exon 2, there are three polymorphic sites that were singletons. The three SNPs were identified at positions 973, 1005 and 1086. Among these SNPs, two were non-synonymous and one was synonymous. The statistical analysis shows that in the population, exon 2 of CyRPA is under purifying selection. The limited number of mutations in the *Cyrpa* gene demonstrates the importance of the gene in the merozoites invasion of the erythrocytes (Favuzza *et al.*, 2016). The limited number of SNPs in exon 2 and lack of SNPs in Exon 1 of CyRPA gene presents CyRPA as a potential malaria vaccine candidate that will not exhibit allele-specific immunity.

The research complements a similar study conducted on samples obtained from Tanzania within Africa (Dreyer *et al.*, 2012). The CyRPA SNPs identified in this study are the first to be reported on field samples from Africa. The obtained data on CyRPA supports the idea that vaccines which are based on this antigen will work in Kenya if developed. Based on the results from this research, *P. falciparum* Ripr gene circulating within the Kilifi populations lacks polymorphism. The results suggest that previous antibodies generated against the Ripr antigen are likely to work if tested against the parasites in Kilifi, since there is no variation in this gene.

## RECOMMENDATIONS

In as much as the results obtained from analysis of CyRPA gene are in support of previous research work, the focus should be shifted towards confirming that the non-synonymous SNPs within exon 2 may not result in conformational changes that hinder the development of a strong vaccine. Also, the results on CyRPA exon 1 covered only 34% of the total samples, implying that the study may have underestimate the number of SNPs. Future studies should increase the number of samples and include the region between the exons that was not amplified in this study. For more diverse study within Kenya, the research should extend to cover other malaria endemic regions such as the Western part of the country. Data from other malaria endemic regions of the country will inform malaria vaccine strategies. Additionally, an increase in the sample size to confirm the Ripr data in this study and include the regions not sequenced in this study will provide more data to inform its potential as a malaria vaccine candidate.

## REFERENCES


Angrisano, F., Riglar, D. T., Sturm, A., Volz, J. C., Delves, M. J., Zuccala, E. S., Turnbull, L., Dekiwadia, C., Olshina, M. A., Marapana, D. S. et al. (2012). Spatial localisation of actin filaments across developmental stages of the malaria parasite. *PLoS ONE*, 7(2), e32188

Baldwin, M.R., Li, X., Hanada, T., Liu, S.C., and Chishti, A.H. (2015). Merozoite surface protein 1 recognition of host glycophorin A mediates malaria parasite invasion of red blood cells. *Blood* 125, 2704–2711.

Baum, J., L. Chen, J. Healer, S. Lopaticki, M. Boyle, T. Triglia, F. Ehlgen, S. A. Ralph, J. G. Beeson, and A. F. Cowman. (2009). Reticulocyte-binding protein homologue 5 - an essential adhesin involved in invasion of human *erythrocytes* by *Plasmodium falciparum*. *Int. J. Parasitol*. 39: 371–380. 54.

Beeson, J. G., Drew, D. R., Boyle, M. J., Feng, G., Fowkes, F. J., & Richards, J. S. (2016). Merozoite surface proteins in red blood cell invasion, immunity and vaccines against malaria. *FEMS microbiology reviews*, 40(3), 343-372.

Boyle MJ, Wilson DW, Richards JS, Riglar DT, Tetteh KKA, Conway DJ, Ralph SA, Baum J, Beeson JG. (2010). Isolation of viable *Plasmodium falciparum* merozoites to define erythrocyte invasion events and advance vaccine and drug development. *Proc. Natl. Acad. Sci.* U. S. A. 107:14378– 14383.

Boyle, M. J., Langer, C., Chan, J. A., Hodder, A. N., Coppel, R. L., Anders, R. F., & Beeson, J. G. (2014). Sequential processing of merozoite surface proteins during and after erythrocyte invasion by Plasmodium falciparum. *Infection and immunity*, 82(3), 924-936.

Bustamante LY, Bartholdson SJ, Crosnier C, Campos MG, Wanaguru M, Nguon C, Kwiatkowski DP, Wright GJ, Rayner JC (2013). A full-length recombinant *Plasmodium falciparum* PfRh5 protein induces inhibitory antibodies that are effective across common PfRh5 genetic variants. *Vaccine*, 31:373–379

Chen, L., Lopaticki, S., Riglar, D. T., Dekiwadia, C., Uboldi, A. D., Tham, W.-H., Cowman, A. F. (2011). An EGF-like protein forms a complex with PfRh5 and is required for invasion of human erythrocytes by *Plasmodium falciparum*. *PLoS Pathogens*, 7(9), e1002199.

Chen, Q., Schlichtherle, M., & Wahlgren, M. (2000). Molecular aspects of severe malaria. *Clinical microbiology reviews*, 13(3), 439-450.

Child, M. A., Epp, C., Bujard, H., & Blackman, M. J. (2010). Regulated maturation of malaria merozoite surface protein-1 is essential for parasite growth. *Molecular microbiology*, 78(1), 187-202.

Cowman, A. F., & Crabb, B. S. (2006). Invasion of red blood cells by malaria parasites. *Cell*, 124(4), 755-766.

Das, S., Hertrich, N., Perrin, A. J., Withers-Martinez, C., Collins, C. R., Jones, M. L., & Wright, G. J. (2015). Processing of Plasmodium falciparum merozoite surface protein MSP1 activates a spectrin-binding function enabling parasite egress from RBCs. *Cell host & microbe*, *18*(4), 433-444.

Douglas, A. D., Williams, A. R., Illingworth, J. J., Kamuyu, G., Biswas, S., Goodman, A. L., & Long, C. A. (2011). The blood-stage malaria antigen PfRH5 is susceptible to vaccine-inducible cross-strain neutralizing antibody. *Nature communications*, *2*, 601.

Douglas, A. D., Williams, A. R., Knuepfer, E., Illingworth, J. J., Furze, J. M., Crosnier, C., & Alanine, D. G. (2014). Neutralization of Plasmodium falciparum merozoites by antibodies against PfRH5. *The Journal of Immunology*, *192*(1), 245-258.

Dreyer, A. M., Matile, H., Papastogiannidis, P., Kamber, J., Favuzza, P., Voss, T. S., and Wittlin, S. (2012). Passive immunoprotection of Plasmodium falciparum-infected mice designates the CyRPA as candidate malaria vaccine antigen. *The Journal of Immunology*, *188*(12), 6225-6237.

Dzikowski, R., & Deitsch, K. W. (2009). Genetics of antigenic variation in *Plasmodium falciparum. Current genetics*, 55(2), 103-110.

Elliott, S. R., & Beeson, J. G. (2008). Estimating the burden of global mortality in children aged< 5 years by pathogen-specific causes. *Clinical infectious diseases*, 46(11), 1794-1795.

Fairhurst, R. M., & Dondorp, A. M. (2016). Artemisinin-resistant Plasmodium falciparum malaria. *Microbiology spectrum*, *4*(3).

Favuzza, P., Blaser, S., Dreyer, A. M., Riccio, G., Tamborrini, M., Thoma, R., & Pluschke, G. (2016). Generation of Plasmodium falciparum parasite-inhibitory antibodies by immunization with recombinantly-expressed CyRPA. *Malaria Journal*, *15*(1), 161.

Galaway, F., Drought, L. G., Fala, M., Cross, N., Kemp, A. C., Rayner, J. C., & Wright, G. J. (2017). P113 is a merozoite surface protein that binds the N terminus of Plasmodium falciparum RH5. *Nature communications*, 8, 14333.

Gao, X., Gunalan, K., Yap, S. S. L., & Preiser, P. R. (2013). Triggers of key calcium signals during erythrocyte invasion by Plasmodium falciparum. *Nature communications*, 4, 2862.

Garcia, C. R., de Azevedo, M. F., Wunderlich, G., Budu, A., Young, J. A., & Bannister, L. (2008). Plasmodium in the postgenomic era: new insights into the molecular cell biology of malaria parasites. *International review of cell and molecular biology*, 266, 85-156.

Geels, M.J., Imoukhuede, E.B., Imbault, N., Van Schooten, H., McWade, T., Troye Blomberg, M., Dobbelaer, R., Craig, A.G., and Leroy, O. (2011). European Vaccine Initiative: lessons from developing malaria vaccines. *Expert Review of Vaccines* 10, 1697 1708.

Gilson, P. R., & Crabb, B. S. (2009). Morphology and kinetics of the three distinct phases of red blood cell invasion by Plasmodium falciparum merozoites. *International journal for parasitology*, 39(1), 91-96..

Haase, S., Cabrera, A., Langer, C., Treeck, M., Struck, N., Herrmann, S., & Cowman, A. F. (2008). Characterization of a conserved rhoptry-associated leucine zipper-like protein in the malaria parasite Plasmodium falciparum. *Infection and immunity*, 76(3), 879-887.

Halbroth, B. R., & Draper, S. J. (2015). Recent developments in malaria vaccinology. In *Advances in parasitology* (Vol. 88, pp. 1-49). Academic Press.

Halliday, K. E., Okello, G., Turner, E. L., Njagi, K., Mcharo, C., Kengo, J., ... & Brooker, S. J. (2014). *Impact of intermittent screening and treatment for malaria among school children in Kenya: a cluster randomized trial*. The World Bank.

Hanssen, E., Dekiwadia, C., Riglar, D. T., Rug, M., Lemgruber, L., Cowman, A. F., ... & Ralph, S. A. (2013). Electron tomography of Plasmodium falciparum merozoites

reveals core cellular events that underpin erythrocyte invasion. *Cellular microbiology*, 15(9), 1457-1472.

Hayton, K., Dumoulin, P., Henschen, B., Liu, A., Papakrivos, J., & Wellems, T. E. (2013). Various PfRH5 polymorphisms can support Plasmodium falciparum invasion into the erythrocytes of owl monkeys and rats. *Molecular and biochemical parasitology*, 187(2), 103-110.

Josling, G. A., & Llinás, M. (2015). Sexual development in Plasmodium parasites: knowing when it's time to commit. *Nature Reviews Microbiology*, 13(9), 573.

Kimura, M. (1983). *The neutral theory of molecular evolution*. Cambridge University Press.

Librado, P., & Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, *25*(11), 1451-1452.

Miller, L. H., Baruch, D. I., Marsh, K., & Doumbo, O. K. (2002). The pathogenic basis of malaria. *Nature*, 415(6872), 673.

Mogeni, P., Williams, T. N., Fegan, G., Nyundo, C., Bauni, E., Mwai, K., ... & Berkley, J. A. (2016). Age, spatial, and temporal variations in hospital admissions with malaria in Kilifi County, Kenya: a 25-year longitudinal observational study. *PLoS medicine*, *13*(6), e1002047.

Moss, D. K., Remarque, E. J., Faber, B. W., Cavanagh, D. R., Arnot, D. E., Thomas, A. W., & Holder, A. A. (2012). Plasmodium falciparum 19-kilodalton merozoite surface protein 1 (MSP1)-specific antibodies that interfere with parasite growth in vitro can inhibit MSP1 processing, merozoite invasion, and intracellular parasite development. *Infection and immunity*, 80(3), 1280-1287.

Ogutu, B. R., Apollo, O. J., McKinney, D., Okoth, W., Siangla, J., Dubovsky, F., Withers, M. R. (2009). Blood stage malaria vaccine eliciting high antigen-specific antibody concentrations confers no protection to young children in Western Kenya. *PloS One*, *4*(3), e4708.

Ogwang, C., Afolabi, M., Kimani, D., Jagne, Y. J., Sheehy, S. H., Bliss, C. M., ... & Anagnostou, N. A. (2013). Safety and immunogenicity of heterologous prime-boost immunisation with Plasmodium falciparum malaria candidate vaccines, ChAd63 ME-TRAP and MVA ME-TRAP, in healthy Gambian and Kenyan adults. *PloS one*, 8(3), e57726.

Osier, F. H., Mackinnon, M. J., Crosnier, C., Fegan, G., Kamuyu, G., Wanaguru, M., ... & Marsh, K. (2014). New antigens for a multicomponent blood-stage malaria vaccine. *Science translational medicine*, *6*(247), 247ra102-247ra102.

Patel, S. D., Ahouidi, A. D., Bei, A. K., Dieye, T. N., Mboup, S., Harrison, S. C., & Duraisingh, M. T. (2013). *Plasmodium falciparum* merozoite surface antigen, PfRh5, elicits detectable levels of invasion-inhibiting antibodies in humans. *The Journal of Infectious Diseases*, *208*(10), 1679–1687.

Reddy, K. S., Amlabu, E., Pandey, A. K., Mitra, P., Chauhan, V. S., & Gaur, D. (2015). Multiprotein complex between the GPI-anchored CyRPA with PfRH5 and PfRipr is

crucial for Plasmodium falciparum erythrocyte invasion. *Proceedings of the National Academy of Sciences*, 112(4), 1179-1184.

Reddy, K. S., Pandey, A. K., Singh, H., Sahar, T., Emmanuel, A., Chitnis, C. E., ... & Gaur, D. (2014). Bacterially expressed full-length recombinant Plasmodium falciparum RH5 protein binds erythrocytes and elicits potent strain-transcending parasite-neutralizing antibodies. *Infection and immunity*, 82(1), 152-164.

Richards JS, Arumugam TU, Reiling L, Healer J, Hodder AN, Fowkes FJI, Cross N, Langer C, Takeo S, Uboldi AD, Thompson JK, Gilson PR, Coppel RL, Siba PM, King CL, Torii M, Chitnis CE, Narum DL, Mueller I, Crabb BS, Cowman AF, Tsuboi T, Beeson JG. (2013). Identification and prioritization of merozoite antigens as targets of protective human immunity to Plasmodium falciparum malaria for vaccine and biomarker development. *The Journal of Immunology*, 1300778.

Richards, J. S., Arumugam, T. U., Reiling, L., Healer, J., Hodder, A. N., Fowkes, F. J., ... & Thompson, J. K. (2013). Identification and prioritization of merozoite antigens as targets of protective human immunity to Plasmodium falciparum malaria for vaccine and biomarker development. *The Journal of Immunology*, 1300778.

Riglar DT, Richard D, Wilson DW, Boyle MJ, Dekiwadia C, Turnbull L, Angrisano F, Marapana DS, Rogers KL, Whitchurch CB, Beeson JG, Cowman AF, Ralph SA, Baum J. (2011). Super-resolution dissection of coordinated events during malaria parasite invasion of the human erythrocyte. *Cell host & microbe*, 9(1), 9-20.

Riglar, D. T., Whitehead, L., Cowman, A. F., Rogers, K. L., & Baum, J. (2016). Localisation-based imaging of malarial antigens during erythrocyte entry reaffirms a role for AMA1 but not MTRAP in invasion. *J Cell Sci*, 129(1), 228-242.

Ruecker, A., Shea, M., Hackett, F., Suarez, C., Hirst, E. M., Milutinovic, K., ... & Blackman, M. J. (2012). Proteolytic activation of the essential parasitophorous vacuole cysteine protease SERA6 accompanies malaria parasite egress from its host erythrocyte. *Journal of Biological Chemistry*, jbc-M112.

Santos, J. M., Graindorge, A., & Soldati-Favre, D. (2012). New insights into parasite rhomboid proteases. *Molecular and biochemical parasitology*, 182(1-2), 27-36.

Savage, C. R., Hash, J. H., & Cohen, S. (1973). Epidermal growth factor location of disulfide bonds. *Journal of Biological Chemistry*, 248(22), 7669-7672.

Silvestrini, F., Lasonder, E., Olivieri, A., Camarda, G., van Schaijk, B., Sanchez, M., & Alano, P. (2010). Protein export marks the early phase of gametocytogenesis of the human malaria parasite Plasmodium falciparum. *Molecular & Cellular Proteomics*, 9(7), 1437-1448.

Sirima, S. B., Cousens, S., & Druilhe, P. (2011). Protection against malaria by MSP3 candidate vaccine. *New England Journal of Medicine*, 365(11), 1062-1064.

Sultana, M., Sheikh, N., Mahumud, R. A., Jahir, T., Islam, Z., & Sarker, A. R. (2017). Prevalence and associated determinants of malaria parasites among Kenyan children. *Tropical medicine and health*, *45*, 25. doi:10.1186/s41182-017-0066-5

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, *123*(3), 585-595.

Takala, S. L., & Plowe, C. V. (2009). Genetic diversity and malaria vaccine design, testing and efficacy: preventing and overcoming 'vaccine resistant malaria'. *Parasite immunology*, 31(9), 560-573.

Tamminga, C., Sedegah, M., Maiolatesi, S., Fedders, C., Reyes, S., Reyes, A., Richie, T. L. (2013). Human adenovirus 5-vectored *Plasmodium falciparum* NMRC-M3V-Ad-PfCA vaccine encoding CSP and AMA1 is safe, well-tolerated and immunogenic but does not protect against controlled human malaria infection. *Human Vaccines & Immunotherapeutics*, 9(10), 2165–2177.

Thera MA, Doumbo OK, Coulibaly D, Laurens MB, Ouattara A, Kone AK, Guindo AB, Traore K, Traore I, Kouriba B, Diallo DA, Diarra I, Daou M, Dolo A, Tolo Y, Sissoko MS, Niangaly A, Sissoko M, Takala- Harrison S, Lyke KE, Wu Y, Blackwelder WC, Godeaux O, Vekemans J, Dubois MC, Ballou WR, Cohen J, Thompson D, Dube T, Soisson L, Diggs CL, House B, Lanar DE, Dutta S, Heppner DG, Jr, Plowe CV. (2011). A field trial to assess a blood-stage malaria vaccine. *New England Journal of Medicine*, 365:1004–1013.

Tran, T. M., Ongoiba, A., Coursen, J., Crosnier, C., Diouf, A., Huang, C., Crompton, P. D. (2014). Naturally acquired antibodies specific for Plasmodium falciparum reticulocyte-binding protein homologue 5 inhibit parasite growth and predict protection from malaria. *The Journal of infectious diseases*, 209(5), 789-798.

Volz, J. C., Yap, A., Sisquella, X., Thompson, J. K., and Nicholas, T. Y. (2016). Essential role of the PfRh5/PfRipr/CyRPA complex during Plasmodium falciparum invasion of erythrocytes. *Cell host & microbe*, 20(1), 60-71.

Weiss, G. E., Gilson, P. R., Taechalertpaisarn, T., Tham, W., Jong, N. W. M. De, Harvey, K. L., Crabb, B. S. (2015). Revealing the sequence and resulting cellular morphology of receptor-ligand interactions during Plasmodium falciparum invasion of erythrocytes. *PLoS pathogens*, 11(2), e1004670.

Withers-Martinez C, Strath M, Hackett F, Haire LF, Howell SA, Walker PA, Christodoulou E, Dodson GG, Blackman MJ (2014). The malaria parasite egress protease SUB1 is a calcium-dependent redox switch subtilisin. *Nature communications*, 5, 3726.

Withers-Martinez, C., Suarez, C., Fulle, S., Kher, S., Penzo, M., Ebejer, J.P., Koussis, K., Hackett, F., Jirgensons, A., Finn, P., and Blackman, M.J. (2012). Plasmodium subtilisin-like protease 1 (SUB1): insights into the active-site structure, specificity and function of a pan-malaria drug target. *International journal for parasitology, 42(6), 597-612.*

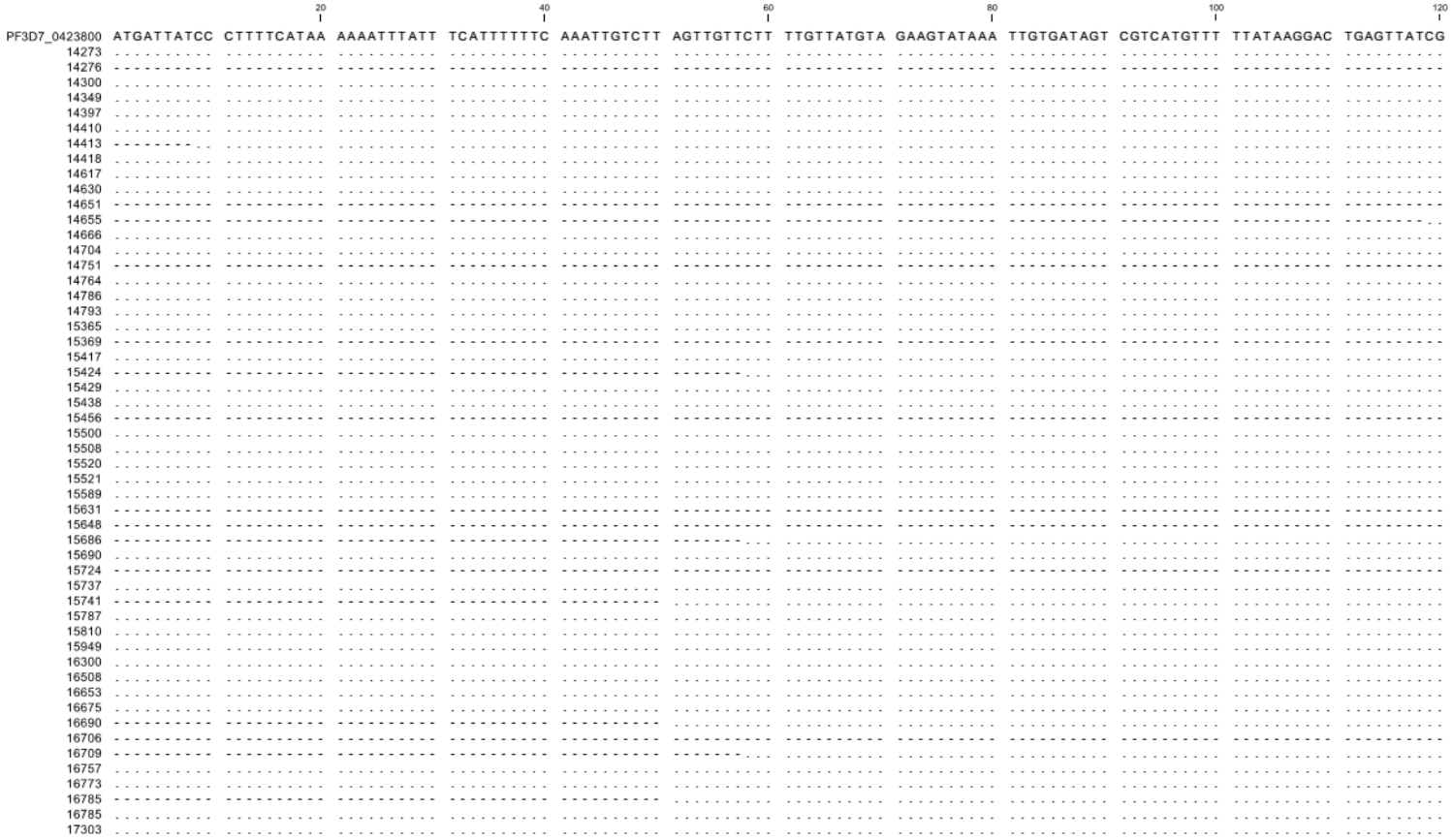
World Health Organization. (2015). *Global technical strategy for malaria 2016-2030*. World Health Organization.

Wright, KE, Hjerrild, KA, Bartlett, J, Douglas, AD, Jin, J, Brown, RE, Illingworth, JJ, Ashfield, R, Clemmensen, SB, de Jongh, WA, Draper, SJ and Higgins, MK. (2014). Structure of malaria invasion protein Rh5 with erythrocyte basigin and blocking antibodies. *Nature*, 515(7527), 427–430.

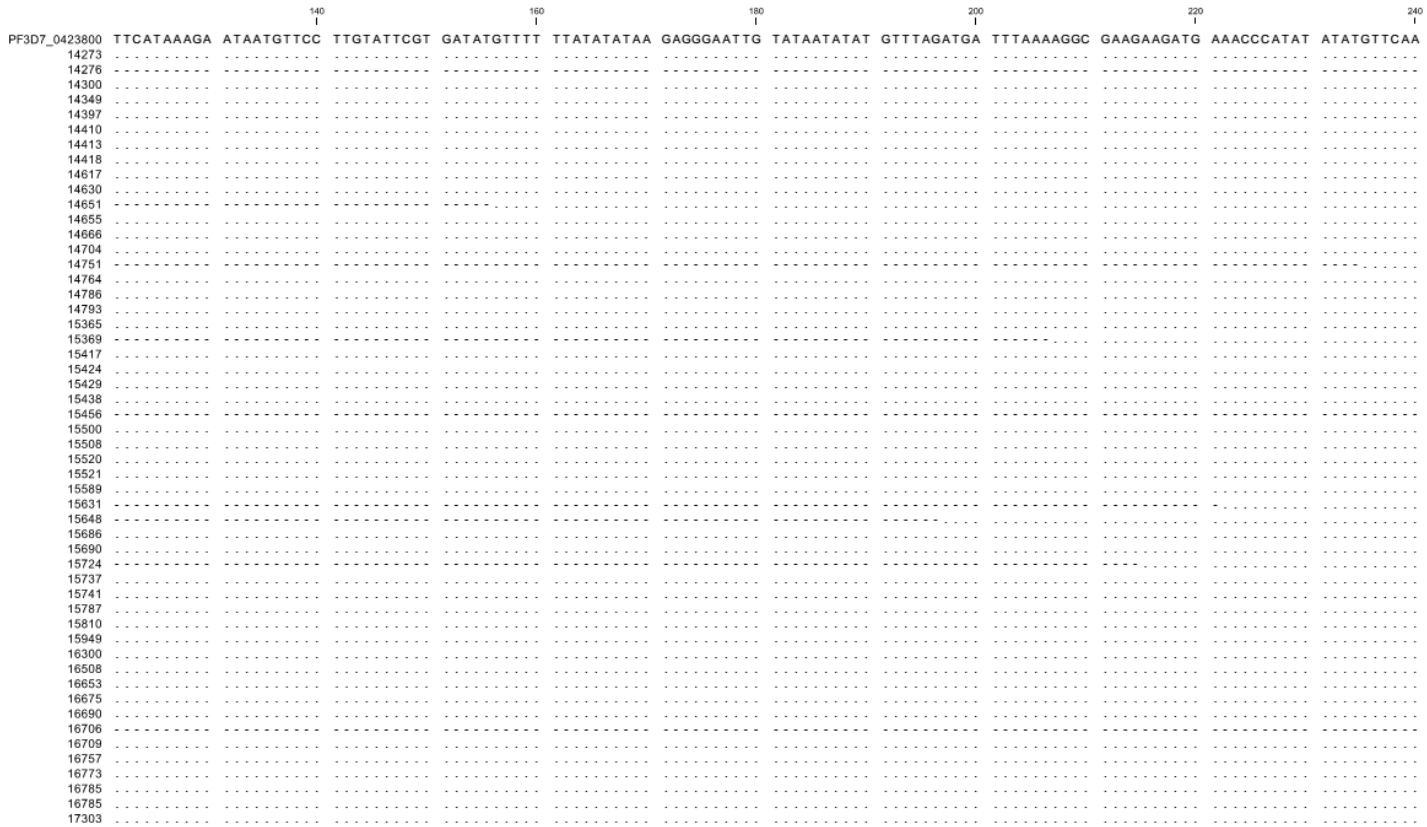
Zuccala ES, Gout AM, Dekiwadia C, Marapana DS, Angrisano F, Turnbull L, Riglar DT, Rogers KL, Whitchurch CB, Ralph SA, Speed TP, Baum J. (2012). Subcompartmentalisation of proteins in the rhoptries correlates with ordered events of erythrocyte invasion by the blood stage malaria parasite. *PloS one*, 7(9), e46160.

## APPENDICES

## Appendix 1. Nucleotide alignment for CyRPA Exon 1 (F1R2)

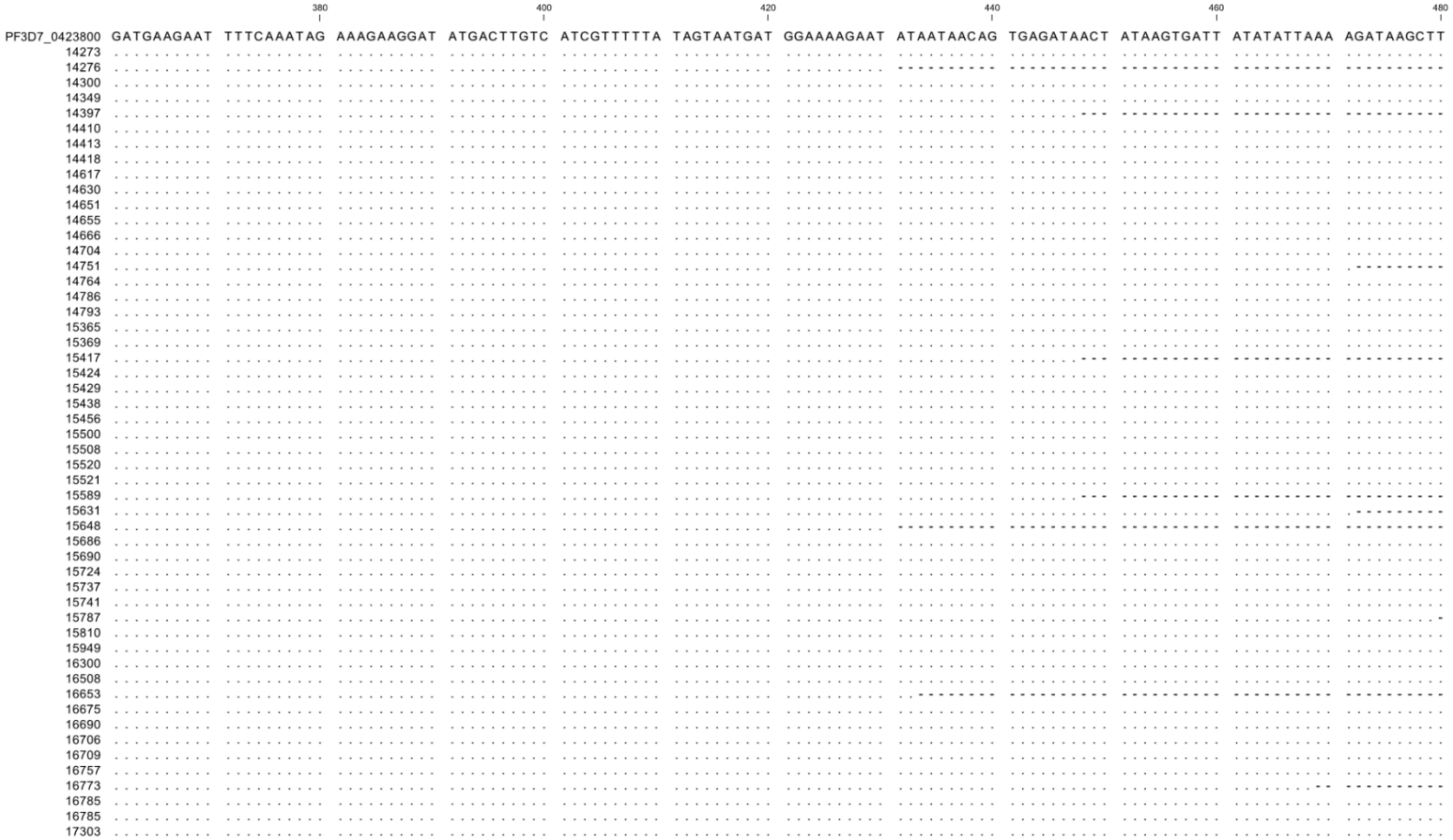


## Appendix 2. Nucleotide alignment for CyRPA Exon 1 (F1R2)

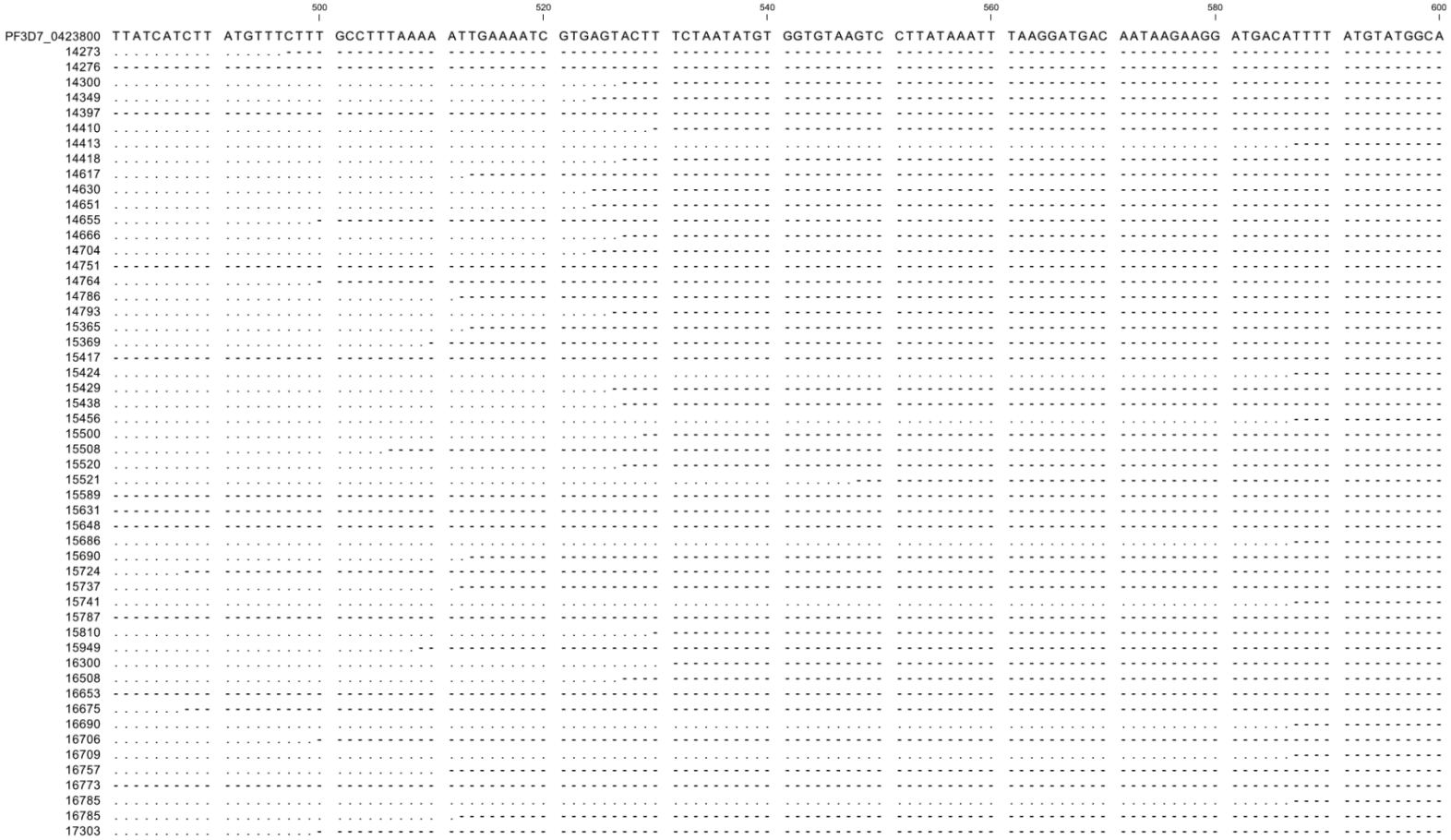


## Appendix 3. Nucleotide alignment for CyRPA Exon 1 (F1R2)

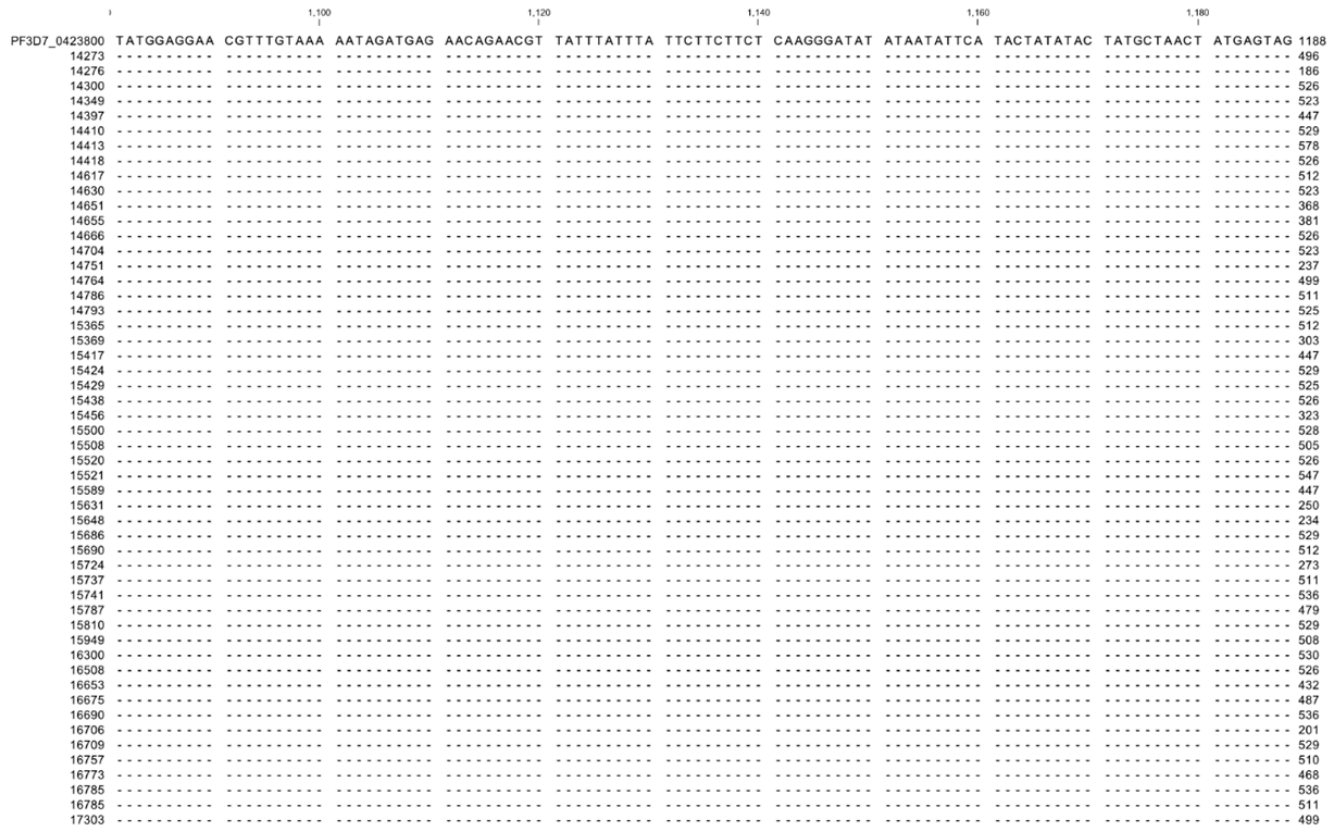
## Appendix 4. Nucleotide alignment for CyRPA Exon 1 (F1R2)



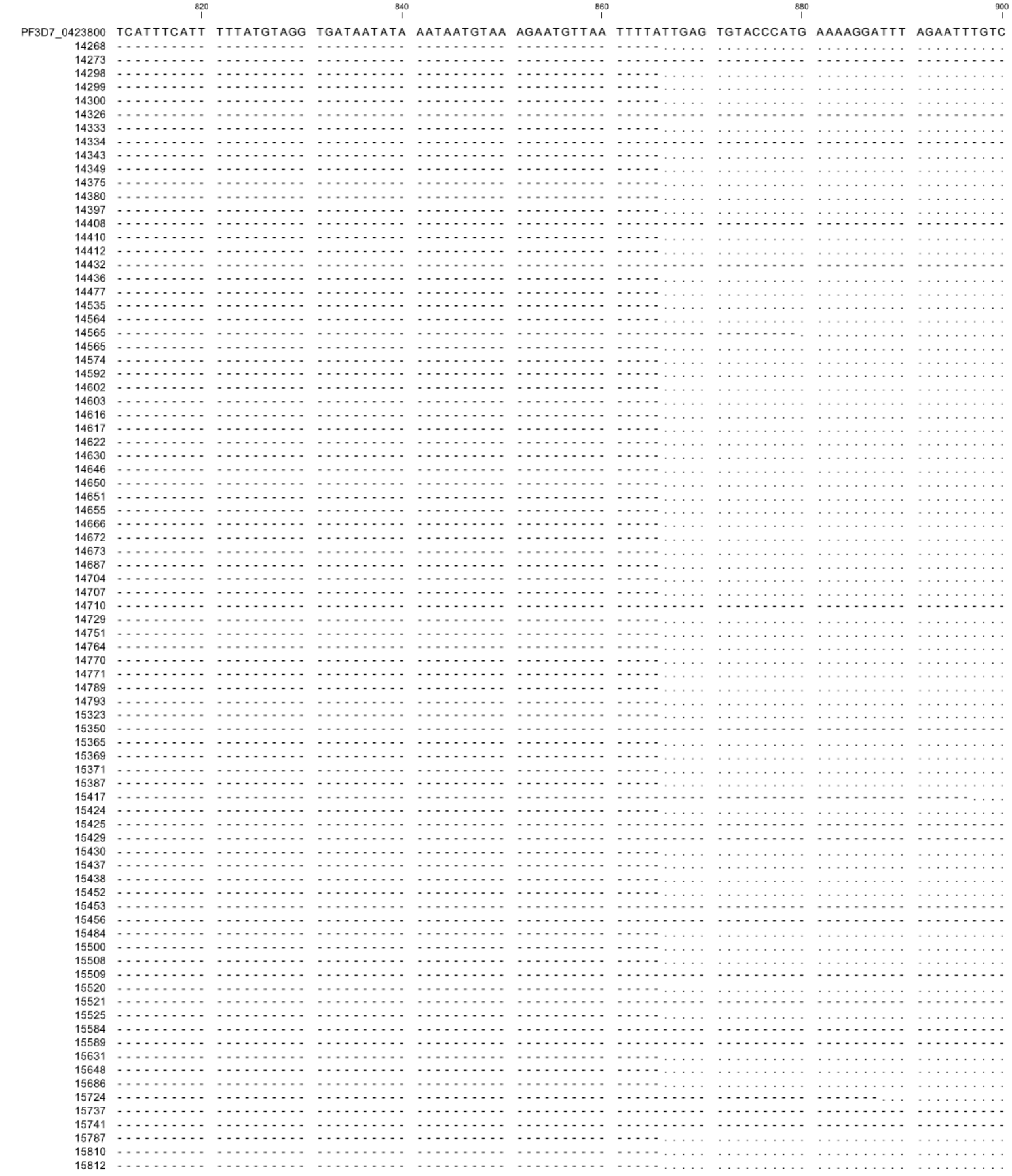
## Appendix 5. Nucleotide alignment for CyRPA Exon 1 (F1R2)



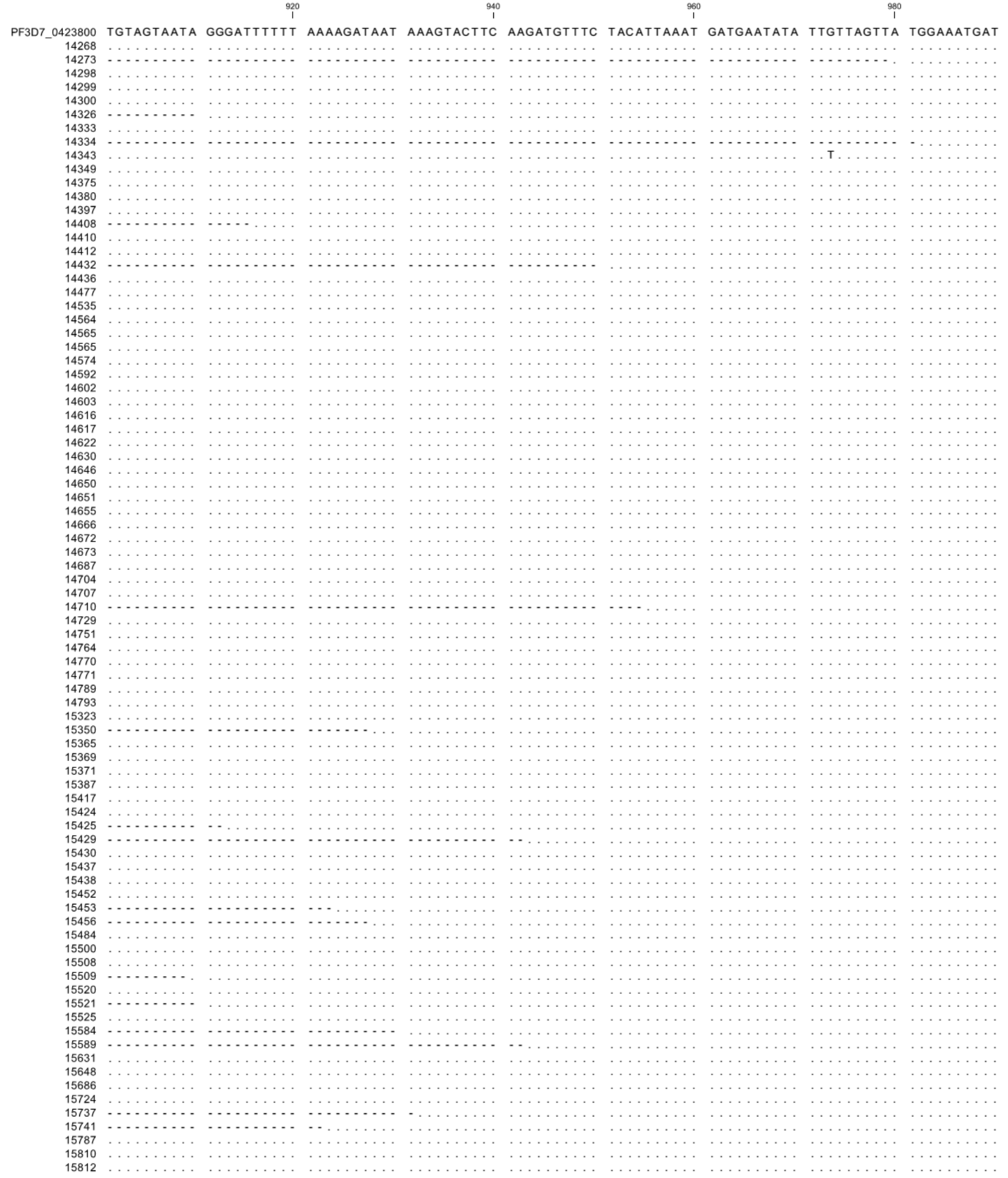
## Appendix 6. Nucleotide alignment for CyRPA Exon 1 (F1R2)



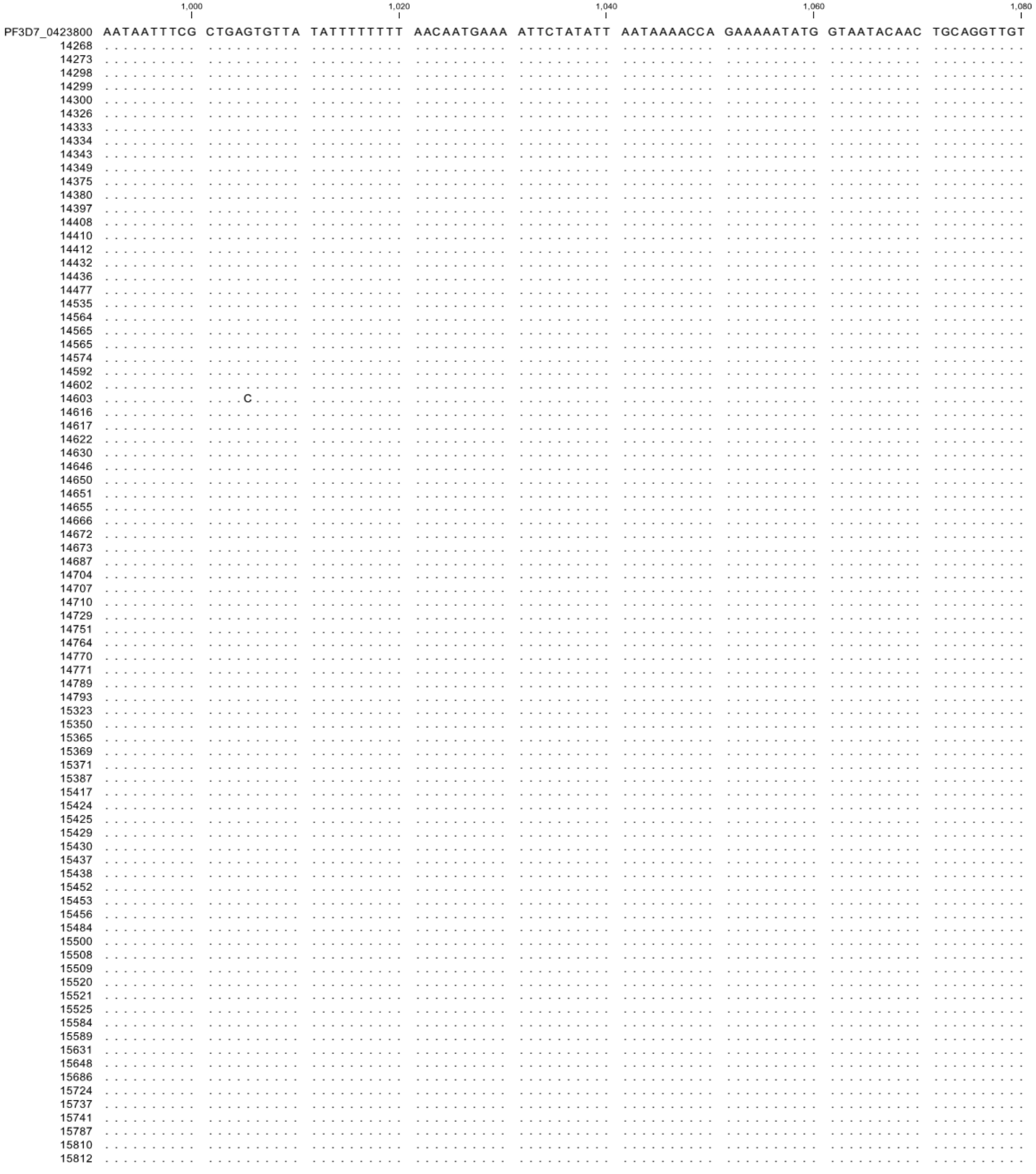
## Appendix 7. Nucleotide alignment for CyRPA Exon 2 (F3R1)



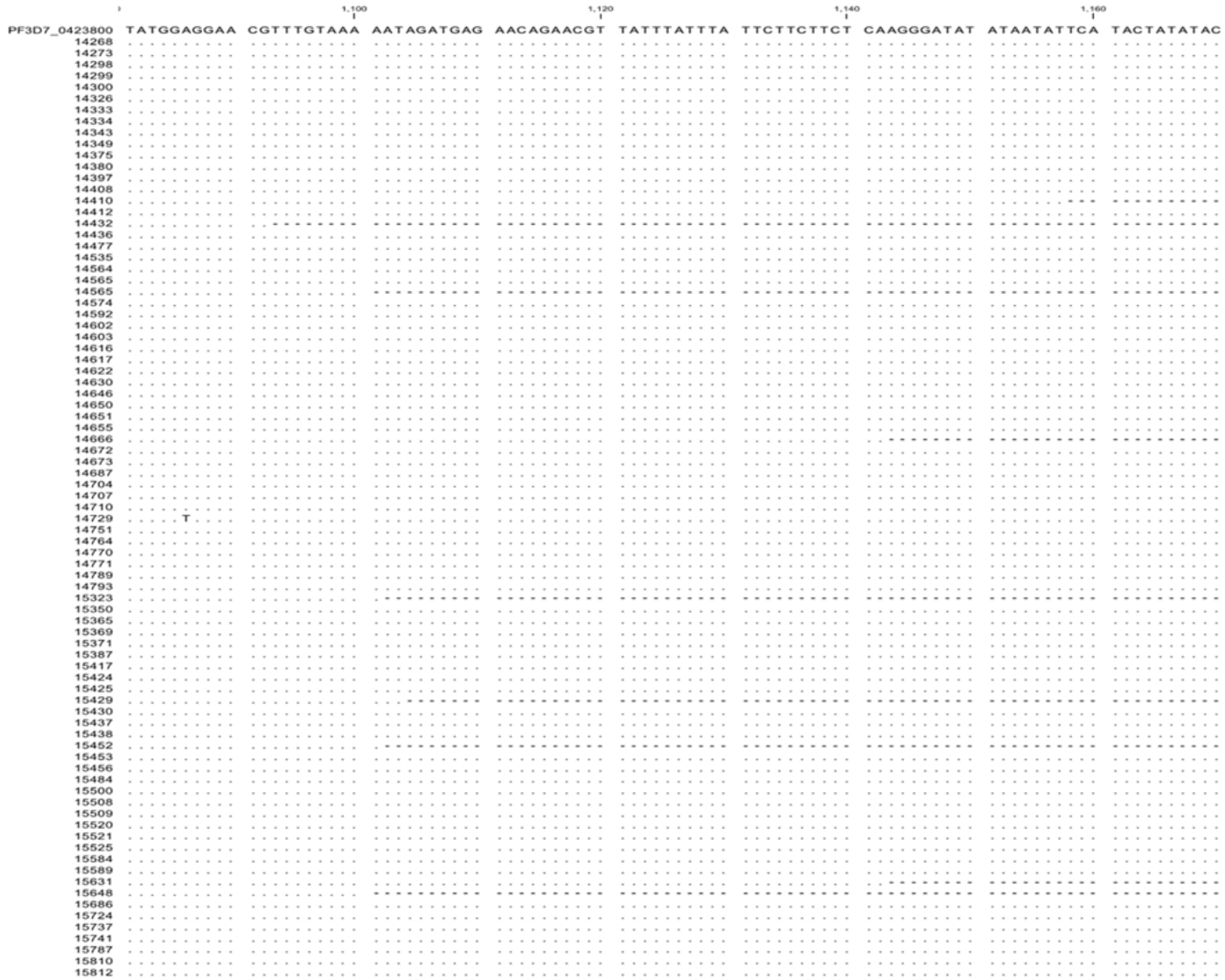
## Appendix 8. Nucleotide alignment for CyRPA Exon 2 (F3R1)



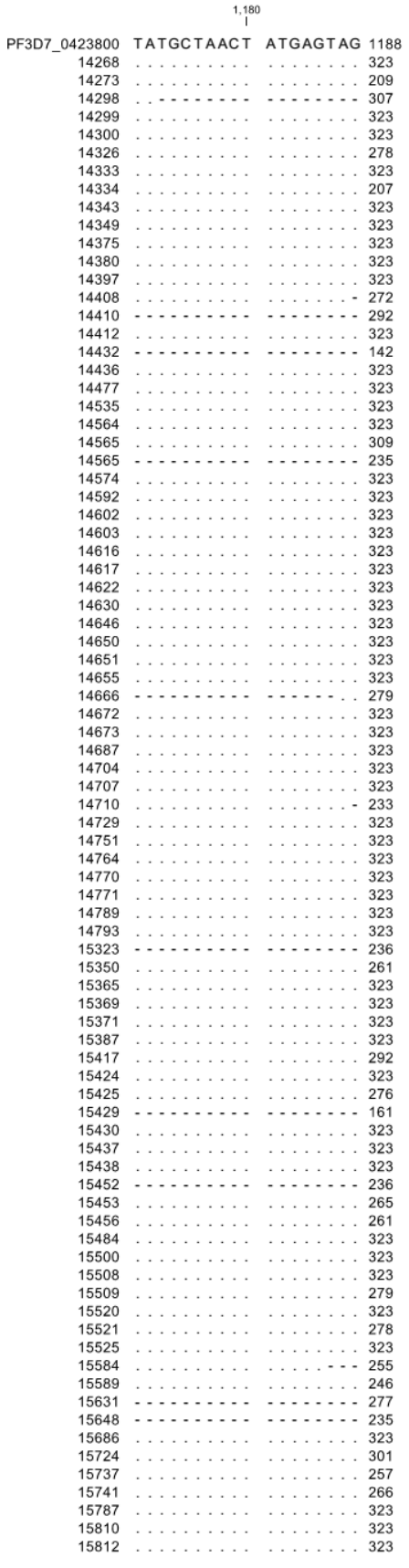
## Appendix 9. Nucleotide alignment for CyRPA Exon 2 (F3R1)



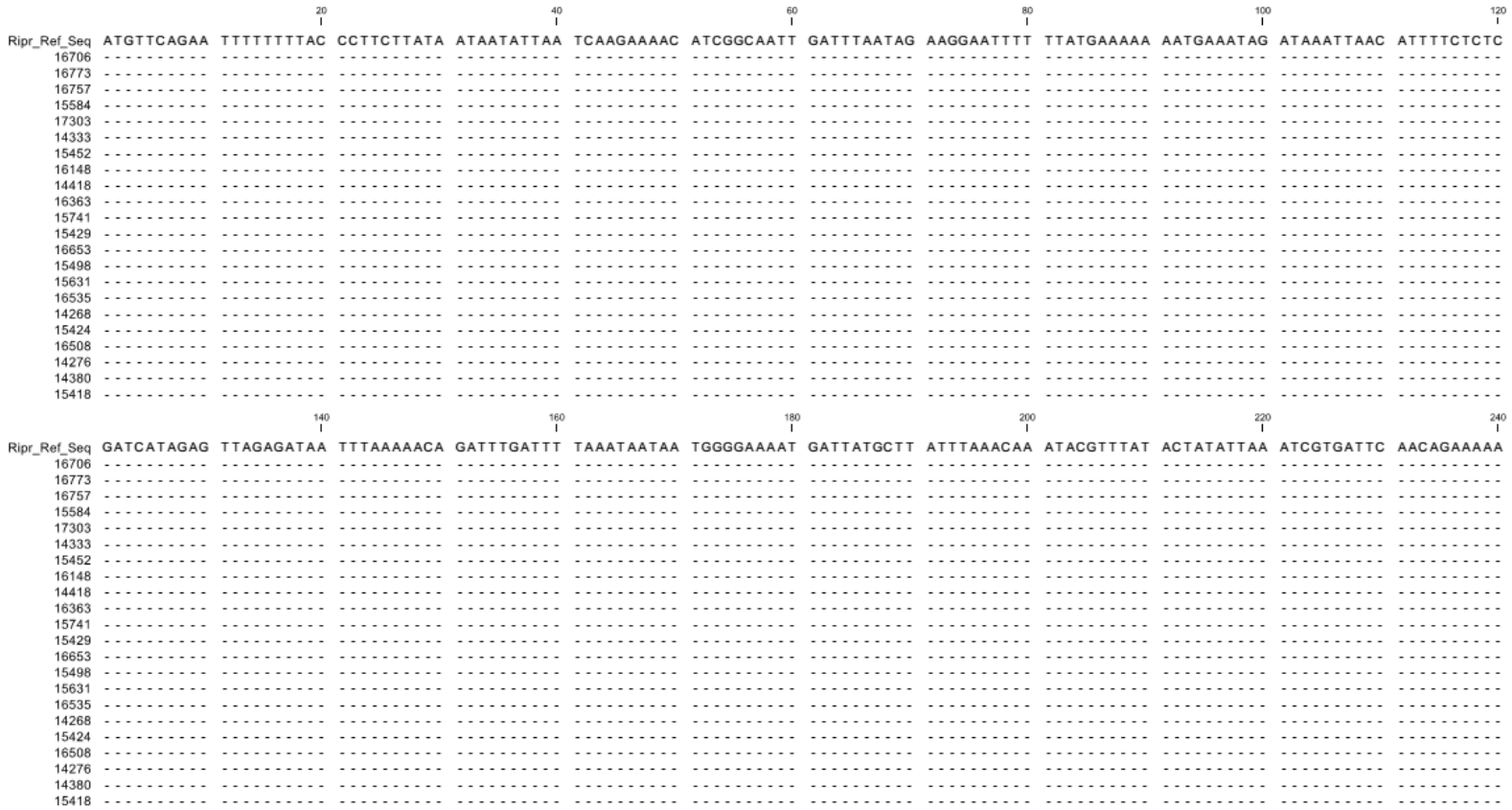
## Appendix 10. Nucleotide alignment for CyRPA Exon 2 (F3R1)



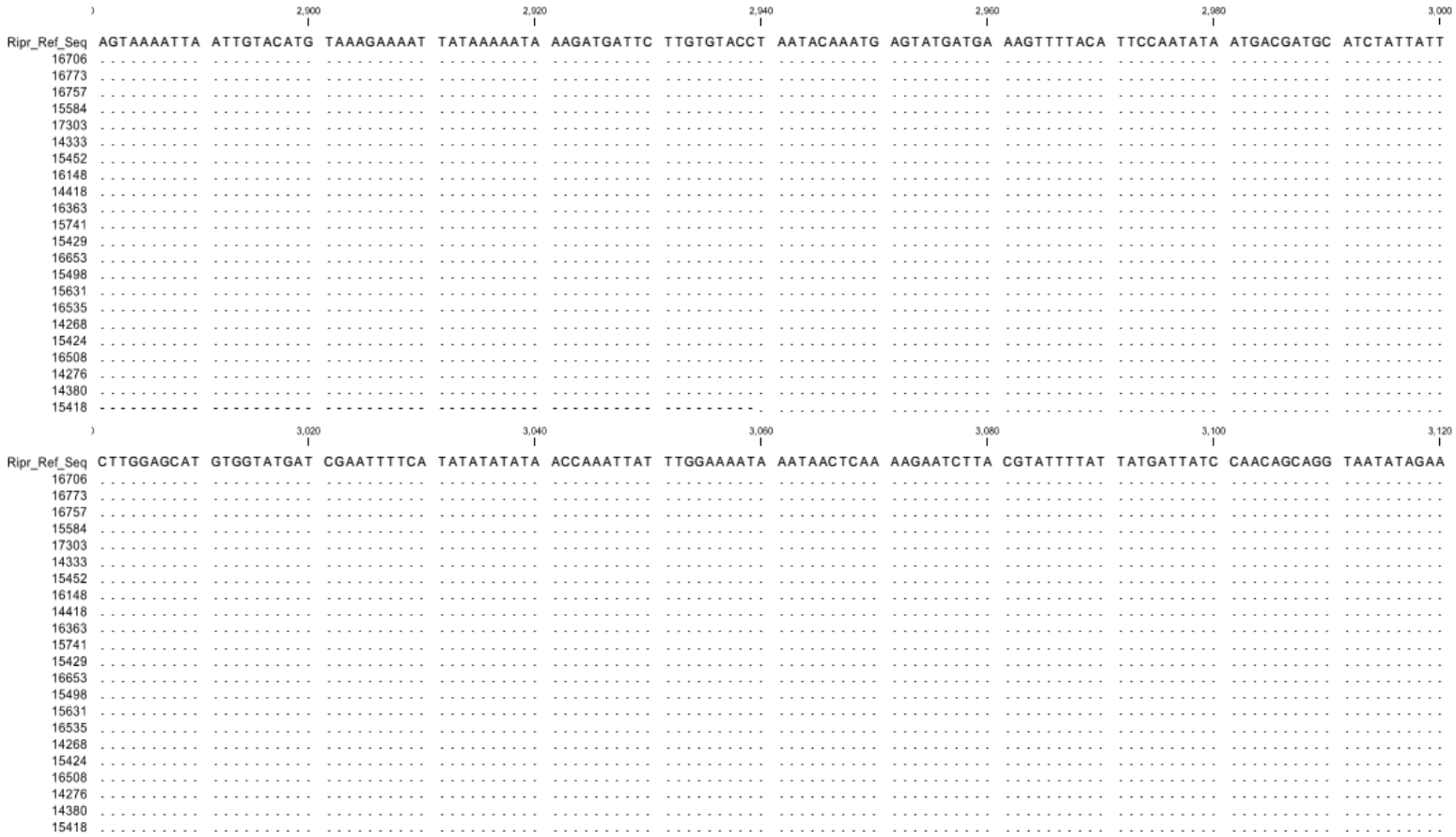
## Appendix 11. Nucleotide alignment for CyRPA Exon 2 (F3R1)



## Appendix 12. Nucleotide alignment for Ripr (F3R1)



## Appendix 13. Nucleotide alignment for Ripr (F3R1)



## Appendix 14. Nucleotide alignment for Ripr (F3R1)