LONGITUDINAL ANALYSIS OF *PLASMODIUM FALCIPARUM* MEROZOITE ANTIGENIC PROFILESAND ANTIBODY PROFILES OF CHILDREN WITH MULTIPLE MALARIA EPISODES IN KILIFI, KENYA

WACHIURI KELVIN MWANGI

I56/79526/2012

CENTER FOR BIOTECHNOLOGY AND BIOINFORMATICS,

COLLEGE OF BIOLOGICAL AND PHYSICAL SCIENCES,

THE UNIVERSITY OF NAIROBI.

Thesis submitted to the board of postgraduate studies in partial fulfilment for the award of Master of Science degree in Biotechnology of the University of Nairobi.

November 2015

DECLARATION

I declare that this is my original work and to the best of my knowledge, has never been submitted to any other University for award of a degree.

Wachiuri Kelvin Mwangi,

Registration number **I56/79526/2012**

Signature: **Date:**

With approval from my supervisors,

Professor James Ochanda,

Centre for Bioinformatics and Biotechnology (University of Nairobi),

Signature: **Date:**

Francis Ndungu, PhD,

KEMRI-Wellcome Trust Research Programme (Kilifi-Kenya),

Signature:  **Date: 19/11/2015**

Isabella Oyier, PhD,

Centre for Bioinformatics and Biotechnology (University of Nairobi),

KEMRI-Wellcome Trust Research Programme (Kilifi-Kenya),

Signature: **Date:**

ACKNOWLEDEMENTS

I thank God for the health and wealth during my research period. I also thank my family for the unwavering support during my Masters studies and while compiling this report. To my supervisors, your faith, patience and opportunity offered to do this work under your stewardship will always be sincerely appreciated. To Drs Anthony Ajua, Eunice Nduati, Yaw Bediako, Sam Kinyanjui, JamesTuju and Julian Rayner, without your esteemed contribution intellectually, morally, and materially this work would not have been of this quality. I also acknowledge my seniors and colleagues Esther Kiragu, JedidaMwacharo, Barnes Kitsao, Irene Nailaine, Ian Oyaro, Musabyimana Pierre, Ruben Mangi, MatakisoMbendo, and staff and colleagues at CEBIB for your immense support in various capacities.

DEDICATION

My family; *A beacon in the darkness…*

ABSTRACT

Plasmodium falciparum, the main cause of malaria, has been implicated as a major driver of human evolution. To counter the deadly effects of malaria, the body mounts an immune response to the parasite. However, this response does not result in sterile immunity, an observation attributed to among other factors, the subversion of host immune cells, and variability of parasites infecting the host. This study, determines whetherparasites obtained from children with multiple malaria episodes are different genetically,and antigenically, with respect to *Plasmodium falciparum*Reticulocyte binding like protein homologue 5 (PfRh5). The Rh5 gene, from a total of 422 *Plasmodium* DNAsamples takenfrom 34 children having sequential episodes of malaria,was sequenced to detect and identify polymorphisms within the infections. Contemporaneous plasma obtained from the infected children was tested for antibodies against recombinant proteins from 5 common Rh5 alleles by Enzyme Linked Immunosorbent Assay (ELISA), to determine if they would display allele-dominant reactivity. The sequencing data showed that most of theinfections contained more than one of the 5 common alleles. Consistent with other studies, the prevalence and levels of anti-PfRh5 antibodies were very low.Together,sequence and ELISA data showed that only 12 out of 92 episodes, which were allfrom variant YHK, showed allele-dominant reactivity. However, 50% of the 12 episodes were from one individual, which would warrant the use of more sensitive techniques to determine the true nature of these allele-dominant responses. Generally, this study suggests that there is little evidence of differential immune responses to the infecting genotype, thus lending weight to PfRh5 being an ideal vaccine candidate.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

1.0 INTRODUCTION

Malaria is an important disease in humans, caused by intracellular parasites of the*Plasmodium* species,(Gerald et al. 2011) such as *P. ovale, P. malariae, P. vivax, P knowlesi and P. falciparum*.Annual malaria cases stand at approximately 219 million, and mortality rates of approximately 1million cases a year, with most of these being in children living in sub-Saharan Africa(Snow et al. 2005; Gupta et al. 2013; Bartholdson et al. 2013). The deadliest of the *Plasmodium* species responsible for most of the morbidities and mortalities is *P. falciparum,*but there isalso a substantial toll from *P. vivax*infections(Derbyshire et al. 2011). Because of the associated high mortality, malaria has been a major driver of human evolution(Derbyshire et al. 2011).Clinical malaria has several manifestations,including febrile clinical malaria, placental malaria, severe malaria, severe anaemia and cerebral malaria (Gerald et al. 2011).

1.1 *Plasmodium*life cycle

Plasmodium has a complex, two host lifecycle with a sexual phase in the mosquito, and an asexual phase in the definitive vertebrate host(Good and Doolan 1999; Langhorne et al. 2008; Gerald et al. 2011). In human malaria, infection starts after an infected female *Anopheles* mosquitoinjectssporozoitesfrom her saliva into the blood during a blood meal(Cox 2010). The sporozoites thencirculate in the blood stream andinvade the liverin a phase called the pre-erythrocyticstage, which lasts for about two weeks. In the liver, sporozoites traverse through several hepatocytes where they then settle. The sporozoite then grow and multiply into thousands of merozoites(Derbyshire et al. 2011). This stage is silent in respect to any of the malaria symptoms, and in the end,merozoitesare released into the blood stream. Infections by *P. ovale* and *P. vivax* have an ability to cause relapses months or even years after the initial infection, due to their ability to hide in the liver as latent or silent hypnozoites(Derbyshire et al. 2011).

Theerythrocyticphase describes the cyclic asexual reproductive phase, during which merozoites continuously invade and develop within the Red Blood Cells (RBCs). During this phase, merozoites first released from the hepatocytes attach to, and invade the RBCs(Abu Bakar et al. 2010). They then develop via erythrocyticschizogony, where one merozoite gives rise to aschizont containing several merozoites via asexual reproduction.Schizontslyse the

erythrocyte cell when mature,and are released as free merozoites in the blood, where the RBC invasion cycle starts again(Olivieri et al. 2011). A few merozoites differentiate into male and female gametocytes, which circulate in the blood (Taylor et al. 2010)and are ingested by a female *Anopheles* mosquito during a blood meal.

The sexual phase of the *Plasmodium* lifecycle takes place within the mosquito vector. In the mosquito gut, the gametocytes develop into mature male and female gametes. They fuse to formdiploid zygotes that then develop and elongate to form motile ookinetes, which burrow into the insect gut wall. There they form oocysts on the outer lining of the gut, which then develops further,and thereafter raptures to release infectious sporozoites. The sporozoites migrate through the mosquito to the salivary glands (Cox 2010). During the next blood meal, the mosquito injects its saliva into the host, which also contains the sporozoites, completing the *Plasmodium* life cycle.A summary of the *Plasmodium falciparum* lifecycle is illustrated in figure 1.1 below.

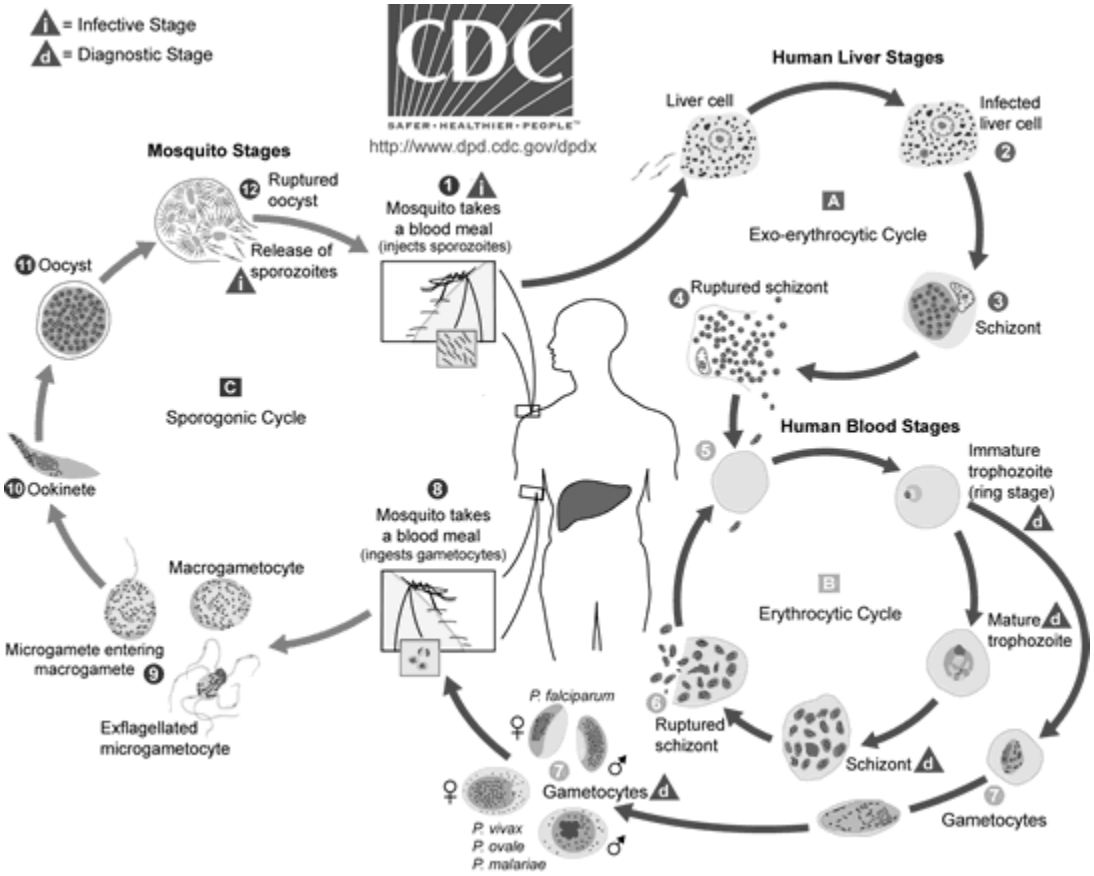


Figure 1.1 Life cycle of *Plasmodium* parasite(http://www.cdc.gov/dpdx/malaria/index.html)

1.2Erythrocyte invasion by *Plasmodium falciparum*merozoites

The merozoite, which is the smallest cell in the *Plasmodium* lifecycle(Cowman et al. 2012), is a pear shaped cell that is used by *Plasmodium* to invade the vertebrate host erythrocytes(Aikawa et al. 1978; Miller et al. 1979).Like the other two invasive forms, the sporozoite and ookinete, the merozoite contains the rhoptries, micronemes and dense granules, special organelles whose contents are thought to enable the parasite to identify, penetrate, and develop within the RBC(Iyer et al. 2007). The structure of the Plasmodium merozoite has been illustrated with more details as shown in figure 1.2 below.

.

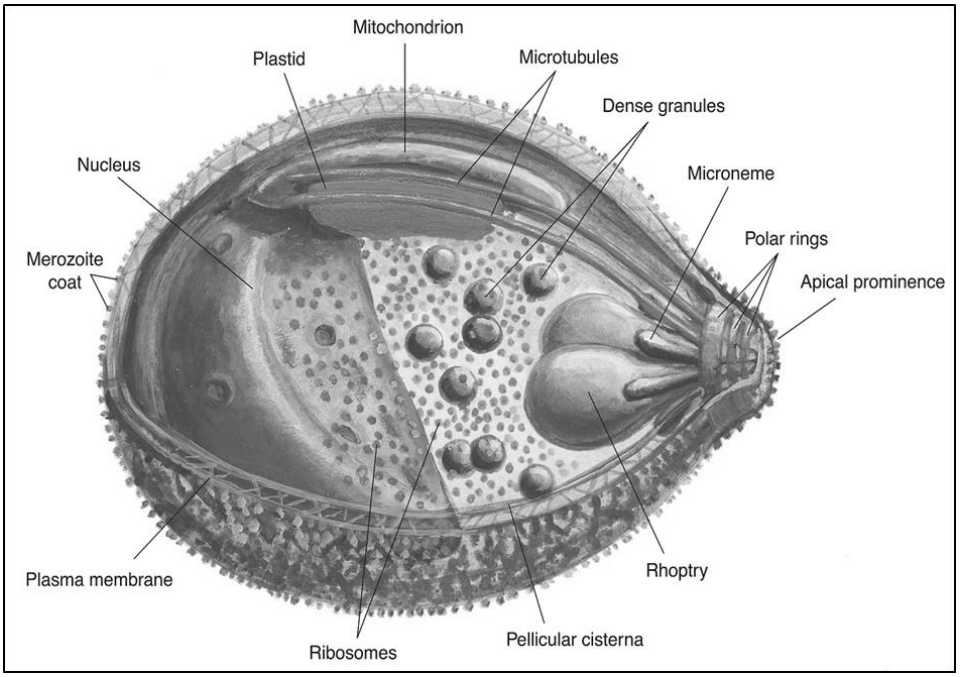


Figure 1.2 Fine structure of *Plasmodium falciparum*merozoite(Bannister et al. 2000)

The invasion is a stepwise process involving the sequential secretion of micronemal and rhoptryproteins(Singh et al. 2010). It includes the initial random interaction, deformation of the erythrocyte membrane, reorientation of the merozoite, formation of the tight junction, and finally, entry into the erythrocyte(Iyer et al. 2007).

During the invasion, merozoitesfirst form long distance, reversible, low affinity initial contacts with unparasitised RBCs(Cowman and Crabb 2006; Tham et al. 2010; Srinivasan et al. 2011). This random interaction precedes merozoitereorientation, such that their apical ends are in contact with the RBCs(Mitchell et al. 2004; Iyer et al. 2007).

A high affinity irreversible contact called the tight junction is then formed. Here, invagination occurs and the paired organelles, the micronemes and rhoptries, releasetheir contents, which are the parasite ligands. The parasite then takes advantage of its actin-myosin motor and the ligand-receptor pairs to propel itself into the vacuole, from its apex to its posterior end. The RBC membrane seals the parasite within, effectively forming a parasitophorous vacuole where the parasite lives, grows, and develops inside the RBC.

When mature, the schizonts are primed for egress as merozoites by proteolytic cleavage events, a process necessary for the next round of invasion(Child et al. 2010; Cowman et al. 2012). The exposure to low potassium concentration in the bloodstream after merozoite egress triggers the release of calcium, which activates the secretion of micronemaladhesins and invasins(Cowman et al. 2012). These then aid in the formation of the long distant, low affinity interactions with erythrocytes, and the invasion cycle proceeds as described.An illustration of the invasion process with some of the ligand-receptor pairs is shown in figure1.3 below.

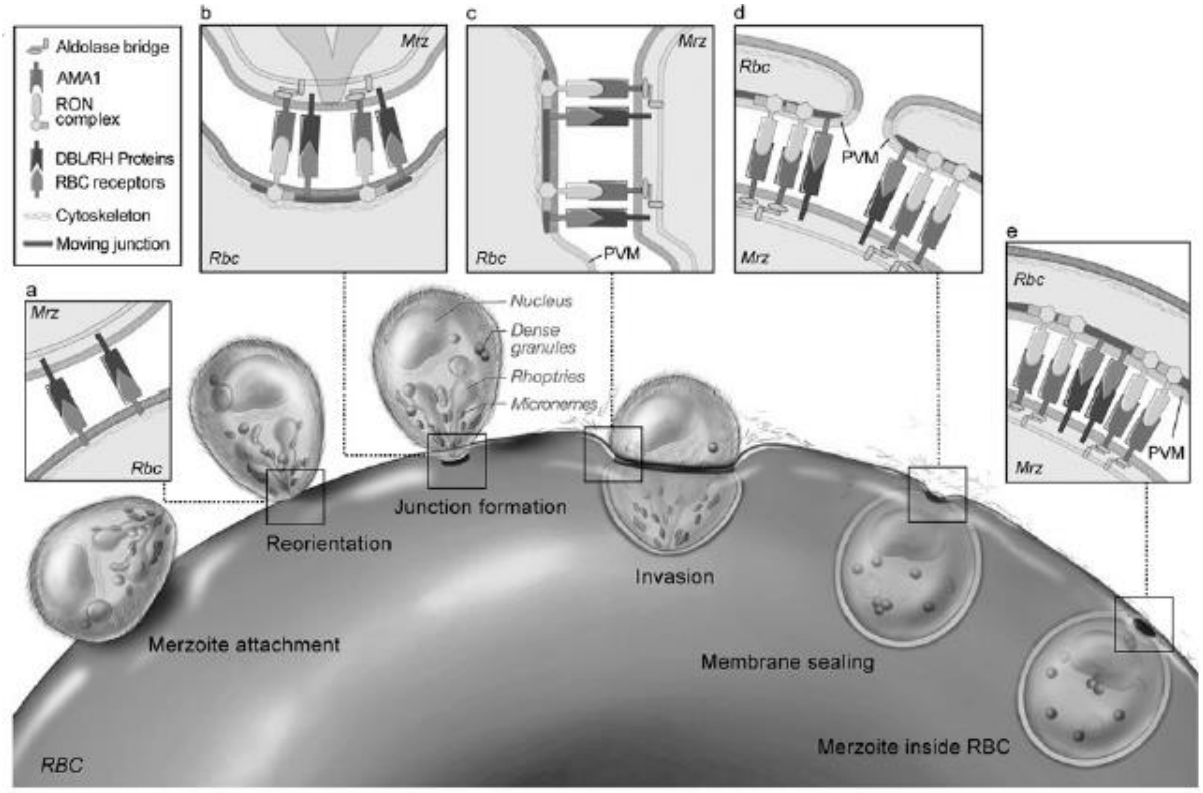


Figure 1.3 Diagram of RBC invasion(Srinivasan et al. 2011)

1.3 Merozoite ligands involved in RBC invasion

Plasmodium falciparum uses a vast array of ligand-receptor pairs to invade host erythrocytes. These ligands fall mainly in two groups; Sialic Acid Dependant (SAD) and Sialic Acid Independent (SAI) pathways (Adam et al. 2013). Ligands in these pathways are redundant, in that the parasite can use the alternative pathway should the need arise, due to reasons like host immune responses or the hosts RBCs lacking a given receptor.The ligands also fall within two super families. These are the Erythrocyte Binding Ligand (EBL) and Reticulocyte-Binding protein homologue (PfRh) families (Cowman and Crabb 2006; Iyer et al. 2007; Riglar et al. 2011).

1.3.1 The Erythrocyte Binding Ligand Super Family

Erythrocyte Binding Ligand (EBL) proteins are all orthologues of Duffy Binding Like (DBL) protein, identified in *Plasmodium vivax,* and include the Erythrocyte Binding Antigens (EBA) 175, 181, 140 and EBL1. EBA 175 was the first to be identified, and the most studied to date(Cowman and Crabb 2006; Healer et al. 2013). They are stored in the micronemes and are type 1 transmembrane proteins, having distinct regions within their structure (Tham et al. 2012). They have cysteine rich domains, Regions II and VI(RII and RVI), which are linked by regions III – V (R III – V). RII of EBA 175, -181 and -140 have a Duffy Binding Like domain with related cysteine rich repeats called F1 and F2, while EBA 175 is dimorphic in RIII to V, with either an F or Callele (Healer et al. 2013). Region VI consists of a small cysteine-rich domain,which is followedby a transmembrane domain and a small cytoplasmic tail(Tham et al. 2012). The different regions of the EBL family are shown in figure 1.4 below.

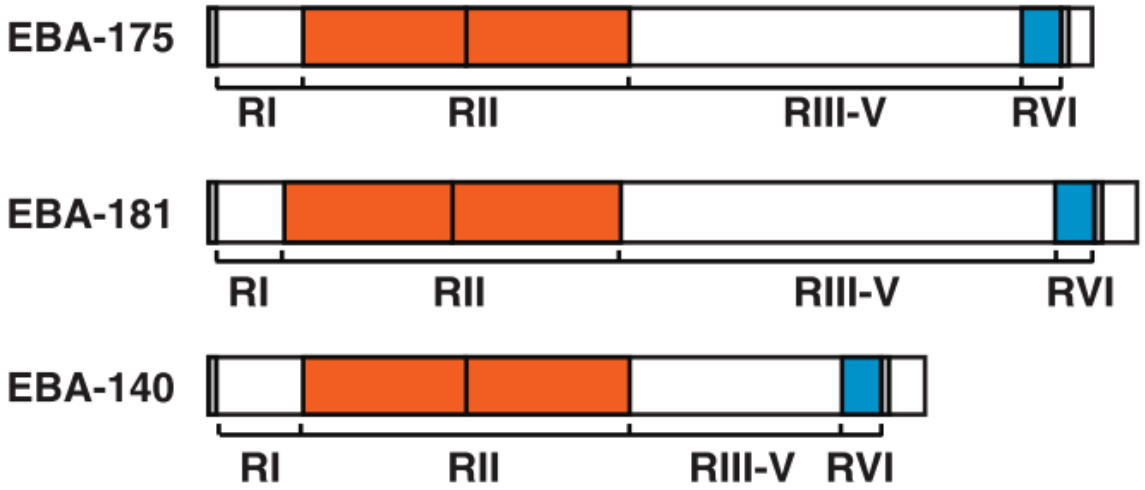


Figure 1.4 Major structures of the EBL super family (Tham et al. 2012). The red region is the receptor binding region within RII while the blue region is the cysteine rich region within RVI. RII is Region 2 while RVI is region six.

During invasion, the EBA family members use the SADpathway, utilising sialic acid residues from receptors in the RBC. EBA 175 binds to Glycophorin A (GPA)(Orlandi et al. 1992), EBL1 to Glycophorin B (GPB)(Mayer et al. 2009), while EBA 140 binds to Glycophorin C (GPC)(Lobo et al. 2003). Treatment of RBCs with neuraminidase removes sialic acid residues from its receptors hence blocking their invasion by merozoites using this pathway(Riglar et al. 2011). However, in spite of the receptor sialic acid residues being necessary for invasion, their protein backbone is also important for specificity(Cowman and Crabb 2006).

1.3.2 Reticulocyte-Binding protein homologue super family

These are large (approximately 220- 350 kDa) type I trans-membrane proteins localised to the merozoiterhoptries. They include Rh1, Rh2a, Rh2b, Rh4 and Rh5. Rh3 is a putative pseudogene on chromosome 12 that is normally transcribed, but not translated in parasite lines tested so far(Baum et al. 2009). They all show a small degree of amino acid sequence conservation (Patel et al. 2013), mainly in the N terminus.

With the exception of Rh1, which utilises the SAD pathway, all use the SAI pathway during invasion. Complement Receptor 1 (CR1) is the receptor for Rh4(Tham et al. 2010; Ord et al. 2012). All of these genes have been disrupted in at least one parasite strain, apart from Rh5,indicating both redundancy in function and strain preference for given pathways. Rh1 is a *P. vivax* Reticulocyte Binding Protein (PvRBP) 1 orthologue while Rh2a and b PvRBP 2

orthologues (Baum et al. 2009). A simple illustration of the Rh family is shown in figure 1.5 below.

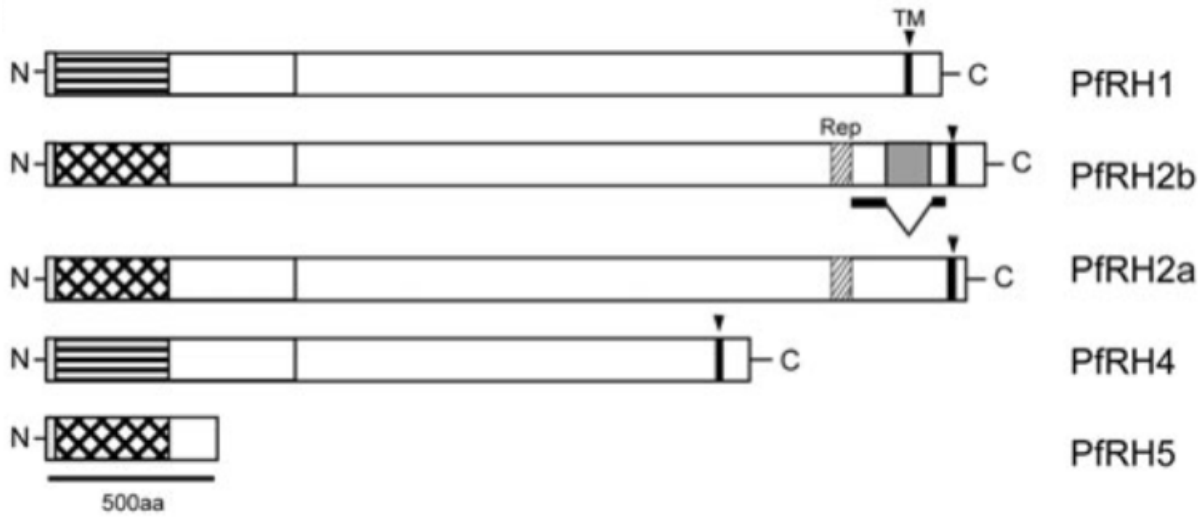


Figure 1.5 Schema of the PfRh family(Patel et al. 2013). The cross hatches show regions of close similarity between PfRh2a, 2b and Rh5 while horizontal lines show areas of similarities between PfRh1 and Rh4.

Rh2a and Rh2b are positioned head to head on chromosome 13,and only differ inthe last 500 amino acids (aa) in the C terminus of the protein, implying gene duplication (Dvorin et al. 2010; Patel et al. 2013). They were also the first in the Rh family to be identified(Rayner et al. 2000), which was then followed by Rh1, Rh4 and Rh5 in that order (Sahar et al. 2011).

1.4 *Plasmodium falciparum* Rh5 gene

Rh5 is expressed as a 63kDa protein that is processed to an approximately 45kDa protein. Unlike other members in its family, it lacks a cytosolic and transmembrane moiety. It is highly conserved across parasite lines with limited Single Nucleotide Polymorphism (SNPs), 14 of them, where only two are synonymous (Hayton et al., 2008; Bustamante et al. 2013). Only five among the non-synonymous SNPs have a frequency of more than 10% in more than one population (Bustamante et al. 2013).Hayton et al (2008) introduced two *P. falciparum*Rh5 polymorphisms from an *Aotusnancymaae* virulent strain (GB4), to an *A.nancymaae* non-virulent stain (7G8), which conferred virulence to the 7G8 strain. This showed that polymorphisms in PfRh5 are able to determine host specificity.

It is also an essential adhesinas it has not been possible to isolate viable parasites after gene knock out (Baum et al., 2009), and invades erythrocytesin a fashion that is trypsin,

chymotrypsin and neuraminidase resistant.In spite of being an essential adhesin, its actual role in invasion has not been elucidated.

Baum et al.,(2009) proposed that due to its small size and lack of both a transmembrane and cytosolic region, it could be part of a protein complex, and indeed, it has been shown to form a complex with at least two other merozoite proteins,the *P. falciparum* Rh5 Interacting Protein (PfRipr) (Chen et al. 2011), and Cysteine rich protective protein (CyRPA) (Reddy et al. 2015). PfRipr is a cysteine rich proteinthat is shed from the merozoite and has a molecular weight of 123kDa. It has 10 Epidermal Growth Factor-like domains localised to the micronemes. CyRPA is anchored to the membrane byGlycosylphosphatidylinositol (GPI) and is also localised to the micronemes. As both Rh5 and PfRipr are secreted, CyRPA anchors the Rh5/PfRipr/CyRPA complex to the *Plasmodium* surface membrane (Reddy et al. 2015).

The Rh5 receptor has been discovered to be basigin on the RBC surface(Crosnier et al 2011).Unlike all other known ligand-receptor pairs involved in invasion, this interaction is essential as its blockage by anti-basigin antibodies led to complete failure in RBC invasion (Crosnier et al 2011). Basigin, also known as CD147 is part of the immunoglobulin superfamily and has three potions, the intracellular, transmembrane, and extracellular regions, with the extracellular region containing two immunoglobulin like domains (Miyauchi et al. 1991; Liao et al. 2011). Work by (Yu et al. 2008) showed that the soluble extracellular region consisted of an N-terminal IgC2 domain anda C-terminal IgI domain, which were connected by a 5-residue flexible linker. Studies have reported that antibodies against the full length Rh5 are both cross-strain neutralising (Douglas et al. 2012) and inhibit cross-strain invasion of parasites to the RBC (Bustamante et al. 2013).

Structurally, Rh5 has been found to have a novel fold composed of 9, mainly antiparallel α-helices forming two, three helical bundles, which enclose a β-hairpin loop near the N terminus(Chen et al. 2014; Wright et al. 2014).Only two disulphide bonds are present, the first being Cys345-Cys 351, and the second Cys224-Cys317. They are important in maintaining the Rh5 structure. Work elucidating the crystal structure of PfRh5 by Wright et al, (2014) showed one free cysteine, Cys329, while that by Chen and co-workers (2014) identified one more residue, Cys203. This difference is attributable to an Y203C mutation at amino acid locus 203 in the 3D7 strain used by Chen as opposed to the 7G8 strain used by Wright, which does not have this mutation. These other cysteine molecules, Cys203 on α-helix 2 and Cys329 on α-helix 3 are free, not forming disulphide bonds, with the Cys203 side

chain only partially exposed while that of Cys329 being totally buried within the structure (Chen et al. 2014). When bound to its receptor, the Rh5 side opposite the C-terminus is in direct contact with both extracellular domains and linker region of basigin(Wright et al. 2014). Figure 1.6 below shows the structure of Rh5 and the Rh5-basigin bond.

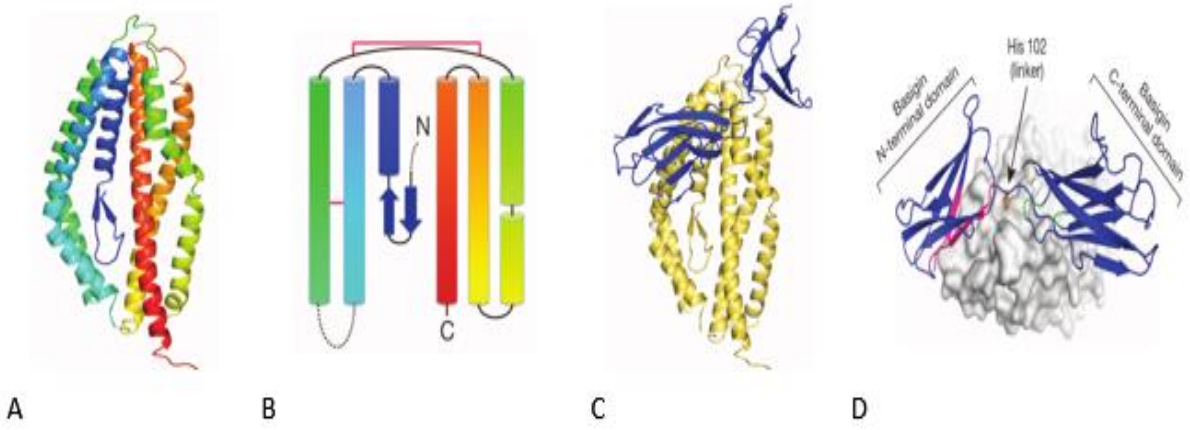


Figure 1.6 Structure of Rh5. Panel A is the structure's ribbon representation while panel B is a schematic topology representation. The red lines on panel B represent the disulphide bonds with the N and C showing the N and C terminus respectively. Panel C is the ribbon structure of Rh5 (yellow), bound to Basigin(blue). Panel D is a zoom of the Rh5-basigin binding sites(Wright et al. 2014).

Interestingly, all of the 5 common SNPs are scattered across the RH5 structure and therefore do not directly interact with the residues directly involved in Rh5-basigin binding(Wright et al. 2014). However, several SNPs identified by Hayton et al. (2008) to be associated with host tropism are within or close to the basigin binding site (Wright et al. 2014)

1.5 Single Nucleotide Polymorphisms (SNPs) and their effects.

As sequence polymorphisms can lead to changes in interactions between ligands and their receptors, it is important to describe the different sequence polymorphisms and some of their effects. There are several types of sequence variations including; Single Nucleotide Polymorphisms (SNPs), Copy Number Variations (CNVs), and insertions and deletions (indels) (Hirakawa et al. 2013). SNPs are among the most common DNA variations in genomic sequences,(Hirakawa et al. 2013)accounting for about 90% of sequence variants in the human genome,(Collins et al. 1998) and are the most abundant sequence variation in plants (Cho et al. 1999). SNPs can officially be defined as alleles in a population of normal individuals, where the least frequency has an abundance of at least 1% (Brookes 1999; Cavallo and Martin 2005). They occur whenthere is a change in the nucleotide sequence,

which leads to a change in the amino acid that is maintained in the population.This is in contrast to single point mutations, which are also changes in the nucleotide sequence which leads to a change in the amino acid but occur randomly in the sequence and have a frequency of less than 1%.

SNPs can be classified either according to where they occur relative to the gene, or with respect to the changes they cause in the amino acid sequence. In relation to the position, they can be further classified intosix types (Hirakawa et al. 2013).There are those that occur within the coding region (cSNP), those that occur in the coding region but do not cause amino acid substitution (sSNPs), those inthe intronic region (iSNP), those within 1kb of the region upstream the start codon (uSNP), those within 1kb of the region downstream the stop codon (dSNP) and those found within the intergenic regions (gSNP). All these SNPs within non coding sequences do not directly influence gene structure and so have a lower probability of affecting the phenotype unlike cSNPs(Cavallo and Martin 2005). However, they can still have an effect, as their occurrence can disrupt expression regulators such as promoter and suppressor regions by changing the sequences in these regions. This can lead to phenotypic variations due to up-regulation or down-regulation of these regulators (Hirakawa et al. 2013).

SNPs in relation to their effect on the protein sequence can further be classified into nonsense, synonymous and non-synonymous SNPs. Nonsense SNPs introduce a premature stop codon leading to a truncated protein that is normally non-functional. Synonymous SNPs (sSNPs) are mutations in the nucleotide sequence which do not cause a change of amino acids in the translated codon.

On the contrary, non-synonymous SNPs (nsSNP) are changes in the nucleotide sequence that lead to a change in the amino acid sequence. If they occur in intron-exon junctions, they can also cause changes in splice sites (Van et al. 2005)or produce truncated proteins (Hirakawa et al. 2013). When found on or near sites that the protein uses for interaction with other proteins or molecules, SNPs could lead to an alteration of the function of the protein complex (Zhao et al. 2014).As summarised by (Zhao et al. 2014), some of the changes caused by nsSNPs in the functional areas could lead to weakening, strengthening, disrupting, preserving, being beneficial, neutral or disruptive to the protein-protein interaction (PPI) as shown in figure 1.7 below.

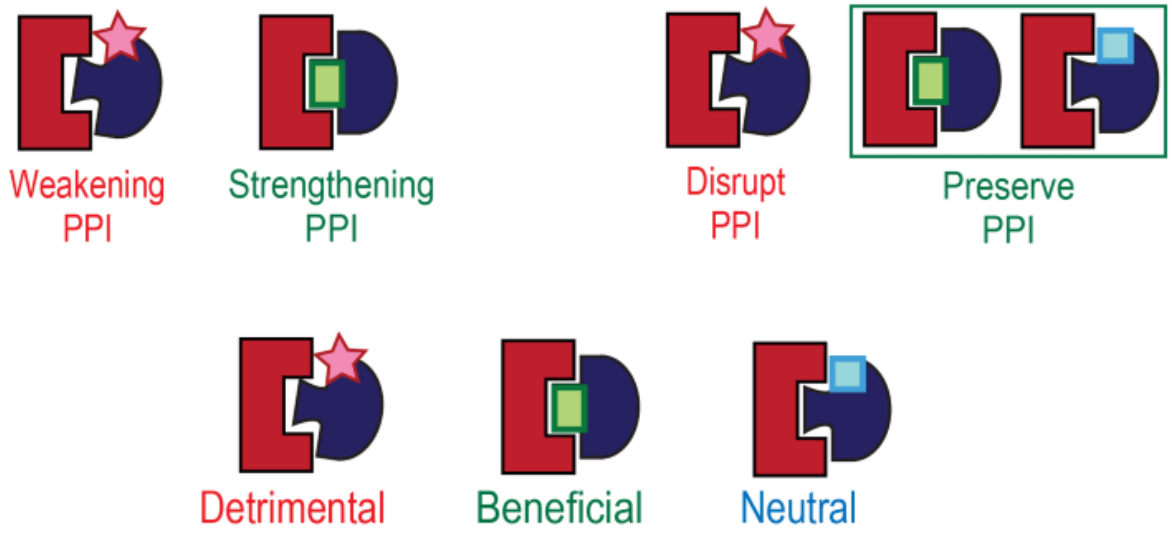


Figure 1.7 Effect of SNPs on Protein-Protein Interactions (PPI) (Zhao et al. 2014)

Some of these alterations to the gene by SNPs are now known to increase or decrease the fitness of individuals carrying them. They may predispose individuals to diseases such as diabetes (Bergholdt et al. 2012) or induce virulence of some parasites strains to new hosts. For instance, (Hayton et al. 2008) showed that 7G8, a *P. falciparum* strain originally not infective to *Aotusnancymaae* became virulent and could invade its erythrocytes after being crossed with a virulent GB4 strain, a change attributed to gaining several Rh5 SNPs from the GB8 strain.

1.6 Immunity to malaria

Malaria pathology is associated with parasitemia arising from cyclic asexual replication in the blood stage of the parasite's lifecycle (Rodriguez et al. 2008; Abu Bakar et al. 2010; Cowman et al. 2012). The body counters this by activating its immune system to fight off the infection as in other diseases. However, unlike other diseases like measles where only one or a few infections are necessary to enable long term protection against the same pathogen, natural immunity against clinical malaria is developed only after multiple exposure (Bull and Marsh 2002; Struik and Riley 2004).

Antibodies have been found to be very important in the responses to malaria infections in humans(Cohen et al. 1961a). Several mechanisms such as blocking of invasion, binding to infected RBCs (iRBC) leading to increased clearance in the spleen and antibody dependant

cellular killing, which is mediated by cytophilic antibodies (Langhorne et al. 2008; Wright and Rayner 2014). Studies done on the role of complement in antibody mediated blocking of invasion showed that antibodies were able to block invasion in the presence of complement, whereas this function was greatly hindered without complement(Boyle et al. 2015).The same study showed that antibodies could even be internalised with the merozoite during invasion in the absence of complement. This demonstrated that invasion inhibition by antibodies is dependent on the deposition of complement.

Immunity to malaria takes different forms, which seem to be dependent on the manifestation of malaria. Thus, there is immunity to severe malaria, mild malaria, and to parasites, which all appear to involve discrete mechanisms(Langhorne et al. 2008).Immune adults inoculated with blood stage parasites are protected from clinical symptoms while passive transfer of immunoglobulins from immune individuals to non-immune individuals reduces theirparasitemia(Cohen et al. 1961b; Bustamante et al. 2013). In endemic areas, infants and children immunologically naive to infections are at risk of continued illness until immunity is developed (Riley 2004). Since immunity to clinical disease increases with age due to more exposure, adults in these areas will have developed immunity to clinical disease, though it is rarely sterile,(Riley 2004; Pierce et al. 2013) as they still have parasitemia to a degree.This phenomenon where the immune systemprotects against new infections while maintaining a low and mainly asymptomatic parasitemia is referred to as premunition(Doolan et al. 2009). Figure 1.8 below illustrates the relationship between the severity of malaria and levels of parasitemia with age, in a malaria endemic region. All these show that no one response to malaria infection is enough, and that successive responses are either additive or synergistic, as proposed by Marsh and Kinyanjui (2006). This is backed up by Osier et al. (2008), who showed that the number of targets, and degree of responses by antibodies to these targets,were important predictors of protective immunity, and that even among the responses, some antibody combinations were more important than others.

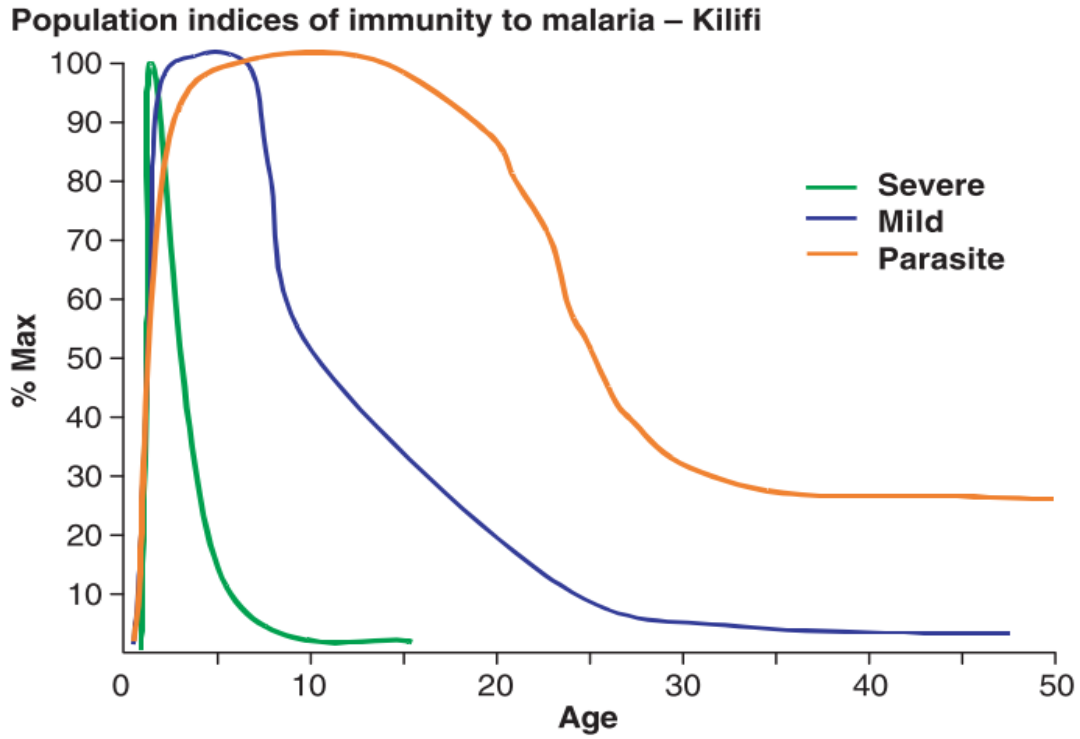


Figure 1.8; Age and malaria parameters(Marsh and Kinyanjui 2006) showing how severity of disease and parasitemia vary with age in malaria endemic regions.

The merozoite has also evolved several mechanisms to avoid destruction by the host's immune responses. Theparasite employs a vast array of redundant erythrocyte invasion ligands. Due to this surplus in ligands, which also utilise different pathways, the merozoite is able to change to another ligand in the presence of strong immune pressure. Most of these ligands are also very polymorphic, with the different variants being immunologically distinct. Thus, it may indicate that after an antibody response to a given set of antigens, parasites in subsequent infections would be variable from those already recognised by the host immune system.The immune responses are also allele-specific, meaning antibodies only recognise the alleles that led to their formation but not different alleles of the same protein. This leads to a lag in immune development as the host has to build up a repertoire of antibodies to the variant antigenic proteins during successive infections (Pierce et al. 2013). Figure 1.9 below shows some of the host immune responses and various mechanisms the merozoite uses to evade the immune responses.

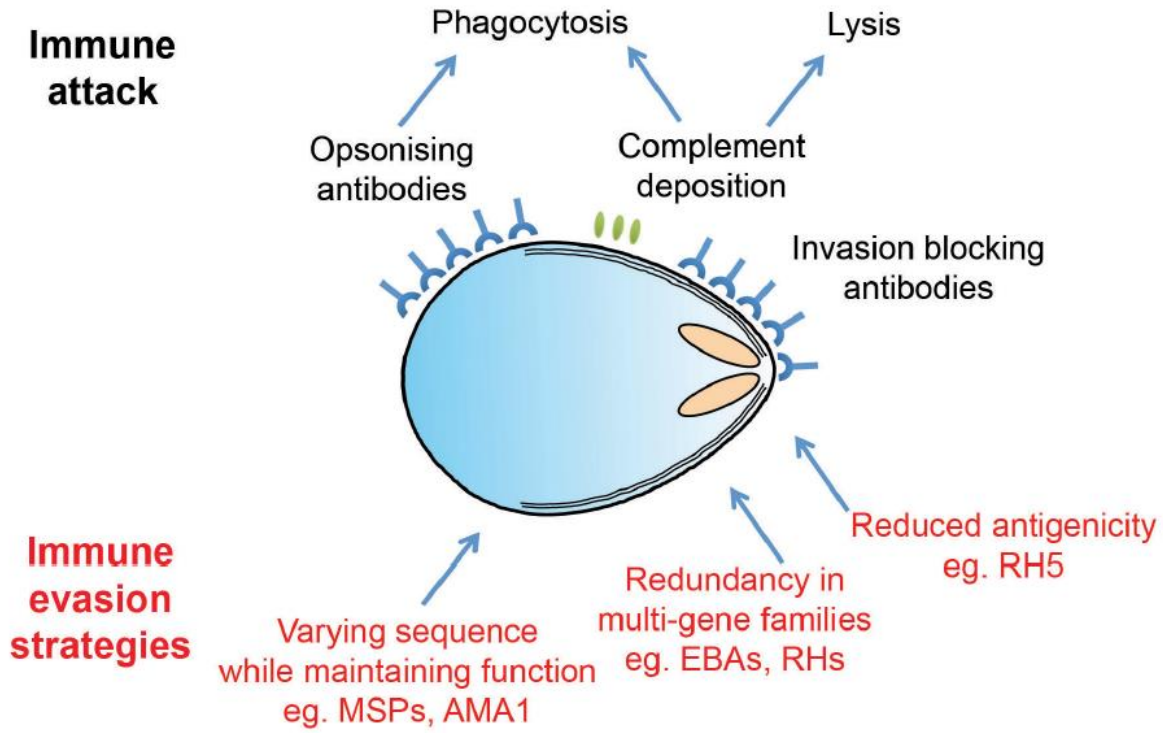


Figure1.9 Mechanisms of immune responses to, and immune evasion by the merozoite(Wright and Rayner 2014).

Also, parasites export some of their proteins to the surface of the iRBC. This has been found to cause rosseting with unparasitised RBCs, and binding with the endothelial lining of the microvesculature in internal organs such as the placenta, lungs and the brain, which may lead to placental and cerebral malaria (Beeson and Brown 2002; Pierce and Crompton 2013). This sequestration is believed to lead to placental malaria (Duffy and Fried. 1999). In areas of high endemicity where adults have already gained premunition, gravid women have an increased succeptibility to malaria occasioned by increased parasitemia and risk of severe disease and death (Mendez 2006). However, as the number of gestations increases, the host is able to control parasitemia better, as indicated by a lower parasite density (Doolan et al. 2009).*P falciparum* adheres to Chondroitin Sulphate A (CSA) on the surface of cells lining the intervillous spaceof the placenta, thus getting sequestered there. Antibodies whichrestrain the adhesion of the parasites to the CSA are found in multigravidas (women with previous pregnancies) but not primigravidas (women without previous pregnancies) (Fried et al. 1998). It is these anti-adhesion antibodies that explain why primagravidae are more susceptible to malaria than multigravidae(Fried et al. 1998).

The merozoite has generated a lot of interest as a potential vaccine target as it is not only directly exposed to the host immune system, but also elicits an immune response as evidenced by the presence of anti-merozoite protein antibodies acquired during exposure in natural infections (Osier et al. 2010; Pierce et al. 2013). Furthermore, in several cases, antibodies to merozoite proteins have been shown to inhibit invasion (Bustamante et al. 2013; Healer et al. 2013).The suggestion by Daubersies et al.,(1996) that parasites in sequential waves of infection were due to allele-specific immunity warrant an investigation to determine if this is the case. Also, in areas that are endemic to malaria, complex parasite populations are displayed in individuals who are infected. It would therefore be important to study the genetic composition of alleles in sequential infections and determine what type of immune response they produce. Of particular interest is Rh5,which has been shown to be highly conserved and produces a strain transcending immune response.

1.7RESEARCH QUESTION

Are there different *Plasmodium falciparum* Rh5polymorphic variants in multiple sequential clinical malaria infections?

1.8HYPOTHESIS

There is no difference in*Plasmodium falciparum* Rh5 polymorphic variants in sequential cases of clinical malaria infections.

1.9MAIN OBJECTIVE

Determine if multiple malaria episodes are infected by different*P. falciparum* Rh5 variants.

1.10SPECIFIC OBJECTIVES

1.Identify *Pf*Rh5 gene polymorphismsin sequentialclinical malaria infections

2. Test if sera from each sequential infection generate an allele-dominant immune response.

1.11JUSTIFICATION

Since the discovery of *Plasmodium* species as the causative agent of malaria in 1880, there has been slow progress in either eradicating, or reducing the global burden of malaria(Derbyshire et al. 2011) in Sub-Saharan Africa, South East Asia and South America. The development of an effective antimalarial vaccine has also been unsuccessful. This has been due, in part,to immune system subversion by *Plasmodium* parasitesand possible genetic variability of *Plasmodium*parasites, which may lead to antigenically distinct proteins.

Several studies have shown that PfRh5 overcomes the major challenges in malaria vaccine development by being essential for invasion and eliciting allele-transcending antibodies in-vitro and in animal studies. However, PfRh5 genetic variation and its effect to the host immune system in sequential infections is not yet known. Theaim of this work was toidentifyif there werePfRh5 gene polymorphismsin multiple sequential infections and test antibodies in sera obtained from theseinfections to determine if there wasallele-dominantreactivity to the differentvariants.This will help us understand the impact of variation in multiple infections, which will give us an insight as to how merozoite antigen variation plays out in natural infections with respect to Rh5.

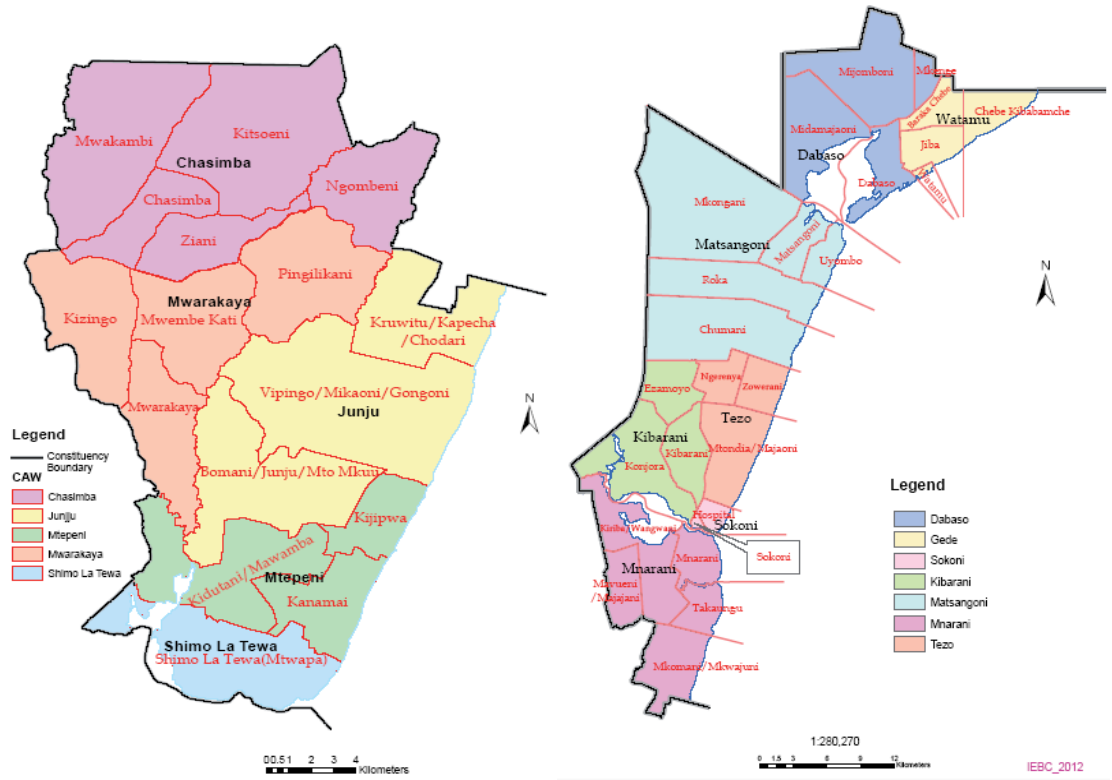
CHAPTER 2

MATERIALS AND METHODS

2.1 Study site and population

The study was conducted in Junju location, a malaria endemic region within Kilifi County.Within the Junju cohort, field workers visited homesteads once a week and measured ancillary temperatures from consenting participants. All participants with a fever (axillary temperature ≥ 37.5° C) or a history of fever were referred to the study clinic, Kilifi County Hospital, (KCH) where blood was drawn and a thick and thin blood smear done.Thick smears contain a larger volume of blood than thin smearsand were thus used for diagnosis, while thin films were used to determine the morphology and species of the infecting parasites. Parasites and plasma were then isolated from the blood. For this study, thetarget group was composed of 34 children aged between 0 and 5 years who had visited KCH with over 8 sequential episodes of malaria.8 episodes were chosen as they were the minimum number of episodes that would give a large enough sample size for the study A total of 422 DNA samples from 34 individualswere obtained for this study. However, only 397plasma samples from 33 of the same individuals were obtained instead of 422 samples as the other 25 plasma samples were not found among the archived samples. Figure 2.1 below shows the Junju and Ngerenya study locations within KilifiCounty.

A



B C

Figure 2.1 Maps showing the study areas. Panel A shows the location of KilifiCounty within Kenya while panels B shows Junju, which is located within Kilifi South sub-county. Panel C shows Ngerenya which is found within Tezo ward within Kilifi North Sub-county. (IEBC 2012; NordNordWest 2015)

Laboratory work was conducted both at the Centre for Biotechnology and Bioinformatics (CEBIB) laboratories, University of Nairobi, and KEMRI-Wellcome Trust laboratories in Kilifi.Ethical clearance for this study had already been obtained from the KEMRI ethics review board for integrated studies of natural immunity to malaria.

2.2 *Plasmodium falciparum* Rh5 gene amplification and cleaning

2.2.1 DNA amplification

The Rh5 gene was amplified from the 422 DNA samples obtained. Pre-designed primers (forward primer CGAAGAATCAAGAAAATAATC and reverse primer CCATGTTTTGTCATTTCATTG) were used to amplify by Polymerase Chain Reaction (PCR) the full Rh5 gene sequence, which is approximately 1800 base pairs (bp). The Expand high fidelity taq system from Roche was used for this study.For each PCR reaction, two master mixes,each of 5µl, were prepared as follows. In an Eppendorf tube, the Taq mix contained 0.13µl of Taq and 1µl of buffer 2 (Roche) containing magnesium chloride. 3.87µl

of PCR grade water was then used to top-up this mix to 5µl. To a separate Eppendorf, a template mix of total volume 5µl was prepared by adding 1µl of buffer 4 (Roche), 0.2µl of dNTPs (Bioline), 0.3µl each of forward and reverse primers and 2.7µl of PCR grade water to top-up the mix to 5µl. To run the PCR, 5µl of the Taq mix was added to a PCR tube containing 4.5µl of the template mix and 0.5µl of a DNA sample. The tubes were placed on an MJ Research PTC-100 heated lid thermocycler and the running conditions set with the following steps; 94°C for 2 min, 94°C for 15 seconds (sec), 50°C for 30 sec, 72°C for 2min, go to step 2 nine times, 94°C for 15 sec, 52°C for 30 sec, 72°C for 2 min, plus 5 sec per cycle, go to step six 24 times, 72°C for 7 min, 4°C for 5 min 12, end of program. The amplicons were then run on a 1% agarose gel to confirm amplification.

One percent agarose gels were prepared by dissolving 0.3g of agarose in 30ml of Tris Boric EDTA (TBE) buffer in a conical flask. The agarose was dissolved in the buffer by boiling, then let to cool. Before gelation, 2µl of Ethidium Bromide was added to aid in visualisation of the DNA amplicons. The dissolved agarose was then poured on a casting rack with combs and left to set. Afterwards, the racks were placed in an electrophoresis chamber where running buffer (1X TBE) was added, then the combs removed carefully. 2µl of the amplicon was added to 2µl of loading dye and mixed,then loaded to the wells in the gel. A 1 kilobase (kb) molecular ladder was loaded on one well and the gel run with a power pac (BioRad) at 80 volts, 400 milli-amperes for 40 minutes. The gel was then visualised on a Gel Doc™ XR+ imaging system (BioRad) and the image captured.

2.2.2 PCR amplicon clean-up

The amplicons were transferred into 96 well plates and cleaned using Ethanol purification. Cleaning was done to remove all unused reagents from the amplicons. 90µl of cleaning solution composed of 3µl of 3M sodium acetate (Ambicon), 62.5µl absolute ethanol (Sigma-Aldrich), and 24.5µl distilled water was used for each reaction. The plates were sealed, incubated at -20°C for 30min, then spun at 3000 x g for 30 min at 4°C. The seals were removed, paper towels placed on top and the plates inverted to drain them. The wet towels were discarded then fresh paper towels added and the plates centrifuged at 50 x g for 1min at 4°C. 150µl of ice cold 70% ethanol was then added to each well. The plates were sealed and spun at 3000 x g for 10min at 4°C. The plates were then inverted over paper towels and centrifuged at 50 x g at 4°C for 1min to drain excess fluid the plates were then left on the bench to air dry. 30µl of EB buffer (Qiagen) was added to each well to resuspend the purified DNA amplicons.

2.3 Amplicon sequencing

Sequencing was done on an Applied Biosystems 3130xl platform using Sanger chain termination chemistry. The 3130xl produces reads of approximately 800bp (Schatz et al. 2010). Since the Rh5 gene is approximately 1800 bp, several primers were used to cover the whole region as demonstrated in Figure 2.2 below.

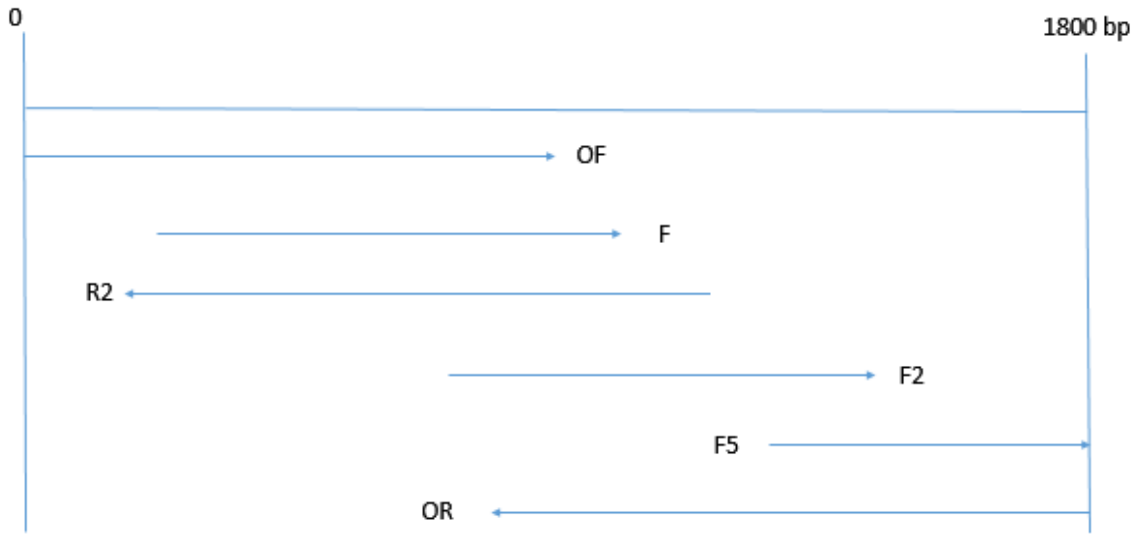


Figure 2.2 Rh5 sequencing strategy showing the 6 sequencing primers used and their coverage relative to the full gene. The forward primers used were OF, F, F2 and F5 while the reverse primers used were OR and R2.

Six primers were used and they were OF, F, F2, F5, R2 and OR (Figure 2.2 above and Table 2.1 below). The arrows show primer orientation relative to the gene with primers OF, F, F2 and F5 having a forward orientation while R2 and OR had a reverse orientation. A unique number was given to each sample and this number fed into the sequencer software. This was to help identify each sample since the six different primers were run separately.

Table 2.1The list of primers used and their sequences

Primer name	Primer sequence
OF	5-CGAAGAATCAAGAAAATAATC-3
F	5-CGAAGAATCAAGAAAATAATCTG-3
F2	5-CATAAGTCCTCTACATATGGA-3
F5	5-AGACATGTCAAATGAATATTC-3
R2	5-TCTTCGGTTTCATCATCTGT-3
OR	5-CCATGTTTTGTCATTTCATTG- 3

For each sequencing reaction using the BigDye terminator v3.1 cycle sequencing kit of a particular sequencing primer, Big-Dye PCR was first done. Here, 0.4µl of BigDye was added to 1.75µl of 5X sequencing buffer, 0.32µl of 10µM primer and 4.53µl of PCR grade water. For each reaction, 3µl of cleaned DNA amplicon was added to make a 10µl reaction volume. The BigDye PCR conditions were set at 96°C for 3sec, 50°C for 15sec and 60°Cfor 4min and this cycle was repeated 25 times. The plates were cleaned as described in section 2.2.2 but with a slight modification in the last step. After air drying, instead of adding EB buffer to the wells, 10µl of HiDi was added to each well and the product denatured by heating the plates in a thermocycler at 96°C for 3 min. The plates were then loaded onto the Applied Biosystems 3130xl genetic analyser for sequencing. The output was loaded into Sequencing Analysis v5.2 (Applied Biosystems) software for base calling after which the sequences were analysed.

2.3.1 Sequence Analysis

All chromatograms and the *P. falciparum* Rh5 3D7 reference sequence PF3D7_0424100 downloaded from PlasmoDB (www.plasmodb.org) were exported to Sequencher software version 5.0.1 for analysis. Apart from the reference, all sequences were trimmed to remove low quality bases using the default parameters set on the software. The cleaned sequence reads from each of the six primers were then grouped by their sample names. The assembly algorithm was set to an 85% minimum match percentage and 20 bases minimum overlap between the reads within a sample then assembled with respect to the reference sequence.

As much as the sequencing strategy planned to cover the full length of the gene for each sample, most contigs had gaps within the full length sequence. Reasons for this included poor

sequencing PCR amplification where a sample was not amplified or poor sequencing where poor or no data was generated for base calling.

Each of the resultant contigs were then manually checked to ensure the correct bases were called to generate the consensus sequence. Ambiguities in the consensus were resolved by referring to the chromatogram files of the respective reads in that consensus sequence.

After correcting all ambiguities, both the consensus and reference sequences were translated. This was to detect if any of the Single Nucleotide Polymorphisms (SNP) caused a change in the amino acid translated from the codon. A report on the number, positions and amino acid change was then made from the software and exported in excel format.

2.4 ENZYME LINKED IMMUNOSORBENT ASSAY (ELISA)

ELISAs against 5 full length recombinant Rh5 variants wererun on each of the 397 plasma samples, to help determine if there was antigenic variation among them. The 7G8 strain, used as our reference sequence had the amino acids Y, H, Y and K at positions 147, 148, 203 and 429 respectively, and was labelled as Variant 1 for this study. The other 4 variants had amino acid mutations at either of the positions 147, 148, and 429 respectively. The amino acid mutations from the reference strain, and their labels as used in this study are shown in table 2.2 below. These variants were obtained from Dr Julian Rayner's laboratory in the Wellcome Trust Sanger Institute, UK.

Table 2.2 The 5 Rh5 variants showing their mutations and labels as used in this study.

Rh5 variants	Label
YHK	Variant 1
HDN	Variant 2
HDK	Variant 3
YHN	Variant 4
YDN	Variant 5

The 5 variants andpooled hyper immune sera (pHIS)were serially diluted and ELISAs run to determine their titre values,which were determined to be 1:2 for variants HDN and YHN while that for variants YHK, HDK and YDN was 1:4. The plates were also incubated for 10, 15, 20, 25 and 30 minutes to determine which incubation time was appropriate. 20min incubation time was picked as it had the highest OD readings without causing errors when

read by the ELISA plate reader due to alkaline phosphatase product saturation in the plate wells.

The concentrations of the 5 variants were measured by Nanodrop and they were as follows; variant YHK, 7.98 mg/ml, HDN, 5.60 mg/ml, HDK, 8.37 mg/ml, YHN, 5.67 mg/ml, YDN, 8.28 mg/ml. The pHIS was used at a dilution of 1:1000 while standards for each plate were made by making two-fold serial dilutions of the pHIS from 1:1 to 1:12800. All dilutions were made using 1%BSA (Sigma) in 1X PBS.

Plasma from twenty Swedish adults and twenty Ngerenyachildren were also included as negative controls.TheSwedish adults were considered to have never experienced malaria in their lifetime, thus had no malaria specific immune response. On the other hand, the Ngerenyachildrenwere from a low malaria transmission area, thus experienced few infections and were expected to have low malaria specific immune responses,and could be used as negative controls.

Each well of a streptavidin coated 96 well plate (NUNC) was blocked with 100µl of 1%BSA-1X PBS for one hour then washed 3 times with 300µl of 0.1% Tween 20 in 1X PBS in an ELX 405 ELISA plate washer (Bio Tek) then rinsed with 200µl of 1X PBS. The variants were then diluted in 1% BSA-1X PBS and 100µl used to coat the plate for 1h at room temperature with gentle shaking. Due to a shortage of the streptavidin coated plates and the large number of samples, each sample was run as a singlicate instead of duplicate or triplicate. Afterwards, the plates were washed as described in section 2.4 and 100µl of serum from the children added at a dilution of 1:1000 with 1% BSA in 1XPBS. In each plate, a blank, plasmas from four Swedish and four Ngerenya individuals, a pHIS positive control and a serially diluted pHIS standard were also included. The arrangement for plasma samples, positive, negative, blank and standard samples are shown in figure 2.3 below. The plates were incubated for 1h at room temperature with gentle shaking then washed as previously described. 100µl of alkaline phosphatase goat anti-human IgG (Sigma) diluted at 1:2000 with 1%BSA in 1XPBS was then added. The plates were then incubated for 1h at room temperature with gentle shaking and afterwards washed as described. The substrate buffer was prepared by diluting 5mg of Para-Nitrophenylphosphate (pNPP) (Sigma), the alkaline phosphatase substrate, in 5ml of 1X Diethanolamine (Zymed laboratories Inc.). 100µl of this buffer was then added to each well and incubated at room temperature for 20min in the dark.

	1	2	3	4	5	6	7	8	9	10	11	12
A	Sample	Sample	Sample	Sample	Sample	Sample	Sample	Sample	Sample	Sample	Eu	
B	Sample	Sample	Sample	Sample	Sample	Sample	Sample	Sample	Sample	Sample	Eu	
C	Sample	Sample	Sample	Sample	Sample	Sample	Sample	Sample	Sample	Sample	Ng	
D	Sample	Sample	Sample	Sample	Sample	Sample	Sample	Sample	Sample	Sample	Ng	
E	Sample	Sample	Sample	Sample	Sample	Sample	Sample	Sample	Sample	Sample	Ng	
F	Sample	Sample	Sample	Sample	Sample	Sample	Sample	Sample	Sample	Sample	Ng	
G	Sample	Sample	Sample	Sample	Sample	Sample	Sample	Sample	Sample	Eu	POSITIVE	
H	Sample	Sample	Sample	Sample	Sample	Sample	Sample	Sample	Sample	Eu	BLANK	

Figure 2.3 Plasma schematic for ELISA plate where cells labelled sample had patient plasmas, those labelled Eu had a unique plasma from European adults, Ng had a unique plasma from a Ngerenya child, POSITIVE had the pHIS as the positive control and BLANK did not have any plasma. Column 12 was the standard, which had been serially diluted from the pHIS.

Plates were then read at 405nm using a Synergy 4 ELISA plate reader (Bio Tek) using the Gen 5 v1.05 software. The Optical Densities (OD) obtained were then exported in Excel format for analysis.

2.5 ELISA analysis

Plasma from the two naïve populations of European adults and Ngerenya children were run alongside each other to determine which would be used as the negative control. As the goal was to determine allele-dominant immune responses, cut offs for each of the 5 variants were calculated independently. The cut off values were determined by calculating the mean ODs of the 20 naïve individuals for each Rh5 variant and adding 3 standard deviations (SD).

Cut off= mean OD of negative controls + 3 standard deviations

The negative control represents individuals who are deemed not to have encountered any malaria hence have no malaria specific responses. Thus, by selecting all episodes above the cut off, the formula could be used to segregate infections that had PfRh5 specific responses from background responses.

CHAPTER 3

RESULTS

3.1 DNA amplification by PCR

PCR amplification was done on all the 422 samples received from the 34 children with multiple malaria episodes. Of these, PCR was successful in 317 samples. Amplification success was verified by running 1% agarose gels then visualizing and taking images of the gels from a gel doc as shown in figure 3.1.

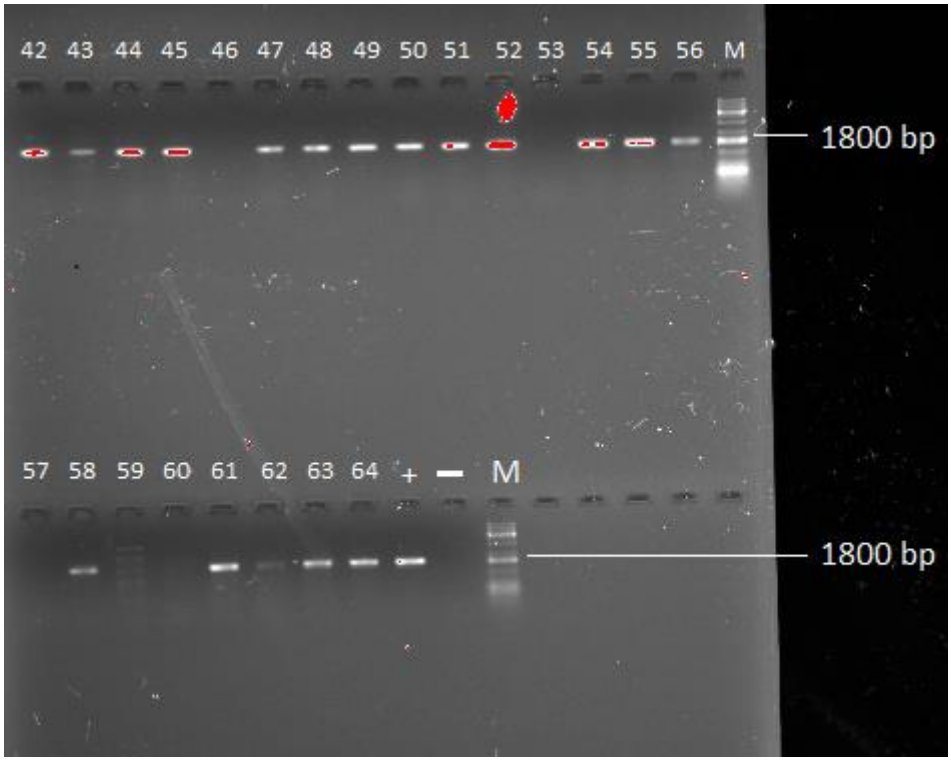


Figure 3.1 Gel image of Rh5 amplification.Lane numbers denote individual samples, + denotes positive control, - denotes negative control, M denotes the ladder and bp stands for base pairs. The banding patterns on the gel shows the amplicons are approximately 1800 base pairs.

3.2 Chain termination sequencing

The 317 successful amplicons were then sequenced by chain termination chemistry. Of these, 302 contigs were obtained after cleaning the individual sequences, assembling them and manually confirming the bases called. These samples were then checked against the ELISA data and it was determined that 269 samples had both sequence and ELISA data. The 269 samples were then checked for presence of SNPs. The respective SNPs, SNP positions and amino acids on the reference and samples are shown in Table 3.1 below.

Table 3.1.The Rh5 SNPs detected in the population of children with multiple malaria episodes.

Codon position	7G8 reference amino acid	Non-reference amino acid	SNP frequency %	% mixed
147	Y	H	11.3	49.3
148	H	D	11.9	50
203	Y	C	5.0	2.3
429	K	N	17.2	0.3

The codon positions are the amino acid positions in the sequence. The total number of samples used to obtain the SNP frequency is 302. The percentage of samples with both amino acids at each codon position (mixed amino acids) are also shown.

The 302 samples were then analysed to find the 5 variants in each malaria episode. Due to incomplete coverage during sequencing, only 100 samples had sequence data from codon positions 147, 148 and 429, which were needed to group samples into the 5 variants. When the 100 samples were grouped into their respective variants, YHK was the most abundant.As shown on table 3.2 below, there were also many episodes that carried mixed variants.

Table 3.2. Frequencies of the Rh5 variants.

Variant	YHK	HDN	HDK	YHN	YDK	YHK& HDK mixed	HDN& YHN mixed	YHK, HDN, HDK& YHN mixed	Others
Percent (n=100)	23	7	4	6	0	26	30	1	3

The frequencies of mixed variants in an episode are also shown. YDK was not detected. The column under others show any variant that was not covered by the 5 common variants.

3.3 ELISA results

Of the 422 DNA samples present, 397 had corresponding plasma samples. ELISAs were thus performed on these 397 plasma samples. The naïve plasmasamples from malaria naïve European adults and Ngerenya children were run on each plate and their ODs compared. This was to determine which of the two would be used as a control for background non-malarial reactivity. As shown on figure 3.3 below, there were no significant differences in their mean ODs to the respective variants except for YDK where the mean of ODs from Ngerenya children was significantly higher than those from the European adults. The p-values for Mann Whitney tests between European and Ngerenya plasma samples for each variant were; 0.0778, 0.392, 0.522, 0.8044 and 0.0475 for alleles YHK, HDN, HDK, YHN and YDK respectively.

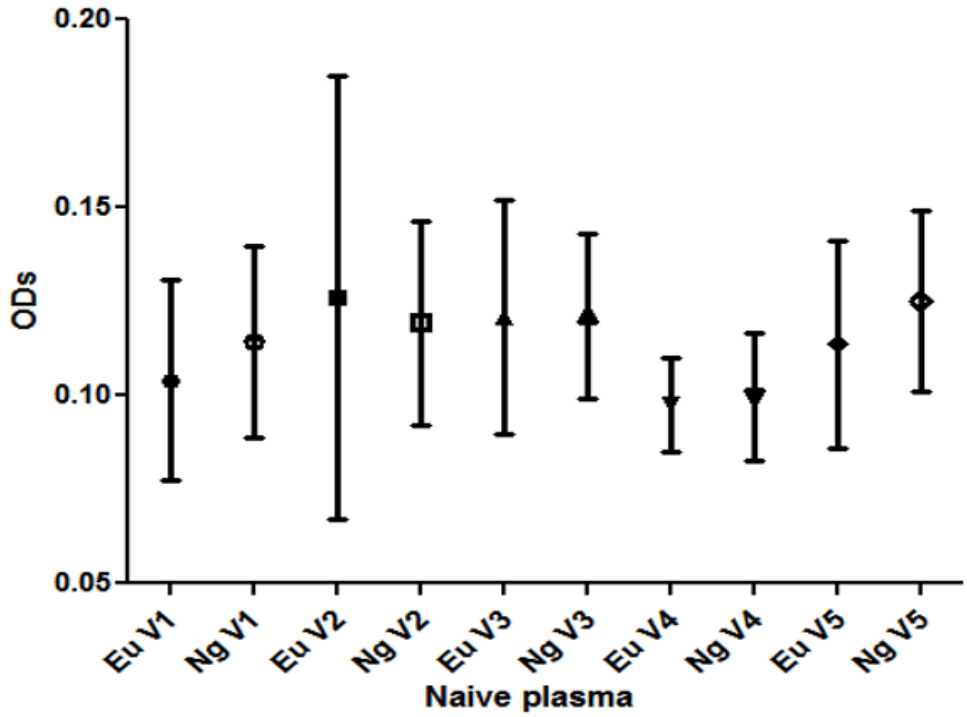


Figure 3.2 ODs against the European and Ngerenyanaive plasma samples. Where Eu represents European, Ng -Ngerenya, V1 - YHK (circle symbol), V2 - HDN (square), V3 - HDK (triangle), V4 - YHN (inverted triangle) and V5 - YDK (diamond) while ODs are the Optical density values measured at 405 nm. The symbols represent the mean while the bars are the standard deviations of the 20 individuals for each variant.

Overall, ODs from European adults were lower than ODs from Ngerenyachildren as shown on figure 3.3above and were thus used for calculation of variant specific cut offs. This trend was expected as the Europeans were not from a malaria endemic region as compared to the Ngerenya children.

Table 3.3Mean ODs, standard deviation and variant specific cut offs, each calculated from a total of 20 plasma samples from naïve individuals.

	YHK	HDN	HDK	YHN	YDK
Eu mean OD	0.104	0.127	0.12	0.098	0.114
SD	0.026	0.059	0.031	0.031	0.028
3X SD	0.078	0.176	0.093	0.038	0.083
Cut-off	0.182	0.303	0.213	0.135	0.197

The mean, standard deviation (SD) and cut-off were calculated separately for each variant as shown. Eu denotes European, OD optical density and SD standard deviation.

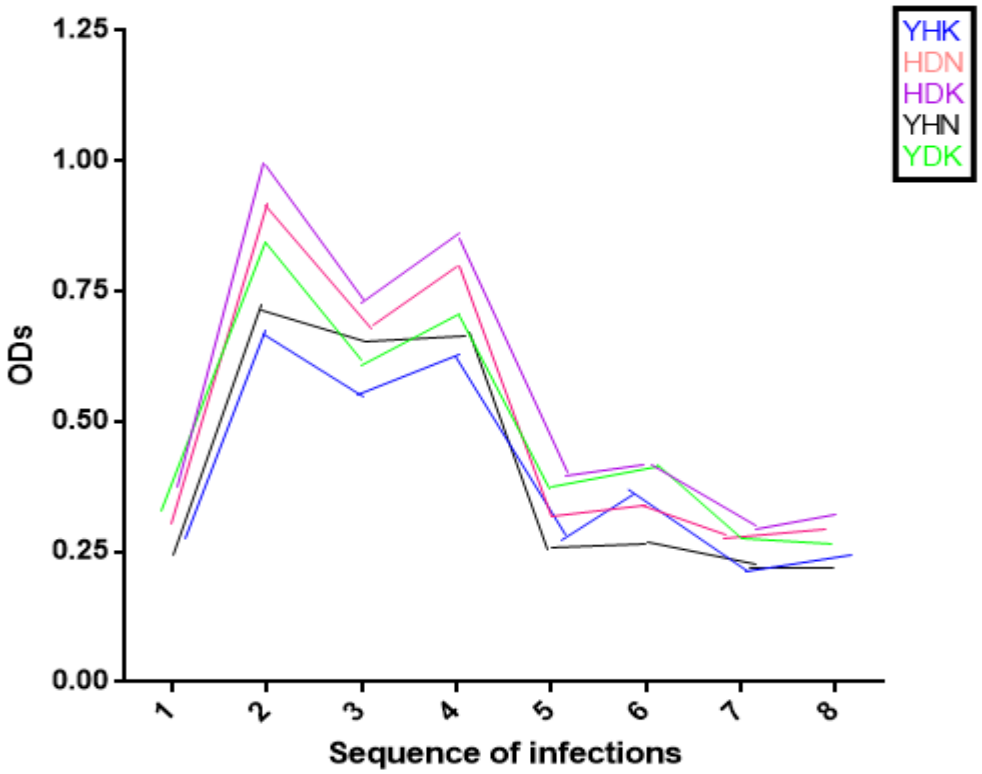
The formulaCut-off= Mean + 3SD was used to calculate the variant specific cut-offs

Using the variant specific cut-offs, as shown in table 3.3 above, all ODs were checked to determine the response by all the plasma samples tested to the 5 variants. Those above the cut off were designated as positive or responders while those that were equal or below were termed as non-responders or negative. YHN had the highest number of responders among the 5 variants with 108 positive episodes out of the possible 397 episodes. This was followed by variant YHK, YDK, HDK and HDN in that order. Table 3.4 below shows the frequencies of the positive responders to the respective variants. A sample of the trend of responses to each variant for two individuals has been shown in figure 3.3 below but the full list for all the 32 individuals in the study has been shown inappendix 1. Of interest, there was variability in the responses within an episode between the different variants.

Table 3.4Total numberof responders per variant.

Variant	YHN	YHK	YDK	HDK	HDN
Total responders	108	93	67	52	16
% responders	27.2	23.4	16.9	13.1	4.03

A total of 397 samples were run for each variant



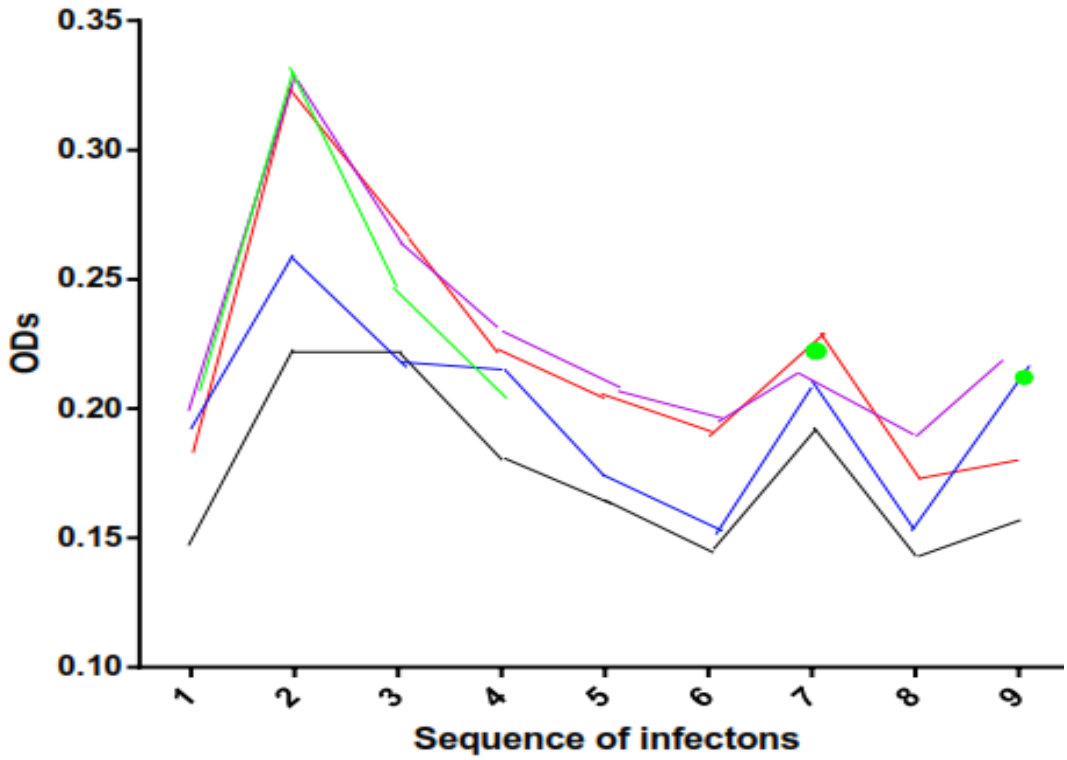


Figure 3.3 Responses to the variants by two individuals. The lines mean there was a response to the given variant between the two subsequent infections joined by the line. Dots mean that a response was only present in only that episode and not subsequent episodes.

3.4 Comparison between ELISA and sequence data

We wanted to find out if a change in the genetic composition of the infecting parasites from one infection to the subsequent ones would lead to a change in the antibody profile of the patient. There were 269 episodes that had both of their ELISA and sequence data available. Of these 269, 92 episodes were ELISA positive to at least 1 of the 5 variants and part of this result is shown in figure 3.4 below. Results for all of the 92 episodes are given in appendix 2.

Patient	Sequence variant	Episode	YHK	HDN	HDK	YHN	YDK
1	HDN & YHN	E					
2	YHK	A					
	YHK & HDK	B					
	HDN & YHN	C					
	HDN & YHN	D					
	YHK & HDK	E					
	YHK & HDK	H					
	YHK	I					
	YHK & HDK	J					
3	YHK	A					
	HDK	F					
	YHK & HDK	G					
	YHK	I					
	YHK	J					
4	YHK & HDK	B					
	YHK & HDK	C					
	YHK & HDK	D					
5	YHK & HDK	B					
	YHK & HDK	E					
6	HDN	A					
	YHK	B					
	YHK & HDK	C					
	YHK & HDK	E					
7	YHK	C					
	YHK	D					
	YHK & HDK	H					
	YHK	I					
	YHK	J					
8	HDK	A					
	YHK & HDK	B					
	HDK	C					
9	HDK	A					
	HDN & YHN	G					
10	YHK	H					
11	YHK & HDK	F					
12	YHK & HDK	B					
	YHK & HDK	D					
	YHK	E					
14	YHK	A					
	HDK	B					

Figure 3.4 Comparisons of ELISA and sequence results. Individual children are shown under the patient column, the parasite variant from sequence data is shown under the sequence variant while the sequential infections for each child are shown in alphabetical order from the first to last under the episode column. The orange boxes indicate positive responses to the particular variant while the white boxes indicate that the responses were below the cut off for the given variant data.

Consistent with high numbers of mixed variants from the sequence data, when the sequence variants were compared to the ELISA data, most infections elicited mixed responses. Of the 92 episodes, only 12 showed a dominant response to the infecting allele aloneand these were all from variant YHK.A dominant response is where the immune response only occurs to the

genotype variant of the infecting parasite. For example, from figure 3.4 above, patient 14 had a dominant response during episode “A” since the infecting parasite sequence variant was YHK and the immune response was only to variant YHK. There were a few other episodes where the infecting parasites were mono-allelic (HDN, 2 episodes and HDK, 8 episodes) but in all these episodes, ELISA responses were to multiple alleles,an example being patient 9 in episode “A” from figure 3.4 above. It is interesting to note that as much as there were few episodes with allele specific responses, all Rh5 variants elicited responses that mostly recognized other variants. For instance, for parasites with variant YHK, only 12 were dominant to YHK while in the other mixed responses, most were to variant YHN. For infecting parasite variant HDN, all responses were mixed but with YHK being recognized the most, while for infecting parasite variant HDK, most responses were mixed but YHK being the mostrecognized.

After observing that individual parasite variants mainly elicited allele transcending responses (87% allele transcending responses), all parasite variants and their responses from the 92 episodes were compared. This was to help determine if there was a trend between the parasite variant in an infection and the response it induces in the host.This comparison using Pearson correlation (r=-0.0037) showed no correlation between the infecting parasite and immune response induced in the host as shown in figure 3.5 below for all the 92 episodes.

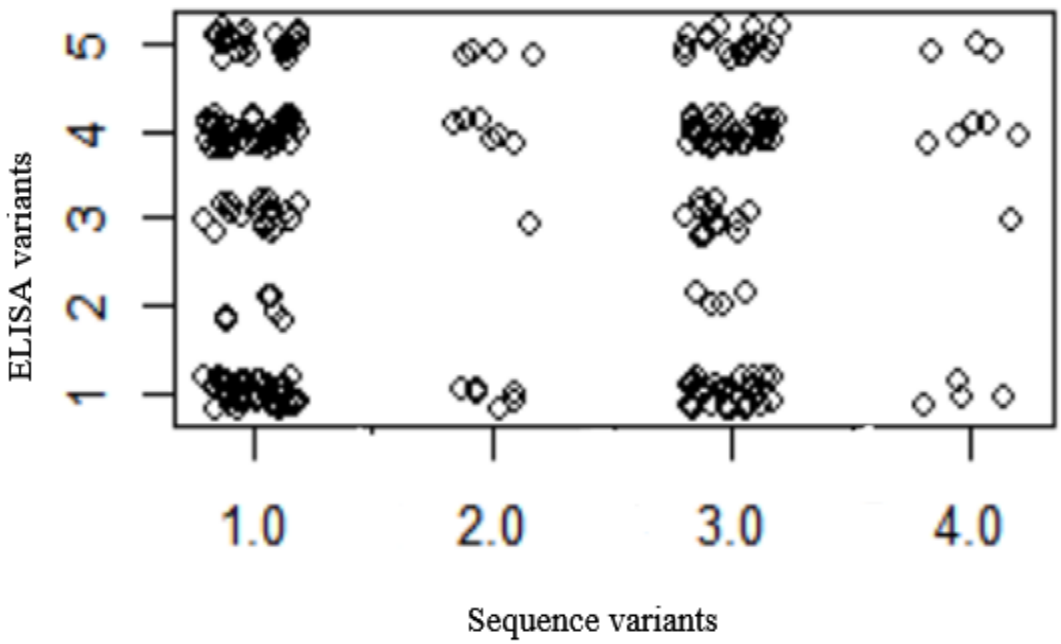


Figure 3.5 Sequence and ELISA comparison. Where 1 represents variant YHK, 2 – HDN, 3 – HDK, 4 – YHN, 5 – YDK. The scatter graph shows no correlation between the sequence and ELISA variants for all the 92 episodes.

CHAPTER 4

DISCUSSION, CONCLUSION AND RECOMMENDATIONS

4.1 Discussion

Some of the most common changes in the DNA sequence are SNPs, which have been linked to the cause of some diseases and inducing virulence in parasites (Hayton et al. 2008; Zhao et al. 2014). Polymorphisms have also been implicated in causing allele-specific immune reactions in some malaria vaccine targets, which could lead to immune escape by variants carrying different alleles from the vaccinating allele. Therefore, there is the need to identify the genetic profile of PfRh5 in parasites found in multiple infections in children to determine if they caused a change in the antibody profile of the children.

302 samples were successfully amplified from the 422 DNA samples. This drop in numbers was due to technical failure during DNA amplification and sequencing. Sequences from this study revealed that PfRh5 had a few SNPs, 9, of which only four had a frequency of at least 10%. This is similar to observations in other studies,Bustamante et al. (2013) and as contrasted to some *P. falciparum* proteins like AMA1 (Ouattara et al. 2010).This would suggest that the protein is not under strong immune selection pressure. Analysis of clinical samples from Africa by Bustamante et al. (2013) showed similar frequencies for the Y147H and H148D mutations. However, data obtained in the same study showed that the C203Y mutation inthe African population had a frequency of 79% while ours was slightly higher at 95%. Wright et al. (2014) showed that this mutation was localized close to, or in the Rh5-basigin binding interface. Hayton et al (2008) identified thatamong the laboratory adapted isolates, the K429N SNPstood out in the Malayan Camp and Palo Alto strains, which had efficiencies of>20%among the parasites able to invade *A. nancymaae* erythrocytes. This SNP has only been reported in laboratory adapted isolates (Bustamante et al. 2013) and this is the first time this mutation is being reported in field isolates.

As we wanted to find out if the alleles were different in sequential infections from the 302 sequenced samples, we grouped the SNPs into the 5 haplotypes for this analysis. However, due to lack of complete coverage of the 3 polymorphic sites, we restricted our analysis to 100 samples. These are the ones that had sequence coverage of all the three positions corresponding to the 5 Rh5 variants. These positions had been chosen as they had shown high frequencies in several geographical locations (Bustamante et al. 2013). Thus, being the major

alleles, they were good candidates for studies in allele-dominant immunity. Analysis of the respective infections showed many instances of mixed infections, which could mean the children were exposed to multiple parasite genotypes during inoculation by the vector. This finding is also consistent with Anderson et al(2000) and Jeffares et al(2007) that in areas experiencing high malaria prevalence, many individuals carry a multiplicity of *P. falciparum* genotypes. Proper analysis of allele-dominant immunity would require using infections where the parasite was mono-allelic. This is to ensure that the observed immune effects are as a result of the infecting allele, and not confounded by the presence of other alleles. Thus, infections carrying mixed alleles were not included for further analysis, which reduced the analytical power further.

To determine if the infecting parasites elicited allele-dominant immune activity, plasma obtained from the children during the acute infection was used in indirect ELISAs with the 5 variants coating the ELISA plate. Each sample OD was subjected to the allele-specific cut-off to determine which variant the plasma responded to. This was important, as if only one cut-off had been used for all variants, it would have been hard to determine which of the variants produced a response. Also, ininstances with responses to multiple variants, it was possible for responses with lower magnitudes to one variant to be masked by those of higher magnitudes.

Samples from 20 European adults and 20 children from Ngerenya were used as naïve samples in this study. When analysed, the mean ODs for the European and Ngerenya samples showed no significant differences apart from ODs for allele YDK, which could indicate pipetting errors or cross-reactivity of this allele with other environmental, or *Plasmodium* proteins.As malaria is predominantly found in Africa (Langhorne et al. 2008; Doolan et al. 2009), it means the European adults are likely to not have seen malaria, hence their plasma was chosen as the negative control. On the other hand, Ngerenya is a village within Kilifi, a county that normally has two malaria seasons per year, which corresponds to the rainy seasons. With this, ODs from both Ngerenya and European plasma samples were similar and this could be explained by the decrease in malaria transmission to date, within Ngerenya. This drop can be attributed several factors like age, vegetation and low transmission rates in the area. In spite of the ODs not being significantly different between the two populations, the ones from Ngerenya children were slightly higher than those from the European adults. This could be attributed to possible cross-reactivity withother *Plasmodium* proteins or environmental exposures within the Ngerenya community. Nevertheless, the European plasma samples had lower ODs than those from Ngerenya and were used in this study as the naïve plasma.

From the sample ELISA data, the number of children who responded to Rh5 and the levels of response were generally low.This mirrors findings in several populations in Kenya, Mali and Papua New Guinea (PNG) (Douglas et al. 2012; Tran et al. 2014; Chiu et al. 2014) where immune responses to Rh5 in natural infections were low, but associated with protection from clinical disease in the Mali and PNG populations. However, vaccines using various delivery techniques such as Freund's adjuvant and Chimpanzee Adenovirus serotype 63 (ChAd63) vector, boosted with modified vaccinia virus Ankara (MVA) both expressing PfRH5 were efficacious against heterologous strains of *P. falciparum* in *Aotusnancymaae*(Douglas et al. 2015).Douglas et al. (2012)showed that rabbits and mice vaccinated with Rh5 mounted a strong immune response with high titre values. This could indicate that Rh5 has either low immunogenicity in natural infections, is exposed to the host immune system for a short time, or as proposed by Wright and Rayner (2014), the parasite protects Rh5 by having low levels of expression.

Further analysis of these results also showed that many of the episodes that were ELISA positive had responses to multiple variants, in line with findings that anti-Rh5 antibodies are strain transcending (Douglas et al. 2012; Bustamante et al. 2013). This is in contrast to the predominant allele-specific acquisition of antibodies to some other merozoite antigens such as AMA1 (Osier et al. 2010) and MSP3 (Osier et al. 2007).

However, since our interest was still on allele-dominant immunity, the observation of responses to multiple alleles had the effect of reducing the numbers of samples that could be analysed. This is for the same reason as when selecting out episodes with infecting parasites having multiple infections. The data also did not show any trend of immune responses when sequential infections within an individual were analysed. This suggests that a positive response at a given time point was not influenced by the preceding infection.

The haplotypes of infecting genotypes were also matched to the responses they elicited on respective infections. Most of the mono-allelic infecting parasites were from variant YHK, which could be explained by it also being the most abundant haplotype in this population. Approximately, only 13% of the infections displayed allele-dominant reactions where the immune response was only to the infecting genotype within an infection. The rest of the infections with a mono-allelic infecting parasite displayed allele transcending immune responses. However, 50% of these allele-dominant reactions were in one individual and thus, this may not be representative of the population as a whole.

Data from this work therefore suggests that immune response with respect to PfRh5, are not dependant on the infecting genotype as Rh5 elicits mainly allele transcending immune responses in multiple infections

4.2 Conclusions

This work set out to determine if *P. falciparum* parasites found in sequential infections were different in relation to the merozoite protein Rh5, and if the Rh5 variants would elicit allele-dominant immune responses in the host. Rh5 variants present in sequential clinical infections were mainly mixed and in spite of an apparently low antibody response in few individuals, many immune responses elicited allele transcending immunity. This is important as previous vaccine candidates have had the challenge of exhibiting stain-specific immunity, which is not the case with PfRh5, as suggested in this work. Thus, the importance of PfRh5 as a viable vaccine candidate has been shown, highlighting its relevance for inclusion in a malaria vaccine. This studyalso reports the presence of the K429N mutation at a high frequency (17.2%) previously described only in laboratory strains. This mutation is present in the Malayan Camp, and Palo Alto laboratory strains which have an increased ability to invade *Aotus*erythrocytes. However, its effect on infectivity in humans is not yet known, and further studies need to be done to determine this. Overall, this study is a good model for studying potential vaccine candidates to determine if they elicit allele-dominant responses, information necessary when determining the suitability of an antigen for inclusion in vaccine studies. This is because it directly compares the immune responses observed to the infective strains within multiple infections.

4.3 Recommendations

Due to a shortage of streptavidin coated ELISA plates, each sample was assayed once instead of being done in duplicates or triplicates. I would thus recommend that if this study was to be repeated or for similar studies in the future, ELISA should be done in duplicates or triplicates to increase the accuracy of results obtained.

The children picked for this study were from a cohort of those having multiple acute infections. Thus, there is the limitation that the observed parasites or immune responses could be from an overlap of previous infections. Thisis because the child may not have had time to either clear the previous parasites or make adequate responses to parasites in the current infection. Similarly, since the children had many infections, at least 8, it could mean that their immune system was already weak, hence influencing the results observed. I would therefore

recommend a similar study, which would also include children with fewer infections and samples from convalescent cases. Including cases with fewer infections would help determine if the observed results in the group with many cases are true by controlling for number of cases. Inclusion of a convalescence group would control for the possibility of observations being from overlaps in infection and immune responses. Within this study, inclusion of an antigen such as AMA1 or MSP3, which elicit high antibody titres should also be included as a control.

Previous studies have also shown that PfRh5 has low immunogenicity in natural infections, elicits strain transcending antibodies and that there is minimal acquisition of anti-PfRh5 antibodies with age (Douglas et al. 2012). The data shows, PfRh5 may not be the best model to study allele-dominant immunity. Therefore, similar studies but with more immunogenic proteins should be used.

The use of children with multiple sequential episodes was important as it determined if the parasite and immune response to it in one infection could affect the type of response in subsequent infections. However, due to the inherent low level responses to Rh5, there was little data from sequential responses for analysis. This thus shows this model of study to be good for studying allele-dominant immunity but best suited toantigensthat are more immunogenic than Rh5.

REFERENCES

Abu Bakar N, Klonis N, Hanssen E, et al (2010) Digestive-vacuole genesis and endocytic processes in the early intraerythrocytic stages of Plasmodium falciparum. J Cell Sci 123:441–50.

Adam AAM, Amine AAA, Hassan DA, et al (2013) Distribution of erythrocyte binding antigen 175 (EBA-175) gene dimorphic alleles in Plasmodium falciparum field isolates from Sudan. BMC Infect Dis 13:469.

Aikawa M, Miller LH, Johnson J, Rabbege J (1978) Erythrocyte entry by malarial parasites. A moving junction between erythrocyte and parasite. J Cell Biol 77:72–82.

Anderson TJC, Haubold B, Williams JT, et al (2000) Microsatellite Markers Reveal a Spectrum of Population Structures in the Malaria Parasite Plasmodium falciparum. Mol Biol Evol 1467–1482.

Bannister LH, Hopkins JM, Fowler RE, et al (2000) A Brief Illustrated Guide to the Ultrastructure of Plasmodium falciparum Asexual Blood Stages. Trends Parasitol 16:427–433.

Bartholdson SJ, Crosnier C, Bustamante LY, et al (2013) Identifying novel Plasmodium falciparum erythrocyte invasion receptors using systematic extracellular protein interaction screens. Cell Microbiol 15:1304–12.

Baum J, Chen L, Healer J, et al (2009) Reticulocyte-binding protein homologue 5 – An essential adhesin involved in invasion of human erythrocytes by Plasmodium falciparum. Int J Parasitol 39:371–380.

Beeson JG, Brown G V (2002) Pathogenesis of Plasmodium falciparum malaria: the roles of parasite adhesion and antigenic variation. Cell Mol Life Sci 59:258–71.

Bergholdt R, Brorsson C, Palleja A, et al (2012) Identification of novel type 1 diabetes candidate genes by integrating genome-wide association data, protein-protein interactions, and human pancreatic islet gene expression. Diabetes 61:954–962.

Boyle MJ, Reiling L, Feng G, et al (2015) Human Antibodies Fix Complement to Inhibit Plasmodium falciparum Invasion of Erythrocytes and Are Associated with Protection against Malaria. Immunity 42:580–90.

Brookes AJ (1999) The essence of SNPs. Gene 234:177–186.

Bull PC, Marsh K (2002) The role of antibodies to Plasmodium falciparum- infected-erythrocyte surface antigens in naturally acquired immunity to malaria. Trends Microbiol 10:55–58.

Bustamante LY, Bartholdson SJ, Crosnier C, et al (2013) A full-length recombinant Plasmodium falciparum PfRH5 protein induces inhibitory antibodies that are effective across common PfRH5 genetic variants. Vaccine 31:373–9.

Cavallo a., Martin a. CR (2005) Mapping SNPs to protein sequence and structure data. Bioinformatics 21:1443–1450.

Chen L, Lopaticki S, Riglar DT, et al (2011) An EGF-like protein forms a complex with PfRh5 and is required for invasion of human erythrocytes by Plasmodium falciparum. PLoS Pathog 7:e1002199.

Chen L, Xu Y, Healer J, et al (2014) Crystal structure of PfRh5 , an essential P . falciparum ligand for invasion of human erythrocytes. Elife 1–10.

Child MA, Epp C, Bujard H, Blackman MJ (2010) Regulated maturation of malaria merozoite surface protein-1 is essential for parasite growth. Mol Microbiol 78:187–202.

Chiu CYH, Healer J, Thompson JK, et al (2014) Association of antibodies to Plasmodium falciparum reticulocyte binding protein homolog 5 with protection from clinical malaria. Front Microbiol 5:314.

Cho RJ, Mindrinos M, Richards DR, et al (1999) Genome-wide mapping with biallelic markers in Arabidopsis thaliana. Nat Genet 23:203–207.

Cohen S, McGregor I, Carrington S (1961) Gamma-Globulin and Aquired Immunity to Human Malaria. Nature 192:733–737.

Collins FS, Brooks LD, Chakravarti A (1998) A DNA Polymorphism Discovery Resource for Research on Human Genetic Variation. Genome Res 8:1229–1231.

Cowman AF, Berry D, Baum J (2012) The cellular and molecular basis for malaria parasite invasion of the human red blood cell. J Cell Biol 198:961–71.

Cowman AF, Crabb BS (2006) Invasion of red blood cells by malaria parasites. Cell 124:755–66.

Cox FEG (2010) History of the discovery of the malaria parasites and their vectors. Parasit Vectors 3:1–9.

Crosnier C, Bustamante LY, Bartholdson SJ, et al (2011) Basigin is a receptor essential for erythrocyte invasion by Plasmodium falciparum. Nature 480:534–7.

Daubersies P, Sallenave-Sales S, Magne S, et al (1996) Rapid turnover of Plasmodium falciparum populations in asymptomatic individuals living in a high transmission area. Am J Trop Med Hyg 54:18–26.

Derbyshire ER, Mota MM, Clardy J (2011) The Next Opportunity in Anti-Malaria Drug Discovery : The Liver Stage. PLoS Pathog 7:e1002178.

Doolan DL, Dobaño C, Baird JK (2009) Acquired immunity to malaria. Clin Microbiol Rev 22:13–36

Douglas AD, Baldeviano GC, Lucas CM, et al (2015) A PfRH5-Based Vaccine Is Efficacious against Heterologous Strain Blood-Stage Plasmodium falciparum Infection in Aotus

Monkeys. Cell Host Microbe 130–139.

Douglas AD, Williams AR, Illingworth JJ, et al (2012) The Blood-Stage Malaria Antigen PfRH5 is Susceptible to Vaccine-Inducible Cross-Strain Neutralizing Antibody. Nat Commun 2:601.

Dvorin JD, Bei AK, Coleman BI, Duraisingh MT (2010) Functional diversification between two related Plasmodium falciparum merozoite invasion ligands is determined by changes in the cytoplasmic domain. Mol Microbiol 75:990–1006.

Gerald N, Mahajan B, Kumar S (2011) Mitosis in the Human Malaria Parasite Plasmodium falciparum. Eukaryot Cell 10:474–482.

Good MF, Doolan DL (1999) Immune effector mechanisms in malaria. Curr Opin Immunol 11:412–9.

Gupta AP, Chin WH, Zhu L, et al (2013) Dynamic Epigenetic Regulation of Gene Expression during the Life Cycle of Malaria Parasite Plasmodium falciparum. PLoS Pathog 9:e1003170.

Hayton K, Gaur D, Liu A, et al (2008) Erythrocyte binding protein PfRH5 polymorphisms determine species-specific pathways of Plasmodium falciparum invasion. Cell Host Microbe 4:40–51.

Healer J, Thompson JK, Riglar DT, et al (2013) Vaccination with conserved regions of erythrocyte-binding antigens induces neutralizing antibodies against multiple strains of Plasmodium falciparum. PLoS One 8:e72504.

Hirakawa H, Shirasawa K, Ohyama A, et al (2013) Genome-wide SNP genotyping to infer the effects on gene functions in tomato. DNA Res 20:221–233.

http://www.cdc.gov/dpdx/malaria/index.html CDC - DPDx - Malaria. http://www.cdc.gov/dpdx/malaria/index.html. Accessed 10 Feb 2014

IEBC (2012) Maps & Locations :: Kilifi County Government. http://kilifi.go.ke/content.php?com=4&com2=44&com3=#.VkJat7crK00. Accessed 10 Nov 2015

Iyer J, Grüner AC, Rénia L, et al (2007) Invasion of host cells by malaria parasites: a tale of two protein families. Mol Microbiol 65:231–49.

Jeffares DC, Pain A, Berry A, et al (2007) Genome variation and evolution of the malaria parasite Plasmodium falciparum. Nat Genet 39:120–125.

Langhorne J, Ndungu FM, Sponaas A-M, Marsh K (2008) Immunity to malaria: more questions than answers. Nat Immunol 9:725–732.

Liao C, Kong L, Song F, et al (2011) Characterization of Basigin Isoforms and the Inhibitory Function of Basigin-3 in Human Hepatocellular Carcinoma Proliferation and Invasion □. Mol Cell Biol 31:2591–2604.

Lobo C, Rodriguez M, Reid M, Lustigman S (2003) Glycophorin C is the receptor for the Plasmodium falciparum erythrocyte binding ligand PfEBP-2 (baebl). Blood 101:4628–4631.

M. Fried,; F. Nosten, A. Brockman, B. Brabin PD (1998) Maternal antibodies block malaria. Nature 20:851–852.

Marsh K, Kinyanjui S (2006) Immune effector mechanisms in malaria. Parasite Immunol 28:51–60.

Mayer DCG, Cofie J, Jiang L, et al (2009) Glycophorin B is the erythrocyte receptor of Plasmodium falciparum erythrocyte-binding ligand , EBL-1. Proc Natl Acad Sci 106:5348–5352.

Miller LH, McAuliffe FM, Johnson JG (1979) Invasion of erythrocytes by malaria merozoites. Prog Clin Biol Res 30:497–502.

Mitchell GH, Thomas AW, Margos G, et al (2004) Apical Membrane Antigen 1 , a Major Malaria Vaccine Candidate , Mediates the Close Attachment of Invasive Merozoites to Host Red Blood Cells. Infect Immun 72:154–158.

Miyauchi T, Masuzawa Y, Muramatsu T (1991) The Basigin Group of the Immunoglobulin Superfamily : Complete Conservation of a Segment in and around Transmembrane Domains of Human and Mouse Basigin and Chicken HT7 Antigen. J Biochem 110:770–774.

NordNordWest (2015) Kilifi County in Kenya - Wikipedia, the free encyclopedia. https://en.wikipedia.org/wiki/File:Kilifi_County_in_Kenya.svg. Accessed 10 Nov 2015

Olivieri A, Collins CR, Hackett F, et al (2011) Juxtamembrane shedding of Plasmodium falciparum AMA1 is sequence independent and essential, and helps evade invasion-inhibitory antibodies. PLoS Pathog 7:e1002448.

Ord RL, Rodriguez M, Yamasaki T, et al (2012) Targeting sialic acid dependent and independent pathways of invasion in Plasmodium falciparum. PLoS One 7:e30251.

Orlandi PA, Klotz FW, Haynes JD, Reed W (1992) A Malaria Invasion Receptor , the 175-Kilodalton Erythrocyte Binding Antigen of Plasmodium falciparum Recognizes the Terminal Neu5Ac (a2-3) Gal- Sequences of Glycophorin A Pleriodate Oxidation / Borohydride Reduction. J Cell Biol 116:901–909.

Osier FH a, Weedall GD, Verra F, et al (2010) Allelic diversity and naturally acquired allele-specific antibody responses to Plasmodium falciparum apical membrane antigen 1 in Kenya. Infect Immun 78:4625–4633.

Osier FHA, Fegan G, Polley SD, et al (2008) Breadth and Magnitude of Antibody Responses to Multiple Plasmodium falciparum Merozoite Antigens Are Associated with Protection from Clinical Malaria. Infect Immun 76:2240–2248.

Osier FHA, Polley SD, Mwangi T, et al (2007) Naturally acquired antibodies to polymorphic and conserved epitopes of Plasmodium falciparum merozoite surface protein 3. Parasite Immunol 29:387–94.

Ouattara A, Mu J, Takala-Harrison S, et al (2010) Lack of allele-specific efficacy of a bivalent AMA1 malaria vaccine. Malar J 9:175.

Patel SD, Ahouidi AD, Bei AK, et al (2013) Plasmodium falciparum merozoite surface antigen, PfRH5, elicits detectable levels of invasion-inhibiting antibodies in humans. J Infect Dis 208:1679–87.

Pierce SK, Crompton PD (2013) Young Lives Lost as B Cells Falter: What We Are Learning About Antibody Responses in Malaria. J Immunol 190:3039–3046.

Pierce SK, Crompton PD, Portugal S (2013) Young lives lost as B cells falter: what we are learning about antibody responses in malaria. J Immunol 190:3039–46.

Rayner JC, Galinski MR, Ingravallo P, Barnwell JW (2000) Two Plasmodium falciparum genes express merozoite proteins that are related to Plasmodium vivax and Plasmodium yoelii adhesive proteins involved in host cell selection and invasion. Proc Natl Acad Sci 97:9648–9653.

Reddy KS, Amlabu E, Pandey AK, et al (2015) Multiprotein complex between the GPI-anchored CyRPA with PfRH5 and PfRipr is crucial for Plasmodium falciparum erythrocyte invasion. Proc Natl Acad Sci 112:1179–1184.

Riglar DT, Richard D, Wilson DW, et al (2011) Super-resolution dissection of coordinated events during malaria parasite invasion of the human erythrocyte. Cell Host Microbe 9:9–20.

Riley EM (2004) Does malaria suffer from lack of memory ? Immunol Rev 201:268–290.

Rodriguez M, Lustigman S, Montero E, et al (2008) PfRH5: a novel reticulocyte-binding family homolog of plasmodium falciparum that binds to the erythrocyte, and an investigation of its receptor. PLoS One 3:e3300.

Sahar T, Reddy KS, Bharadwaj M, et al (2011) Plasmodium falciparum reticulocyte binding-like homologue protein 2 (PfRH2) is a key adhesive molecule involved in erythrocyte invasion. PLoS One 6:e17102.

Schatz MC, Delcher AL, Salzberg SL (2010) Assembly of large genomes using second-generation sequencing. Genome Res 20:1165–73.

Singh S, Alam MM, Pal-Bhowmick I, et al (2010) Distinct external signals trigger sequential release of apical organelles during erythrocyte invasion by malaria parasites. PLoS Pathog 6:e1000746.

Srinivasan P, Beatty WL, Diouf A, et al (2011) Binding of Plasmodium merozoite proteins RON2 and AMA1 triggers commitment to invasion. Proc Natl Acad Sci U S A

108:13275–80.

Struik SS, Riley EM (2004) Does malaria suffer from lack of memory? Immunol Rev 201:268–290.

Taylor HM, McRobert L, Grainger M, et al (2010) The malaria parasite cyclic GMP-dependent protein kinase plays a central role in blood-stage schizogony. Eukaryot Cell 9:37–45.

Tham W, Wilson DW, Lopaticki S, et al (2010) Complement receptor 1 is the host erythrocyte receptor for Plasmodium falciparum PfRh4 invasion ligand. Proc Natl Acad Sci U S A 107:17327–32.

Tham W-H, Healer J, Cowman AF (2012) Erythrocyte and reticulocyte binding-like proteins of Plasmodium falciparum. Trends Parasitol 28:23–30.

Tran TM, Ongoiba A, Coursen J, et al (2014) Naturally Acquired Antibodies Speci fi c for Plasmodium falciparum Reticulocyte-Binding Protein Homologue 5 Inhibit Parasite Growth and Predict Protection From Malaria. J Infect Dis 209:789–798.

Van K, Hwang E-Y, Kim MY, et al (2005) Discovery of SNPs in soybean genotypes frequently used as the parents of mapping populations in the United States and Korea. J Hered 96:529–35.

Wright GJ, Rayner JC (2014) Plasmodium falciparum erythrocyte invasion: combining function with immune evasion. PLoS Pathog 10:e1003943.

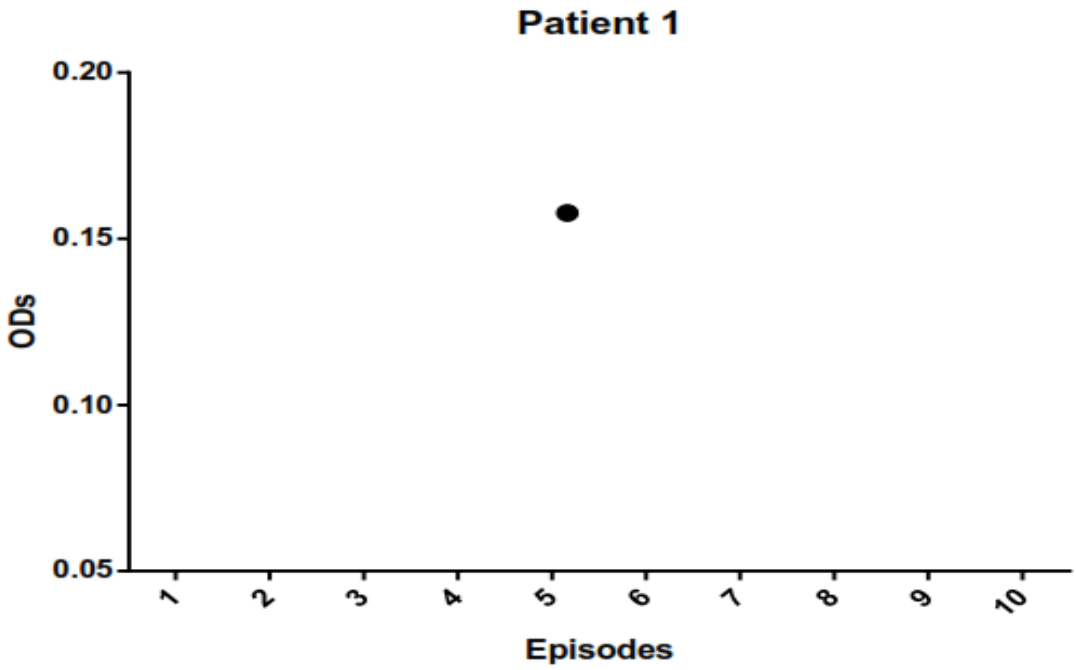
Wright KE, Hjerrild KA, Bartlett J, et al (2014) erythrocyte basigin and blocking antibodies. Nature 515:427–430.

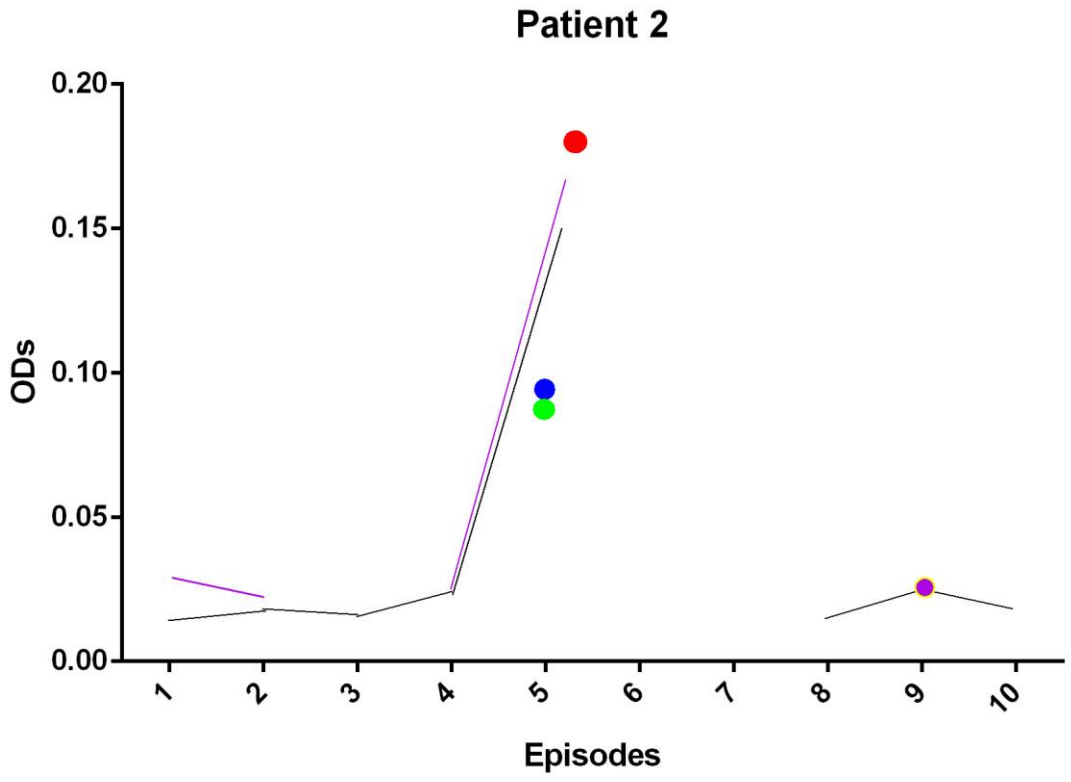
Yu X, Hu T, Du J, et al (2008) Crystal Structure of HAb18G / CD147. J Biol Chem 283:18056–18065.

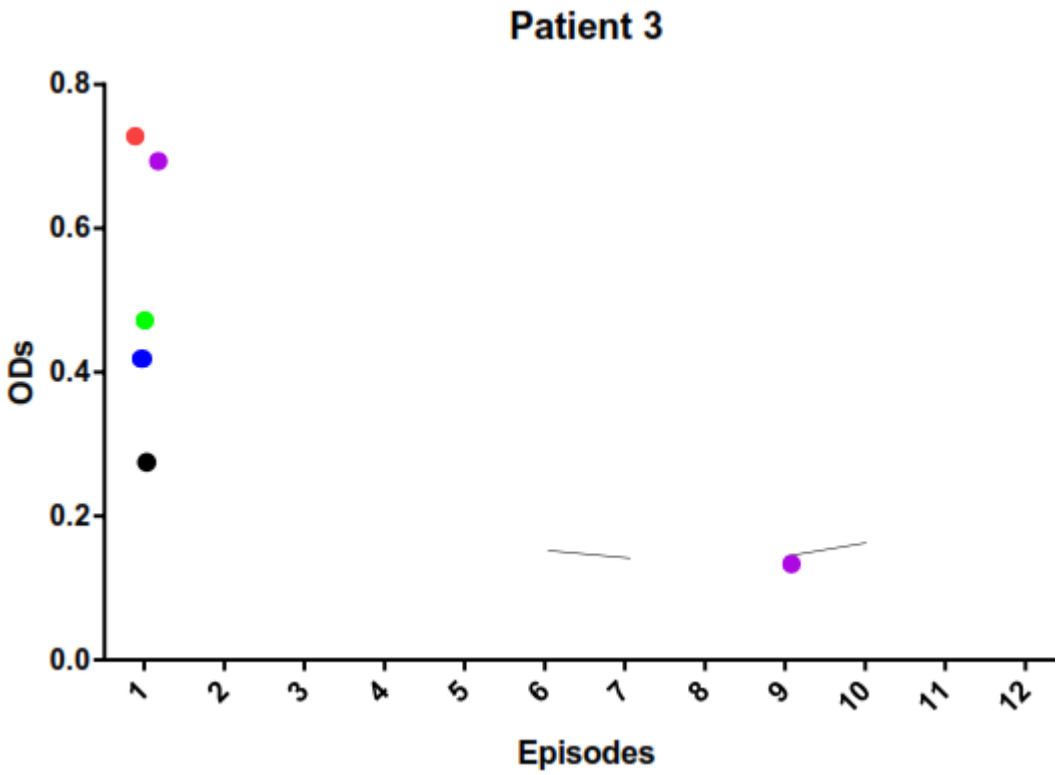
Zhao N, Han JG, Shyu C-R, Korkin D (2014) Determining effects of non-synonymous SNPs on protein-protein interactions using supervised and semi-supervised learning. PLoS Comput Biol 10:e1003592.

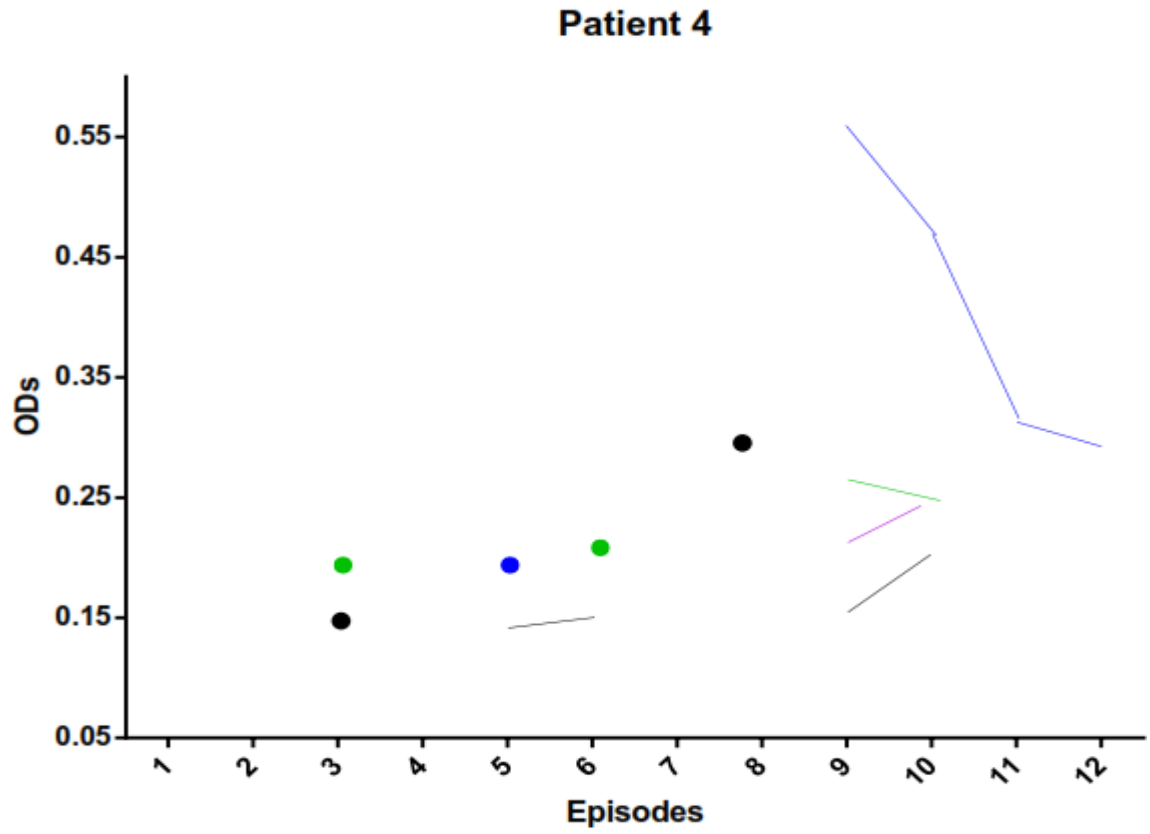
APPENDIX

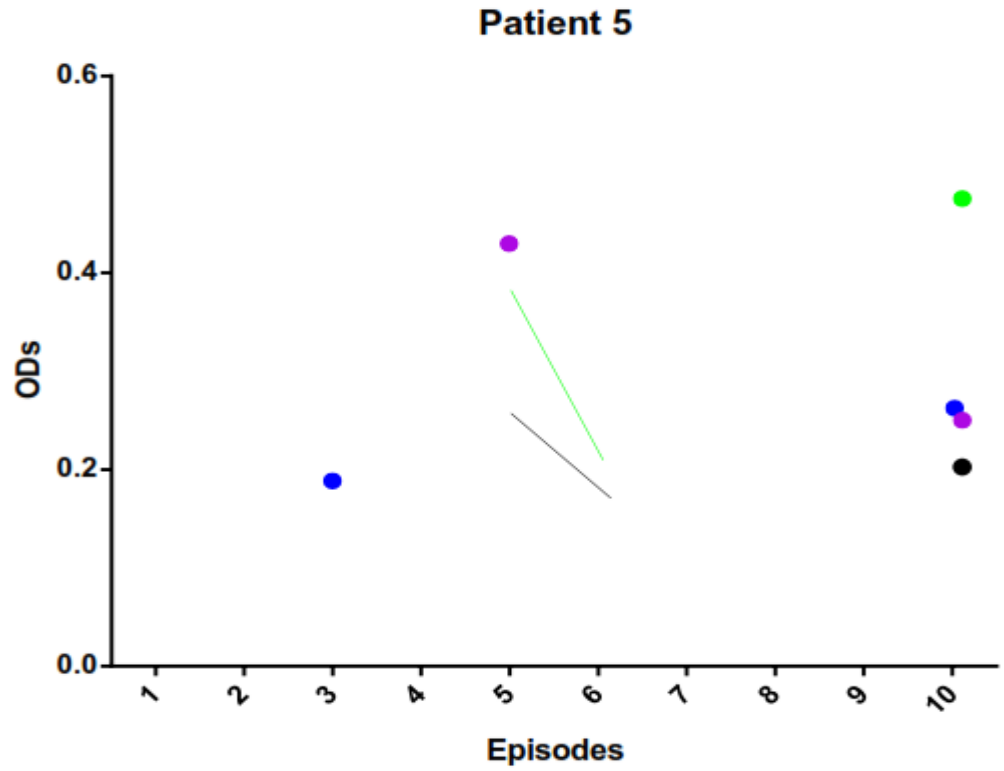
Appendix 1: Trend of responses within individuals. Color keys are; blue YHK, red HDN, purple HDK, black YHN, and green YDK.

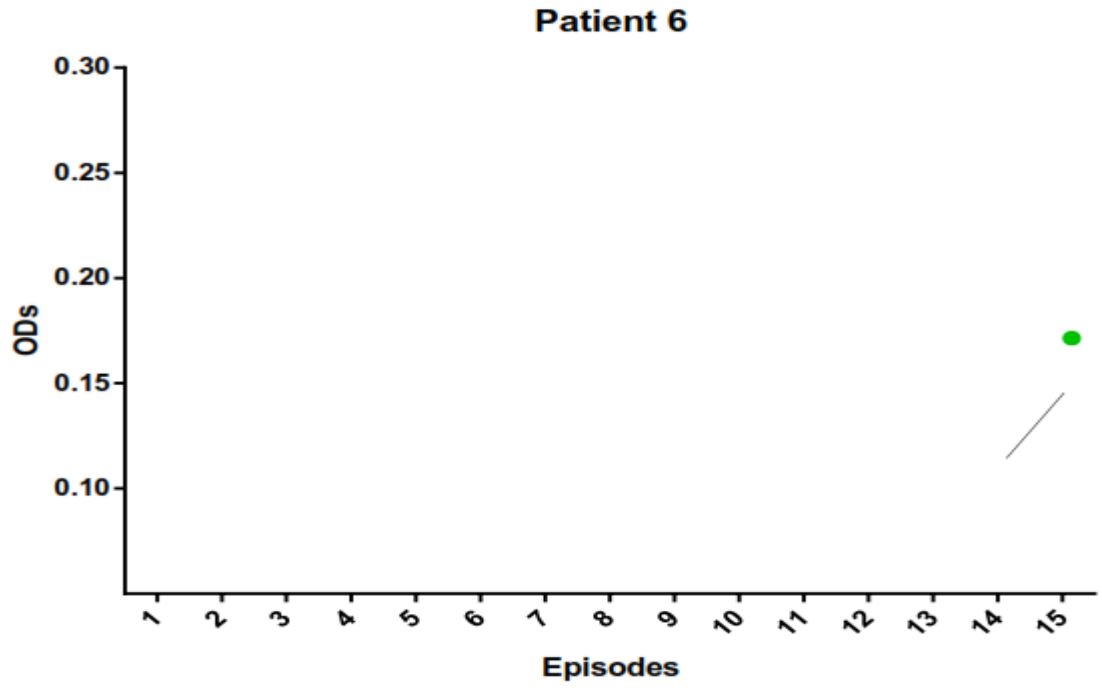


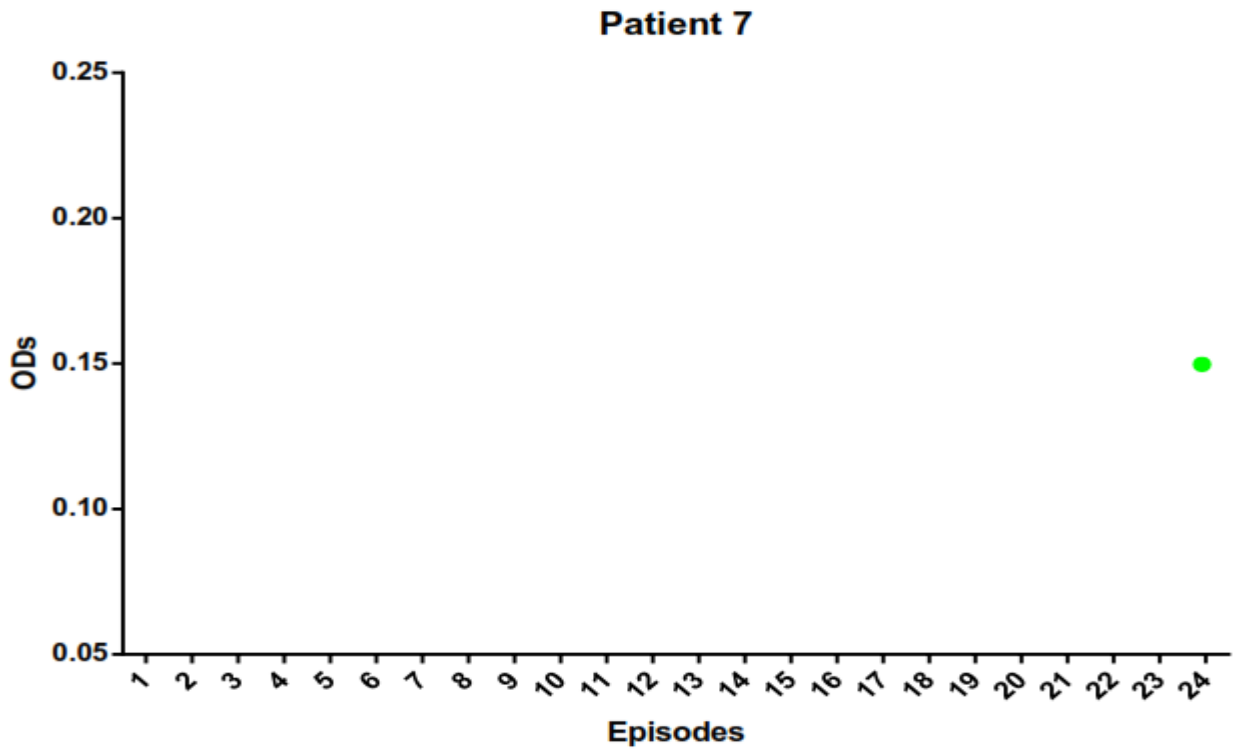


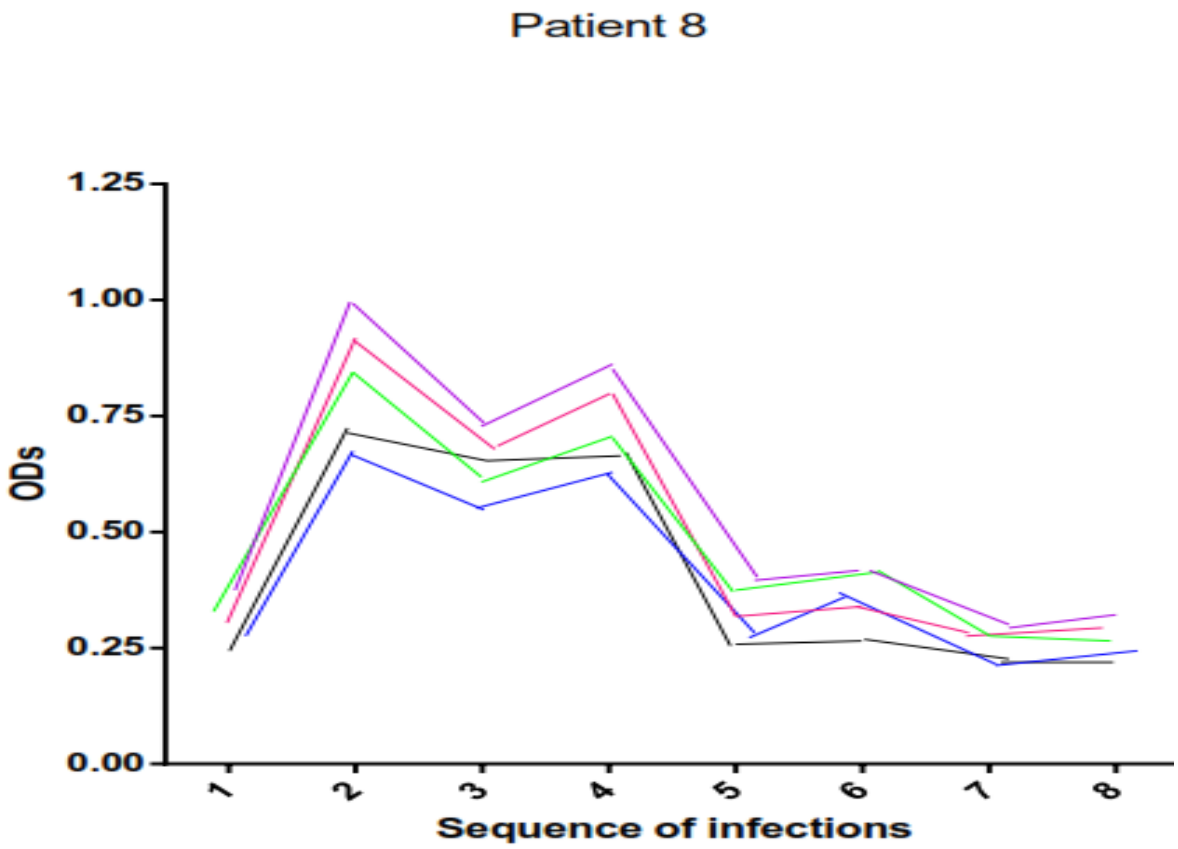

Patient 3
0.8
0.6
0.4
0.2
0.0
ODs
1
2
3
4
5
6
7
8
9
10
11
12
Episodes

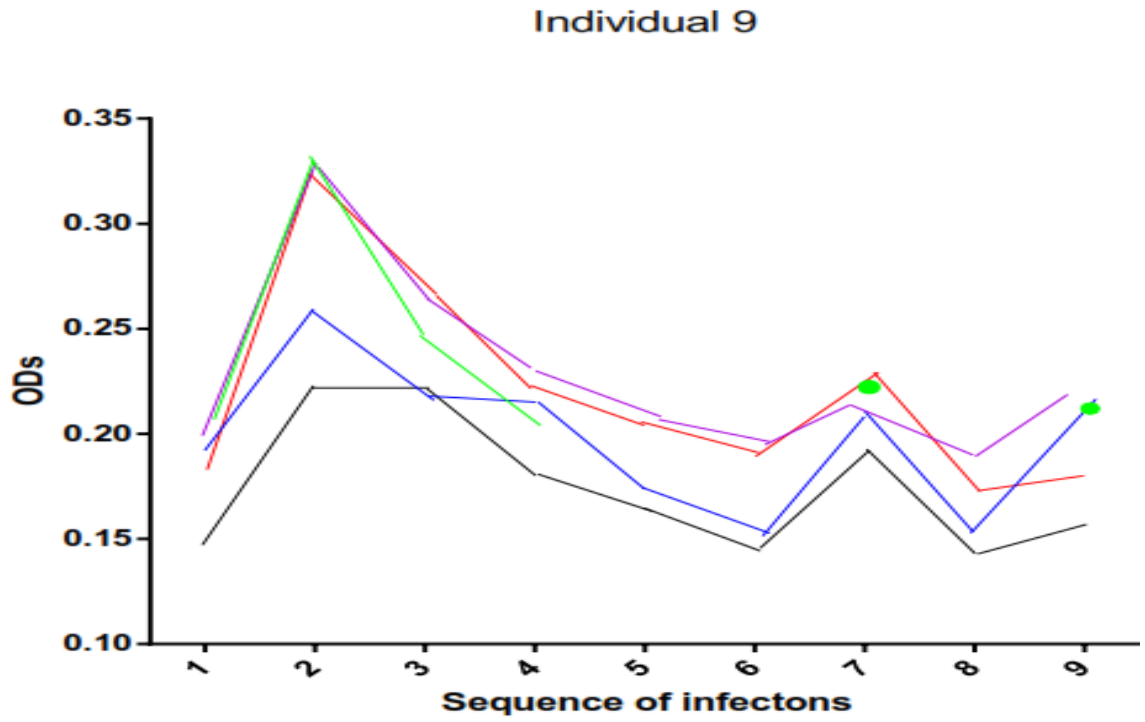

Patient 4
0.55
0.45
0.35
0.25
0.15
0.05
ODs
1
2
3
4
5
6
7
8
9
10
11
12
Episodes

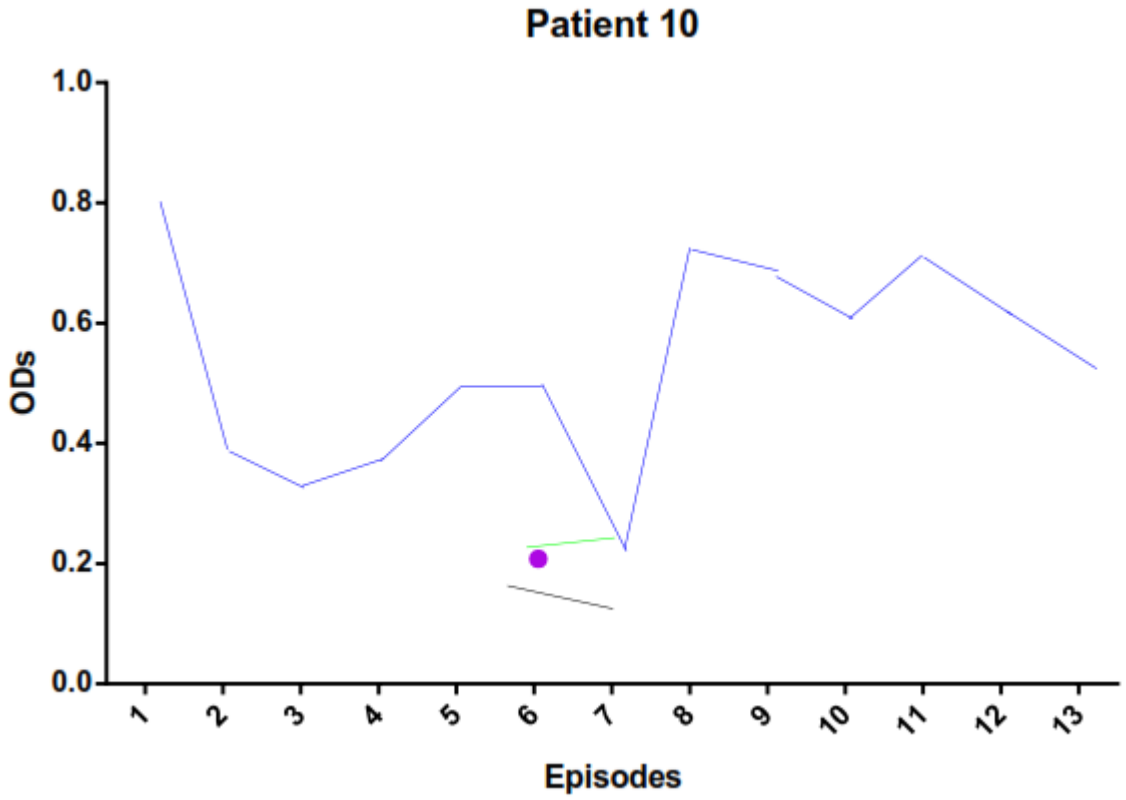

Patient 5
0.6
0.4
0.2
0.0
ODs
1
2
3
4
5
6
7
8
9
10
Episodes

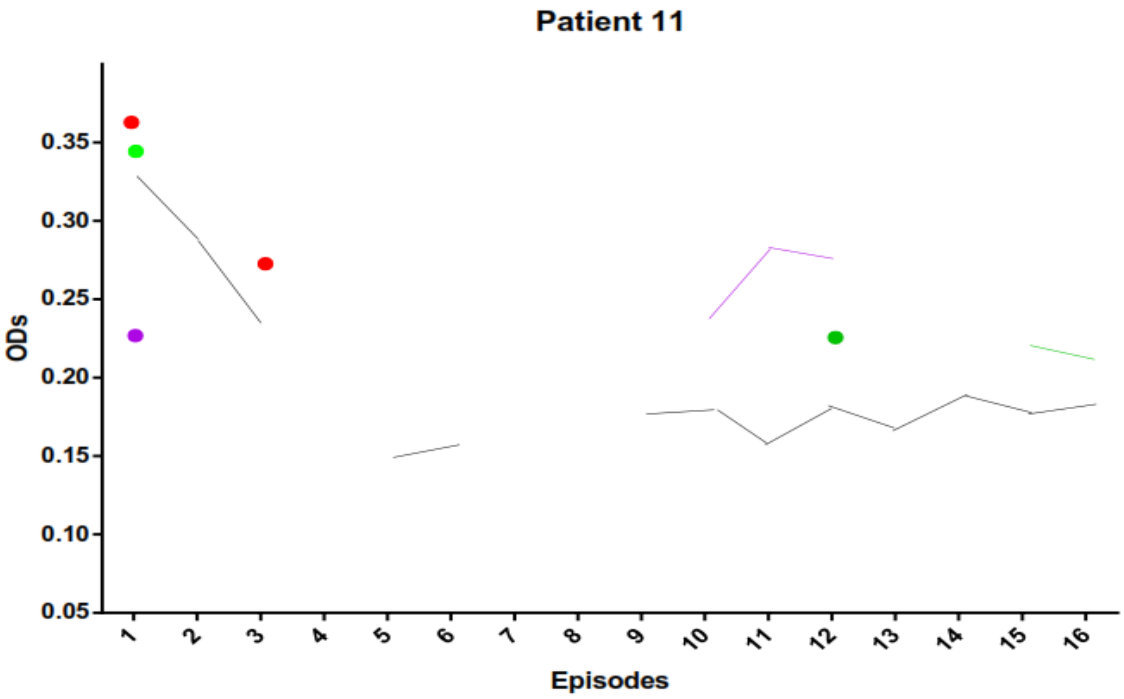

Patient 6
0.30
0.25
0.20
0.15
0.10
ODs
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
Episodes

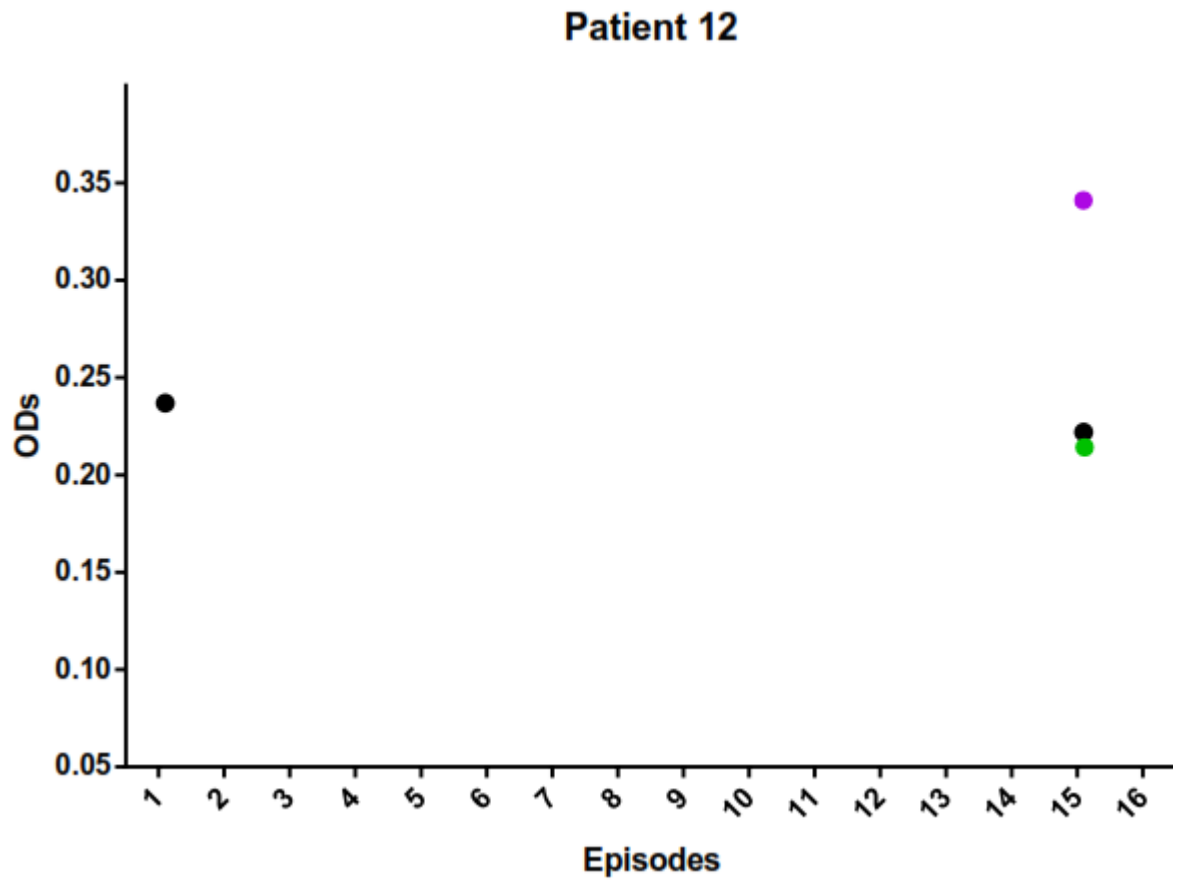

Patient 7
0.25
0.20
0.15
0.10
0.05
ODs
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
Episodes

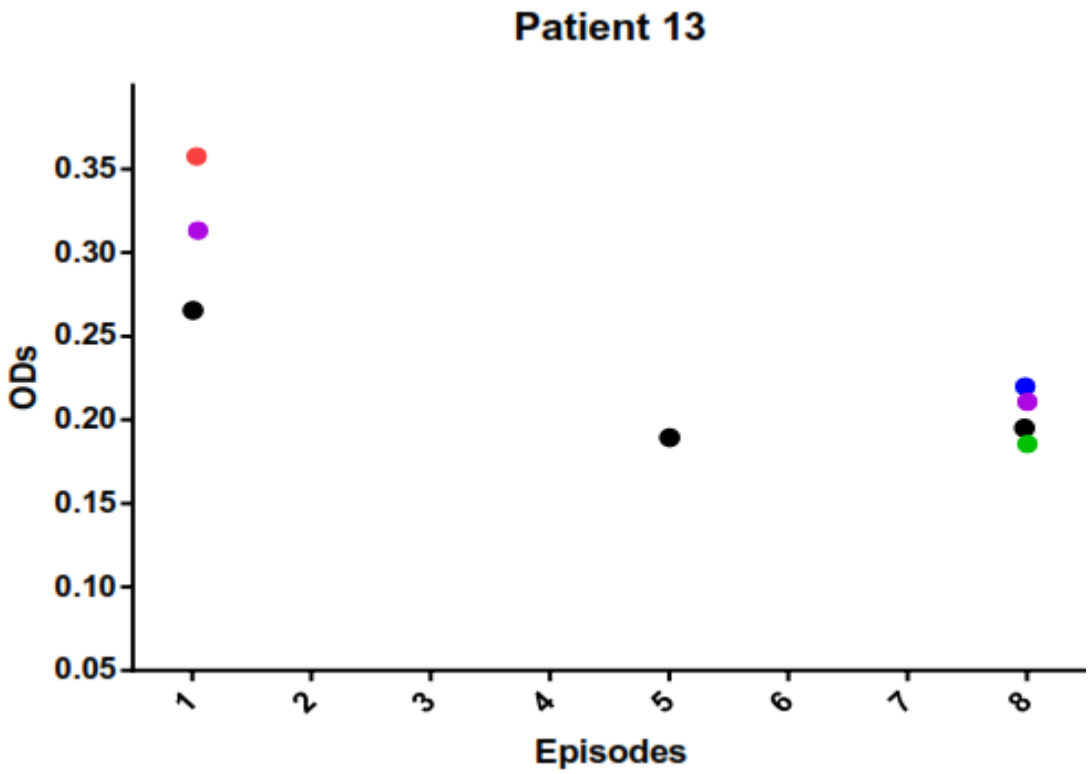

Patient 8
1.25
1.00
0.75
0.50
0.25
0.00
ODs
1 2 3 4 5 6 7 8
Sequence of infections

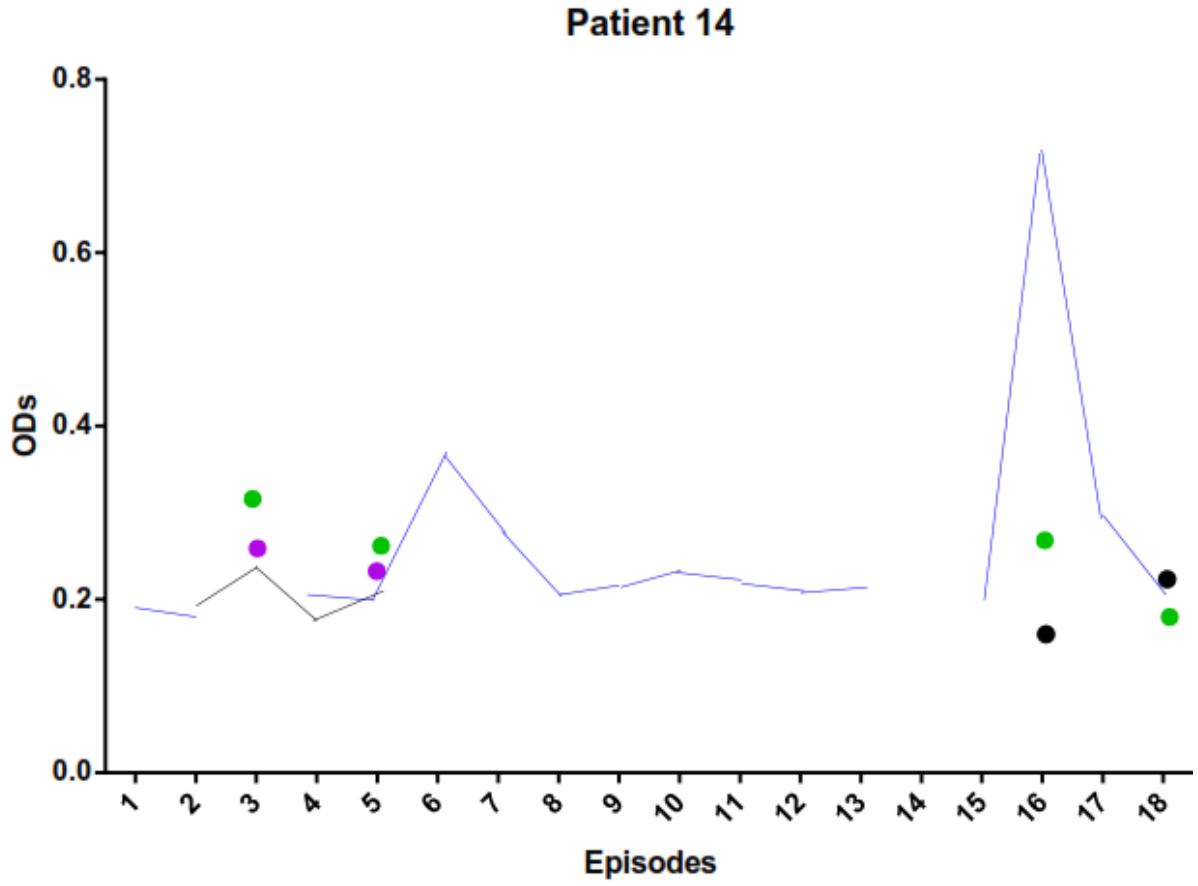

Individual 9
0.35
0.30
0.25
0.20
0.15
0.10
ODs
1
2
3
4
5
6
7
8
9
Sequence of infectons

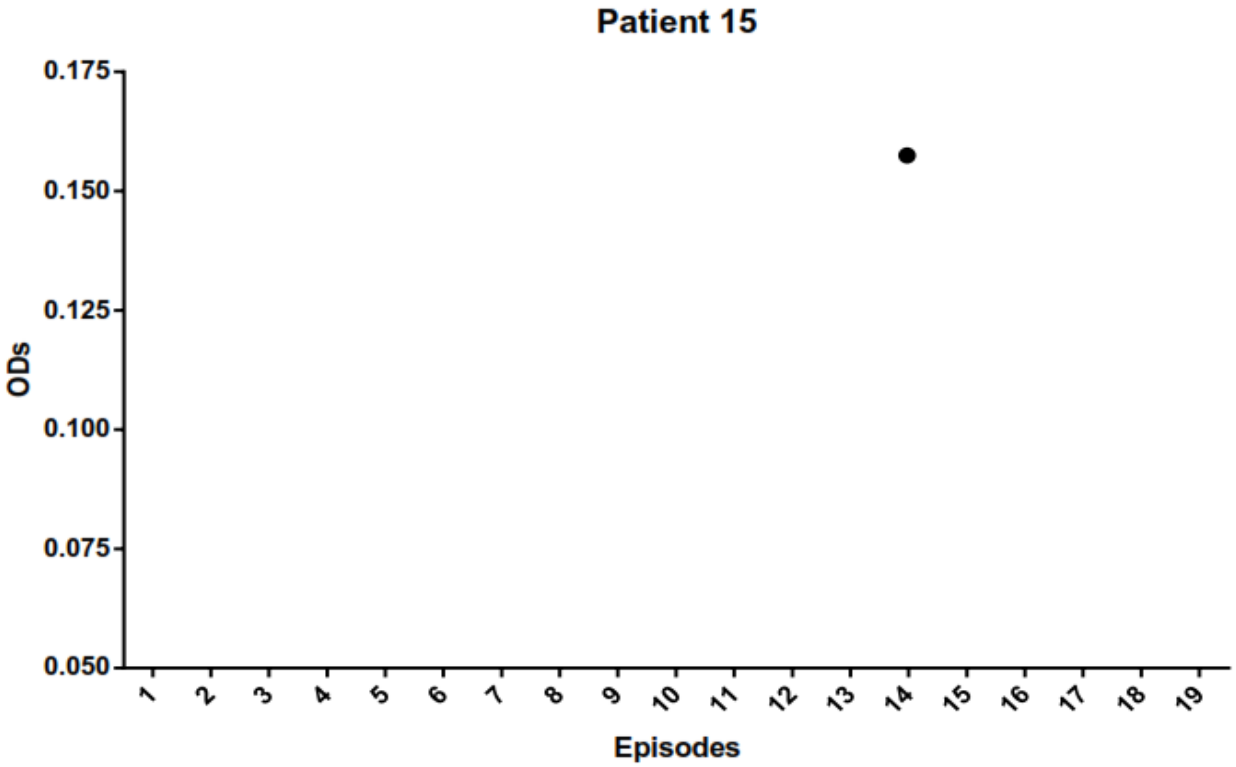

Patient 10
1.0
0.8
0.6
0.4
0.2
0.0
ODs
1
2
3
4
5
6
7
8
9
10
11
12
13
Episodes

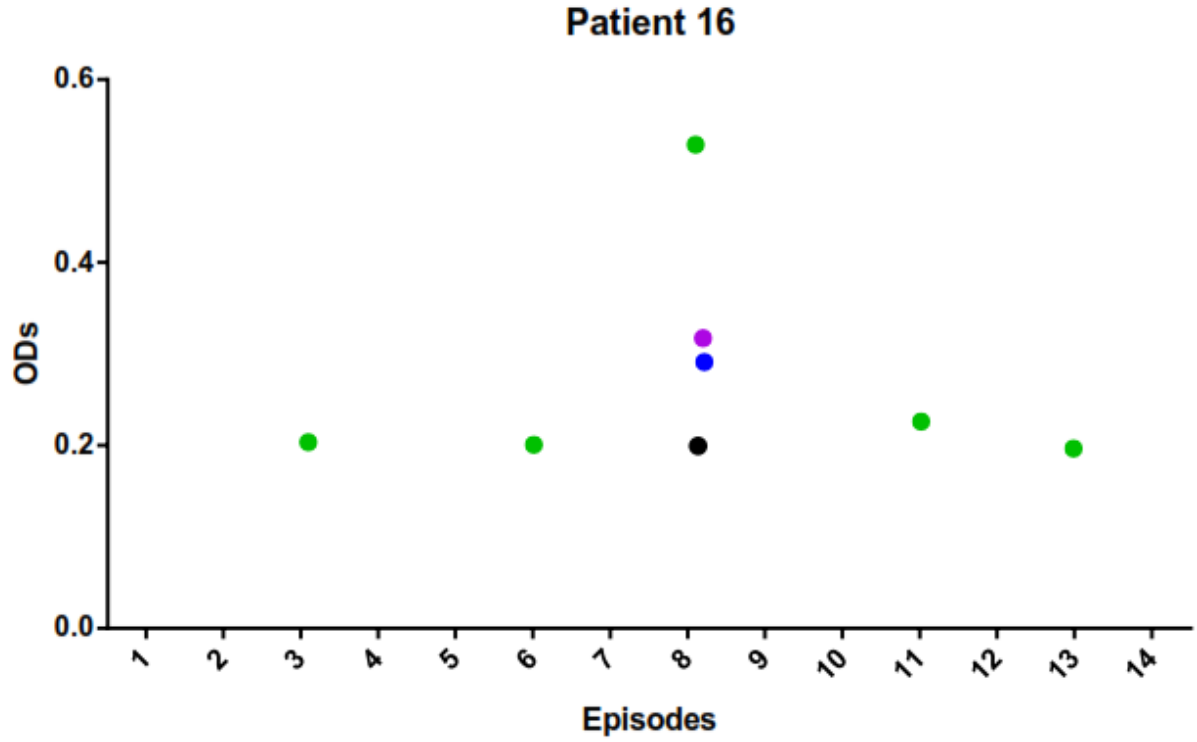

Patient 11
ODs
0.35
0.30
0.25
0.20
0.15
0.10
0.05
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
Episodes

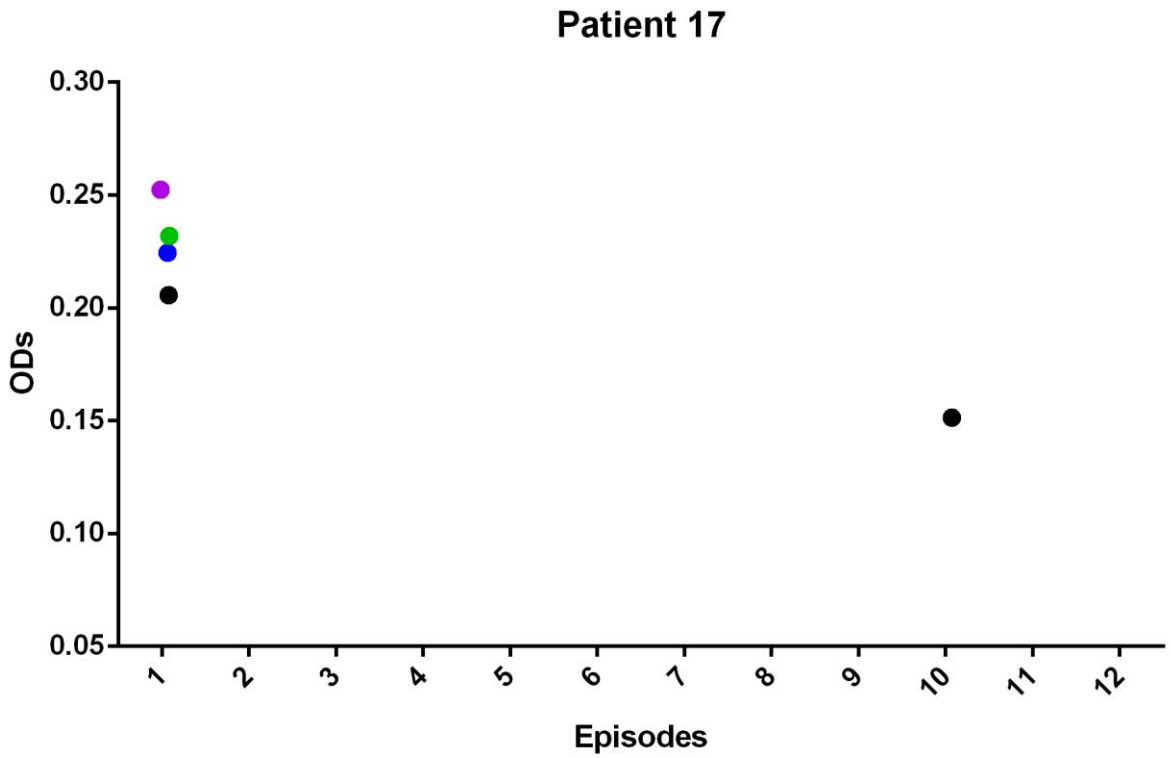

Patient 12
ODs
0.35
0.30
0.25
0.20
0.15
0.10
0.05
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
Episodes

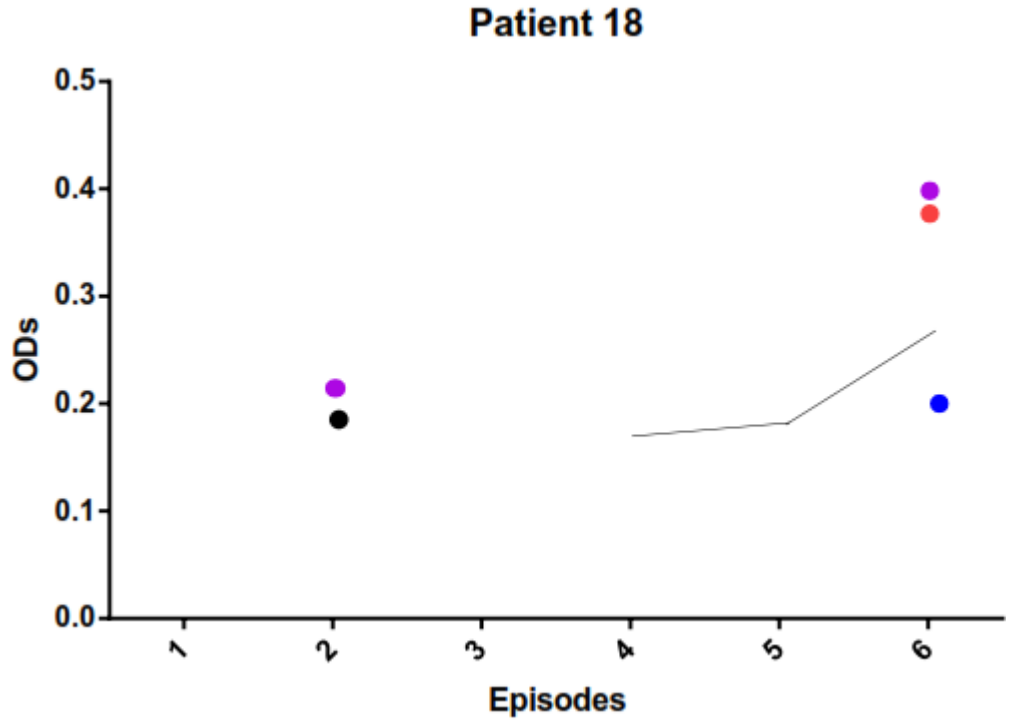

Patient 13
ODs
0.35
0.30
0.25
0.20
0.15
0.10
0.05
1
2
3
4
5
6
7
8
Episodes

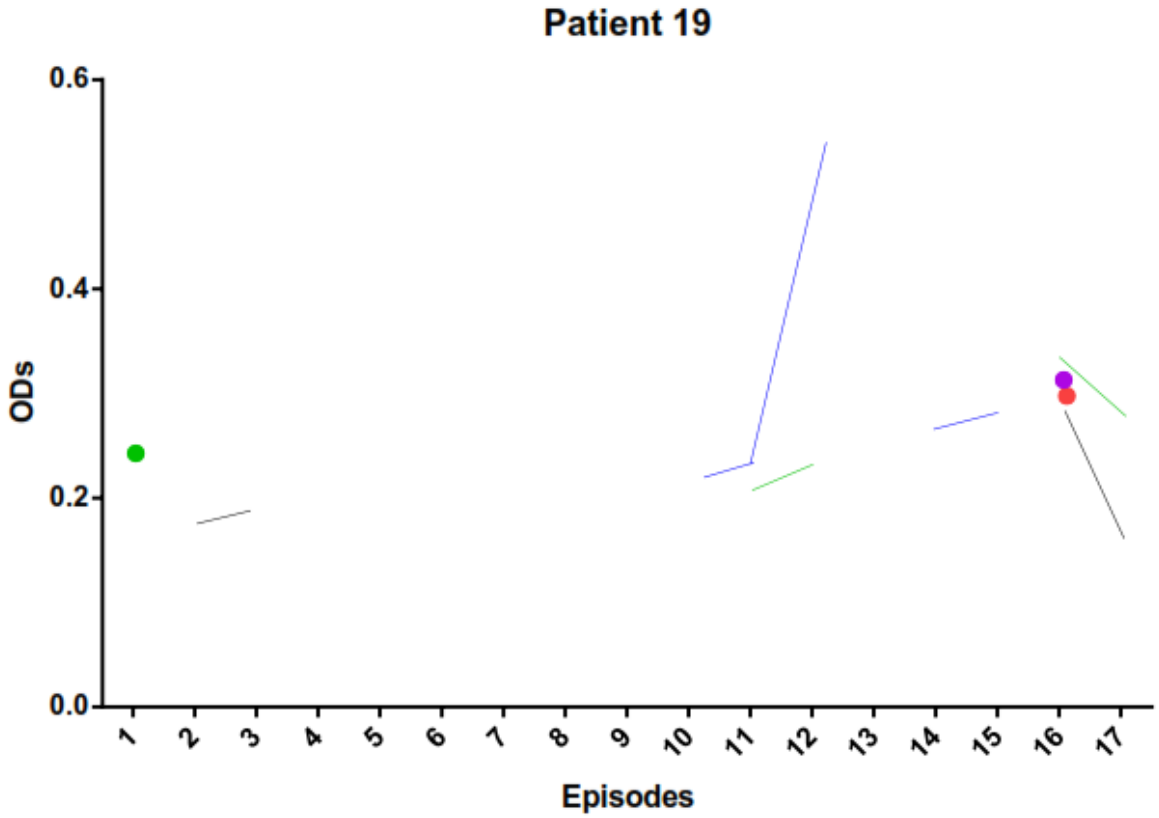

Patient 14
ODs
0.8
0.6
0.4
0.2
0.0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
Episodes

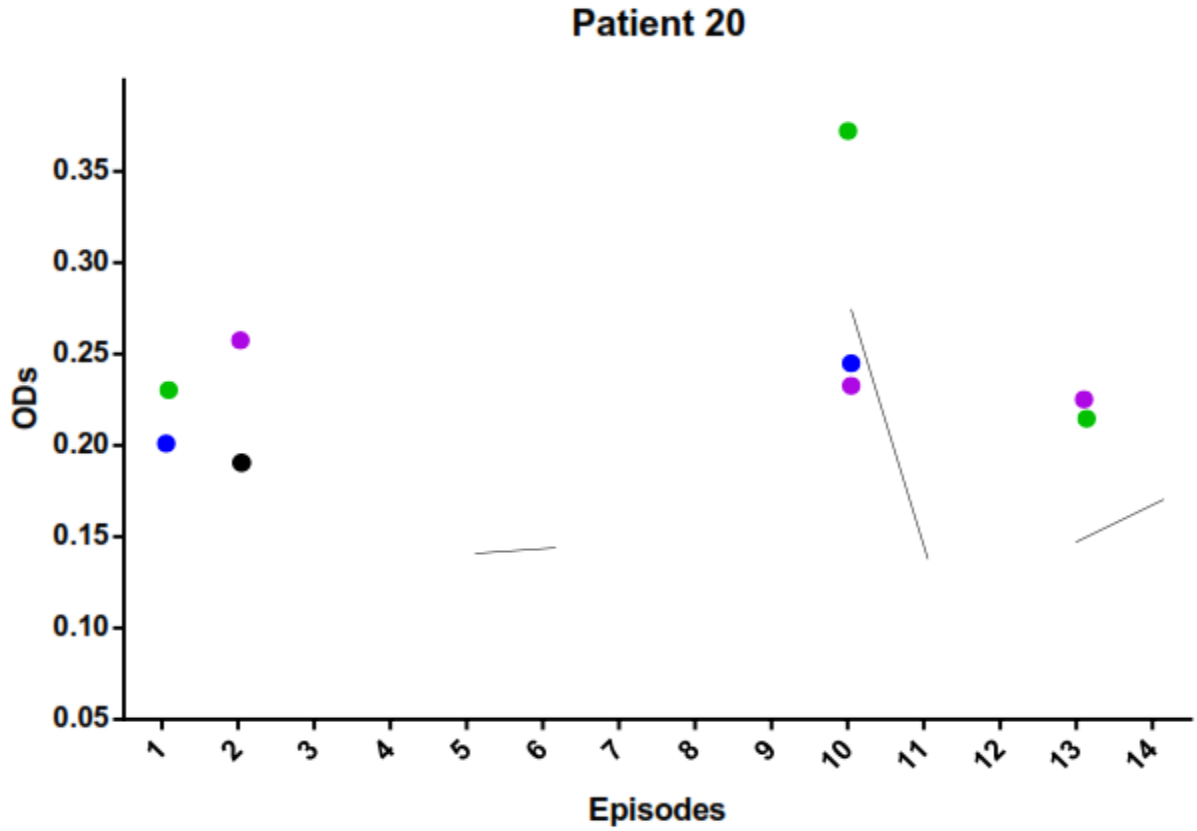

Patient 15
0.175
0.150
0.125
0.100
0.075
0.050
ODs
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
Episodes

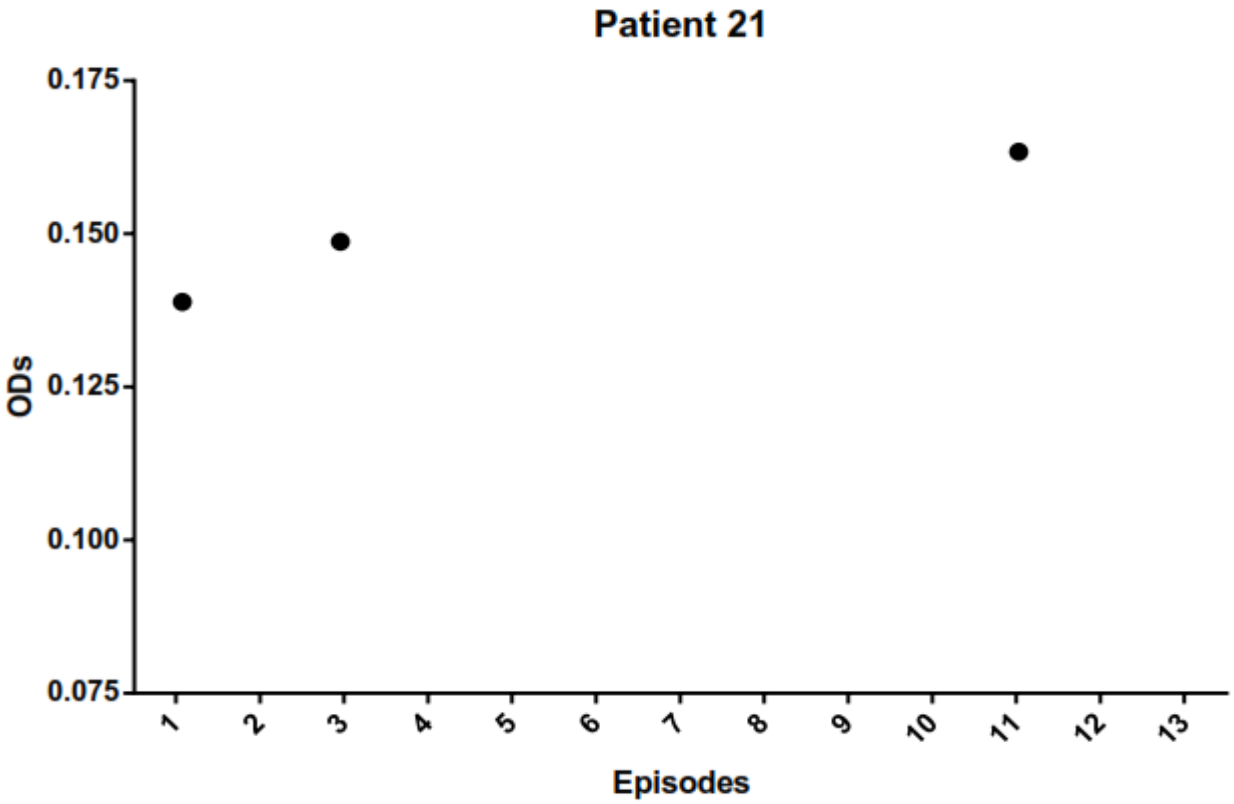

Patient 16
0.6
0.4
0.2
0.0
ODs
1 2 3 4 5 6 7 8 9 10 11 12 13 14
Episodes

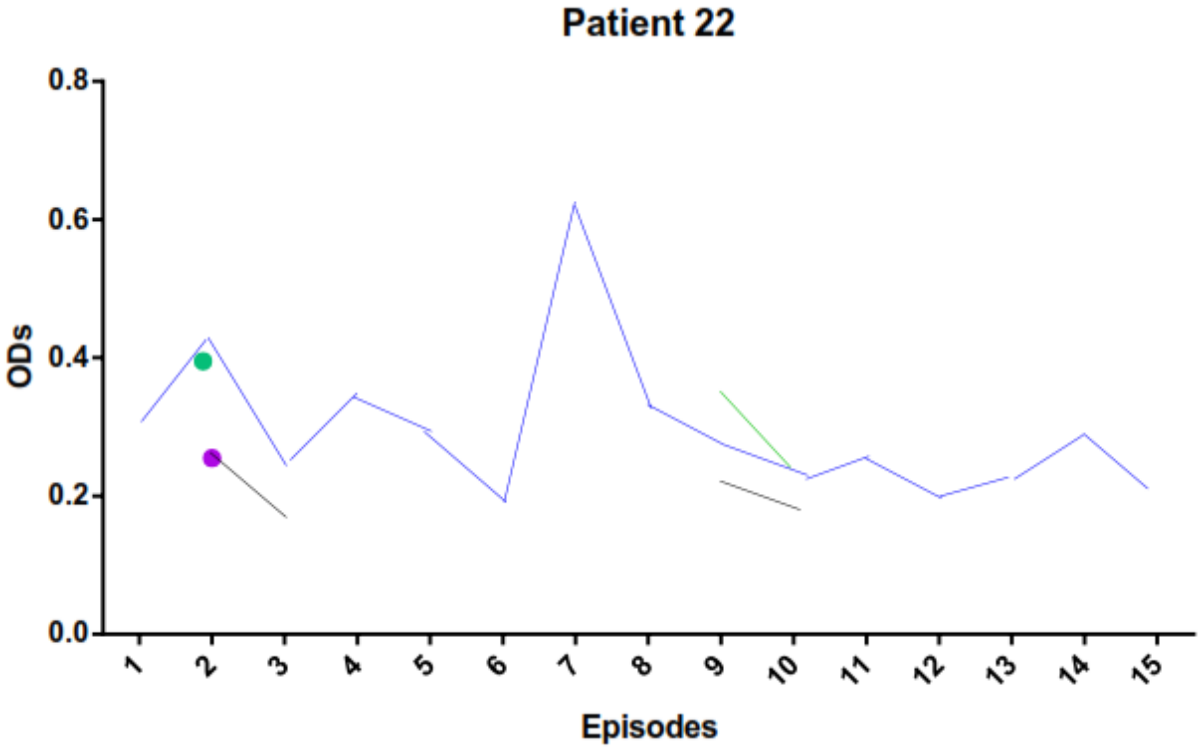

Patient 17
0.30
0.25
0.20
0.15
0.10
0.05
ODs
1
2
3
4
5
6
7
8
9
10
11
12
Episodes

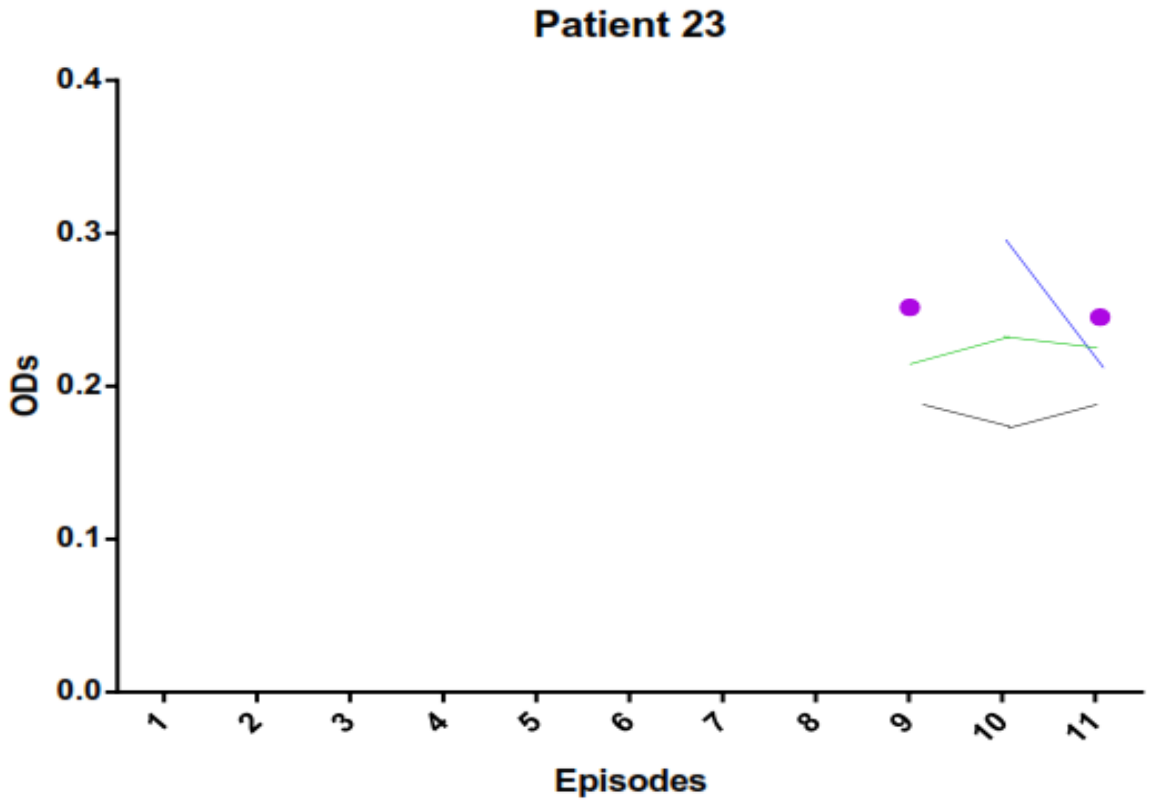

Patient 18
0.5
0.4
0.3
0.2
0.1
0.0
ODs
1
2
3
4
5
6
Episodes

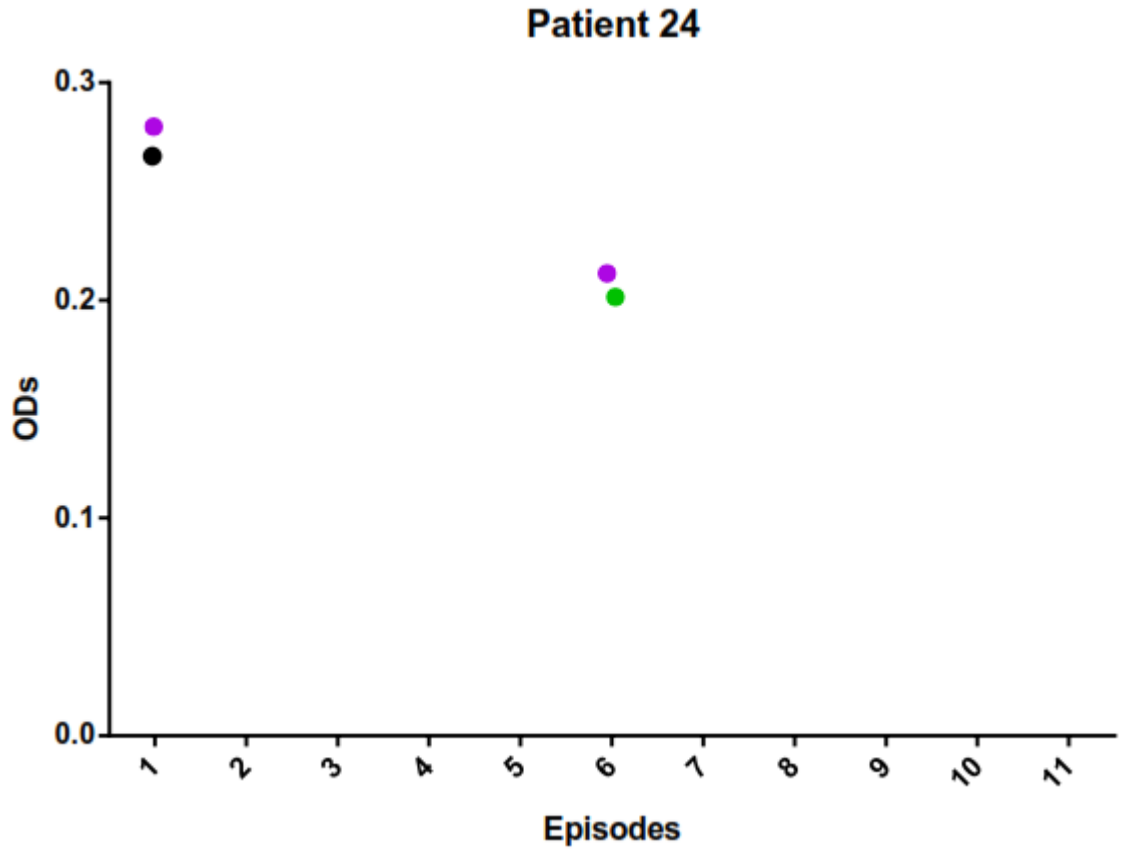

Patient 19
ODs
0.6
0.4
0.2
0.0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
Episodes

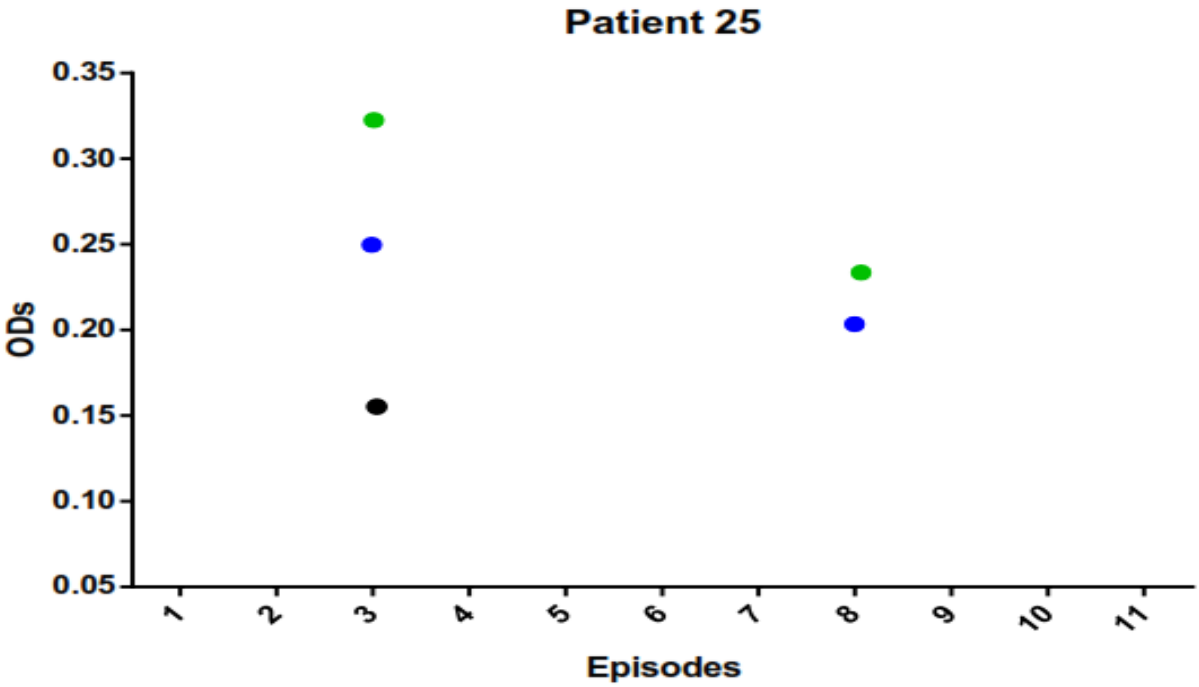

Patient 20
ODs
0.35
0.30
0.25
0.20
0.15
0.10
0.05
1
2
3
4
5
6
7
8
9
10
11
12
13
14
Episodes

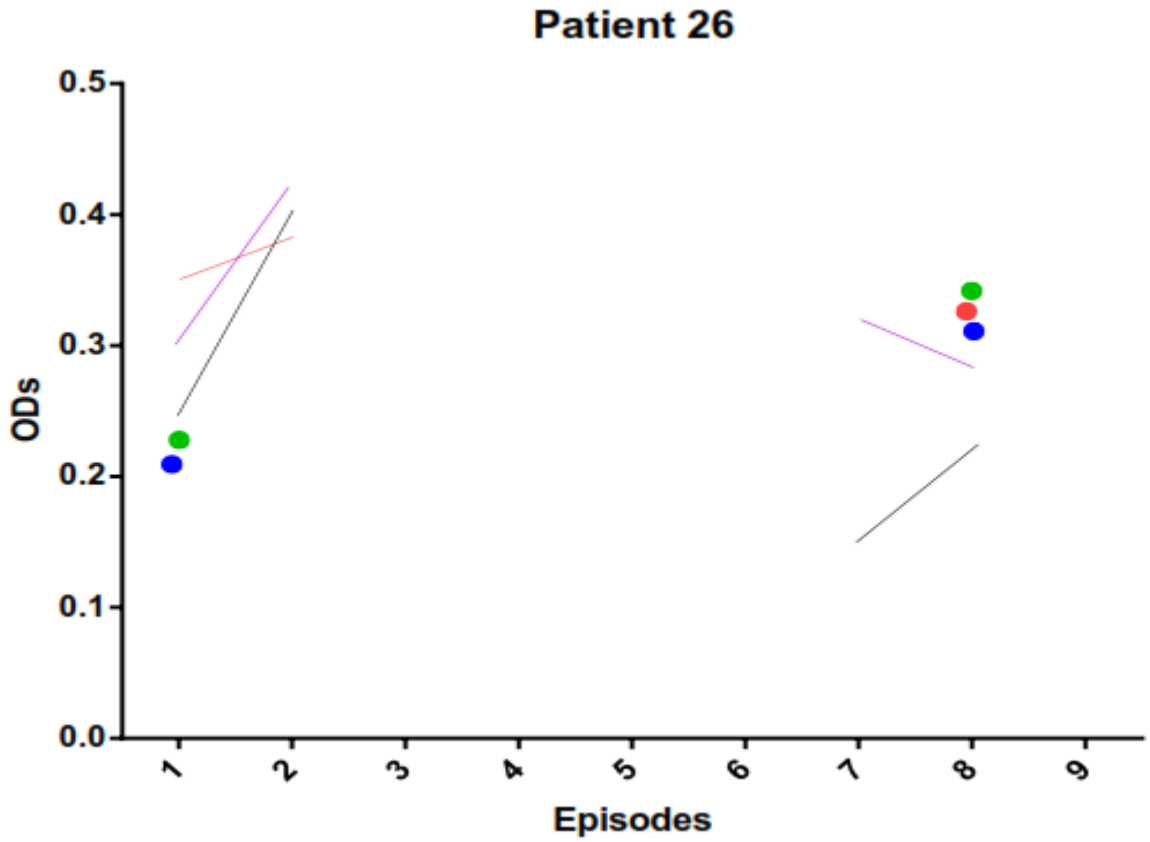

Patient 21
0.175
0.150
0.125
0.100
0.075
ODs
1
2
3
4
5
6
7
8
9
10
11
12
13
Episodes

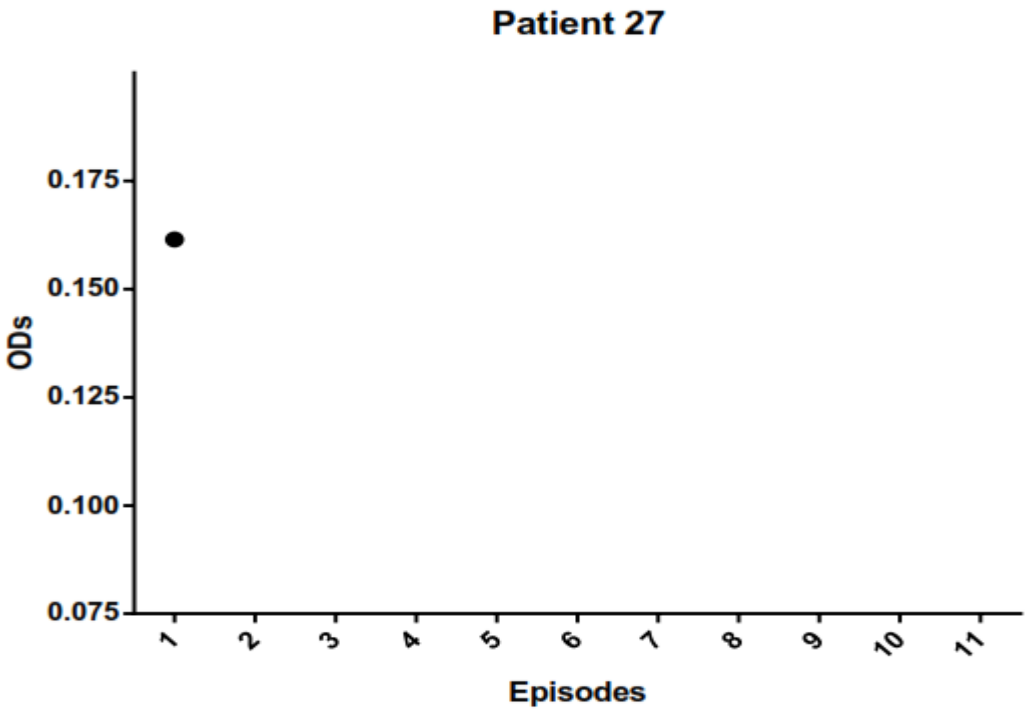

Patient 22
0.8
0.6
0.4
0.2
0.0
ODs
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
Episodes

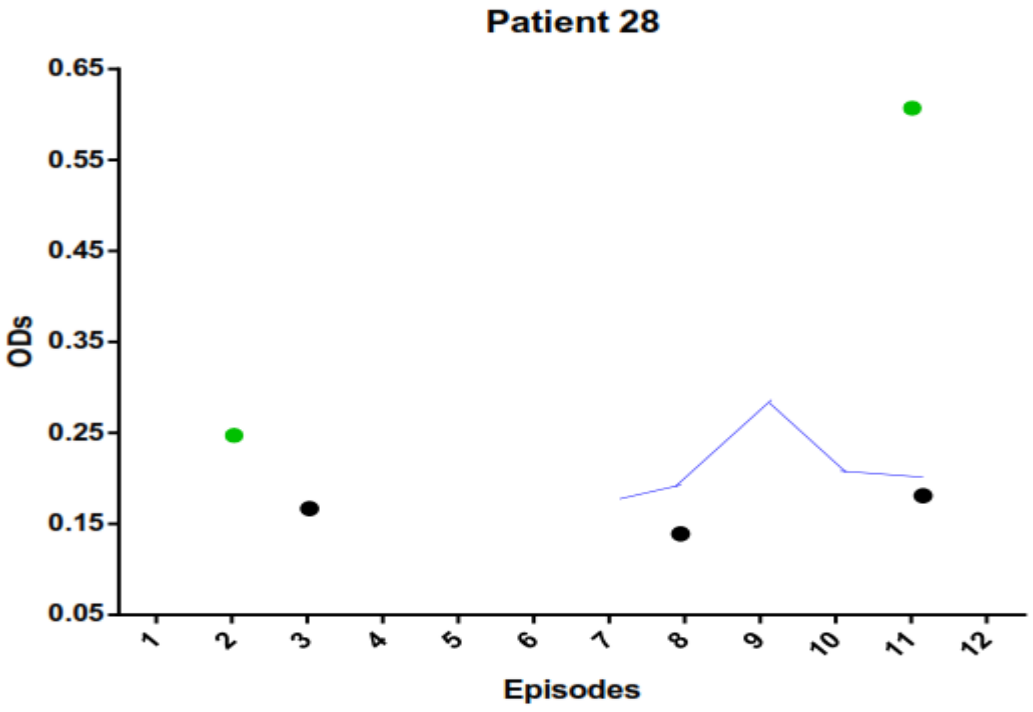

Patient 23
0.4
0.3
0.2
0.1
0.0
ODs
1
2
3
4
5
6
7
8
9
10
11
Episodes

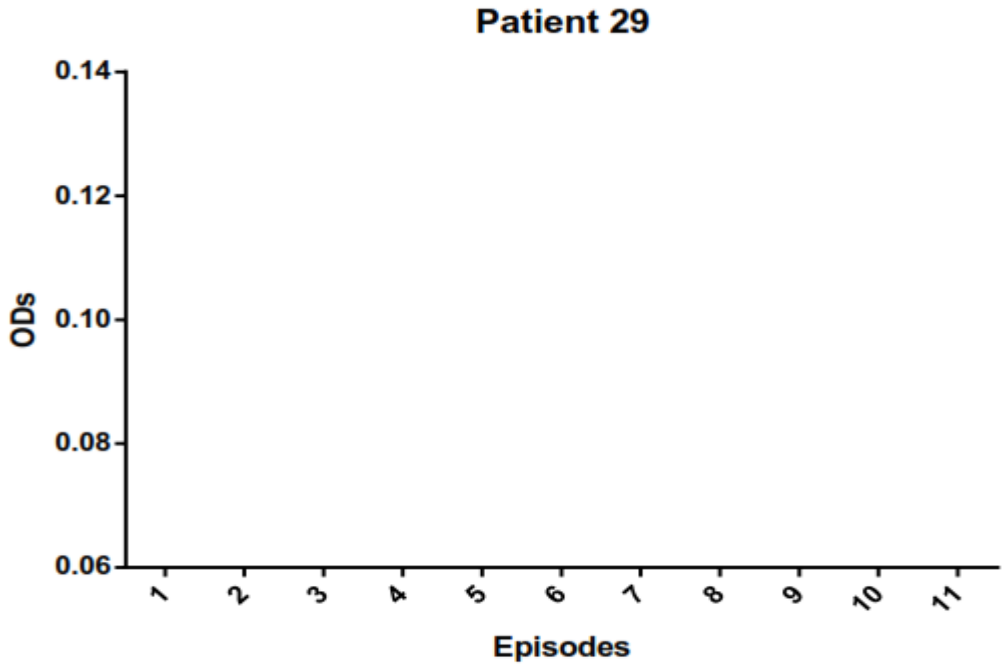

Patient 24
0.3
0.2
0.1
0.0
ODs
1
2
3
4
5
6
7
8
9
10
11
Episodes

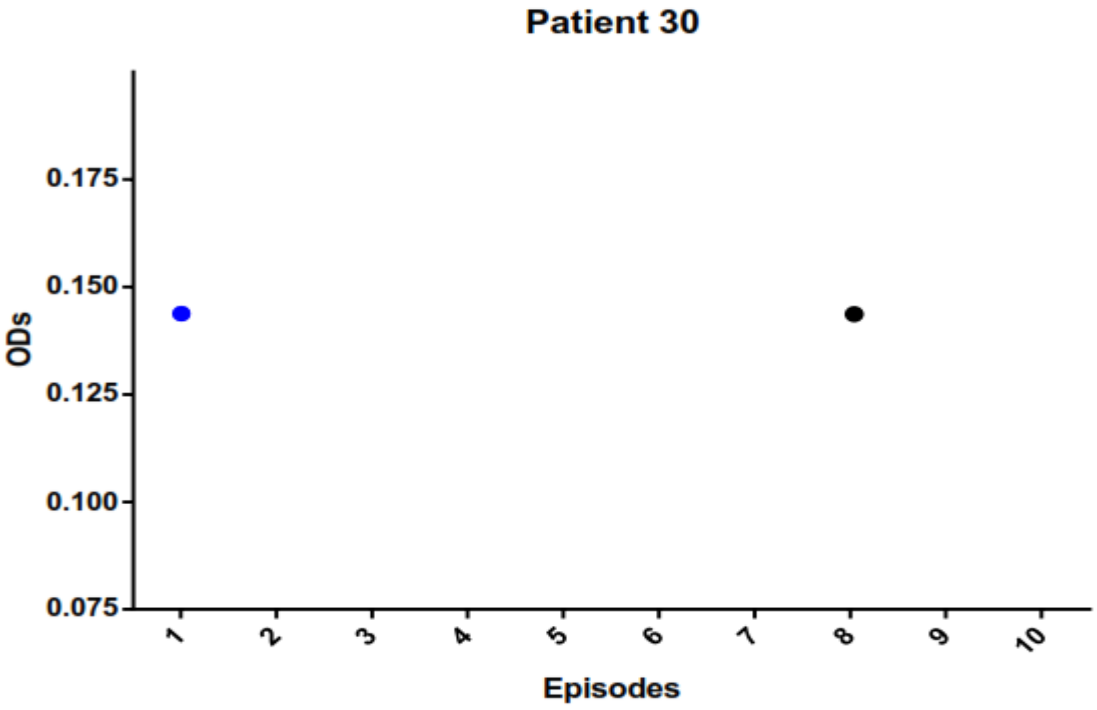

Patient 25
0.35
0.30
0.25
0.20
0.15
0.10
0.05
ODs
1
2
3
4
5
6
7
8
9
10
11
Episodes

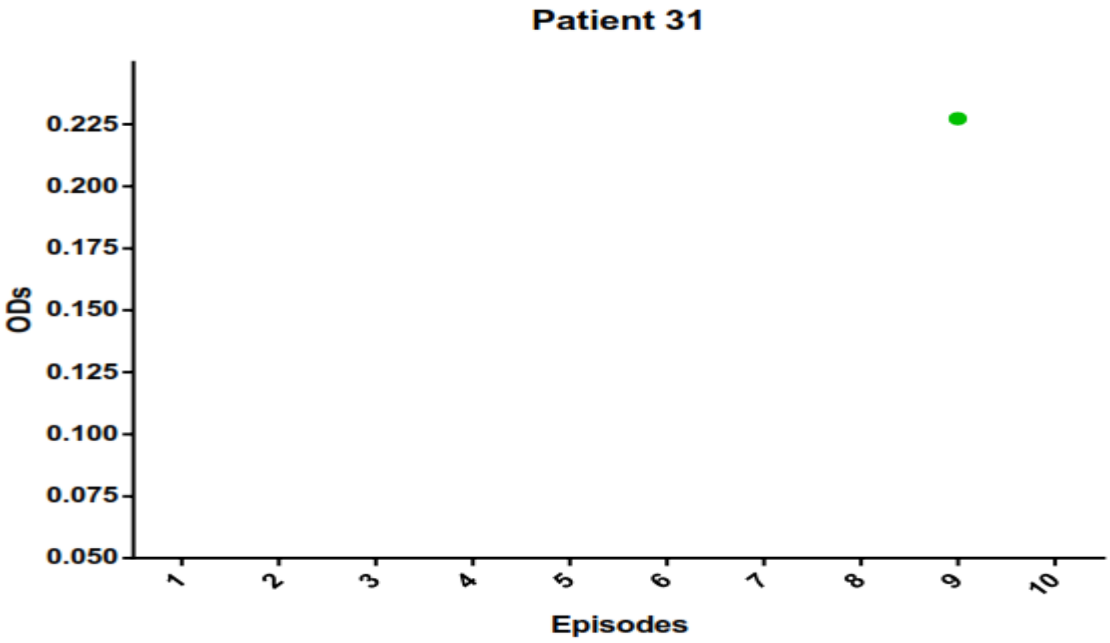

Patient 26
0.5
0.4
0.3
0.2
0.1
0.0
ODs
1
2
3
4
5
6
7
8
9
Episodes

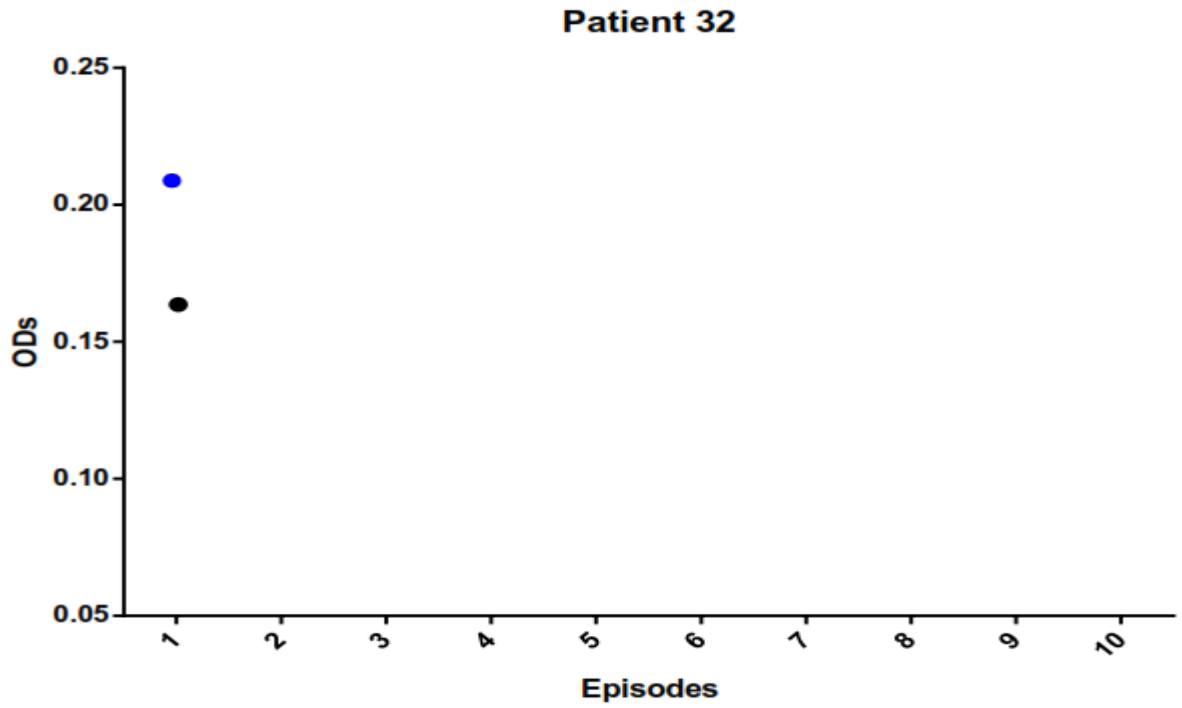

Patient 27
0.175
0.150
0.125
0.100
0.075
ODs
1
2
3
4
5
6
7
8
9
10
11
Episodes


Patient 28
0.65
0.55
0.45
0.35
0.25
0.15
0.05
ODs
1
2
3
4
5
6
7
8
9
10
11
12
Episodes


Patient 29
0.14
0.12
0.10
0.08
0.06
ODs
1
2
3
4
5
6
7
8
9
10
11
Episodes


Patient 30
0.175
0.150
0.125
0.100
0.075
ODs
1
2
3
4
5
6
7
8
9
10
Episodes





Trend of responses within an individual for each of the alleles. Colored lines indicate consecutive episodes where a child had responses to a given variant, colored dots indicate episodes with a response but neither the preceding norsubsequent episodes had responses for the given variant. Episodes are arranged numerically from the first to last for each individual. Color keys are; blue YHK, red HDN, purple HDK, black YHN, and green YDK.

Appendix 2 Comparison of ELISA and sequence data.

Patient	Sequence variant	Episode	YHK	HDN	HDK	YHN	YDK
1	HDN & YHN	E					
2	YHK	A					
	YHK & HDK	B					
	HDN & YHN	C					
	HDN & YHN	D					
	YHK & HDK	E					
	YHK & HDK	H					
	YHK	I					
	YHK & HDK	J					
3	YHK	A					
	HDK	F					
	YHK & HDK	G					
	YHK	I					
	YHK	J					
4	YHK & HDK	B					
	YHK & HDK	C					
	YHK & HDK	D					
5	YHK & HDK	B					
	YHK & HDK	E					
6	HDN	A					
	YHK	B					
	YHK & HDK	C					
	YHK & HDK	E					
7	YHK	C					
	YHK	D					
	YHK & HDK	H					
	YHK	I					
	YHK	J					
8	HDK	A					
	YHK & HDK	B					
	HDK	C					
9	HDK	A					
	HDN & YHN	G					
10	YHK	H					
11	YHK & HDK	F					
12	YHK & HDK	B					
	YHK & HDK	D					
	YHK	E					
14	YHK	A					
	HDK	B					
15	YHK & HDK	A					
	HDK	B					

Patient	Sequence variant	Episode	YHK	HDN	HDK	YHN	YDK
15	YHK & HDK	C					
	YHK & HDK	D					
	YHK & HDK	E					
	HDK	F					
	HDN & YHN	G					
16	YHK	C					
	YHK & HDK	D					
	YHK	E					
	YHK & HDK	F					
	YHK & HDK	G					
17	YHK & HDK	A					
	YHK & HDK	B					
	YHK & HDK	C					
	YHK & HDK	D					
	YHK	E					
	YHK	F					
	YHK	G					
	YHK	H					
	HDK	I					
	YHK	J					
	YHK & HDK	K					
	YHK & HDK	M					
	YHK & HDK	N					
	YHK	O					
18	YHK	C					
	HDN	D					
	HDN & YHN	E					
	YHK & HDK	F					
	HDN & YHN	G					
19	YHK & HDK	A					
	YHK & HDK	B					
	YHK	F					
	HDN & YHN	G					
	YHK	H					
	YHK	I					
20	YHK	G					
21	YHK & HDK	B					
	YHK & HDK	F					
22	YHK	A					
24	YHK & HDK	C					
	HDN & YHN	F					
	YHK	H					

Patient	Sequence variant	Episode	YHK	HDN	HDK	YHN	YDK
24	YHK	K					
25	YHK & HDK	A					
26	YHK	A					
27	YHK & HDK	M					
29	YHK & HDK	A					
	YHK & HDK	C					
	YHK & HDK	K					
30	YHK & HDK	A					

Comparisons of ELISA and sequence data.Episodes shown had at least one positive response using ELISA and had sequence data to make either of the 5 haplotypes. They are arranged alphabetically, with yellow representing episodes which showed a response to the particular variant. White represents episodes which were negative to the particular variant