IDENTIFICATION AND MOLECULAR CHARACTERIZATION OF SIMIAN MALARIA PARASITES IN WILD MONKEYS OF SINGAPORE

LI MEIZHI IRENE

NATIONAL UNIVERSITY OF SINGAPORE

2011

IDENTIFICATION AND MOLECULAR CHARACTERIZATION OF SIMIAN MALARIA PARASITES IN WILD MONKEYS OF SINGAPORE

LI MEIZHI IRENE

(*B.Sci. (Hons.)*), *NUS*

A THESIS SUBMITTED
FOR THE DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF EPIDEMIOLOGY AND PUBLIC HEALTH
YONG LOO LIN SCHOOL OF MEDICINE

NATIONAL UNIVERSITY OF SINGAPORE

2011

ACKNOWLEDGMENTS

I will like to thank the Environmental Health Institute, National Environmental Agency for the fund made available for this study. With special gratitude to my supervisors, ***Dr Vernon Lee***, ***Dr Ng Lee Ching***, ***Dr Indra Vythilingam*** and ***Prof Lim Meng Kin***, for their continuous support and encouragement throughout the Masters Program. I am also indebted to my mentor, ***Mr Wilson Tan***, for his technical assistance, advice and selflessness in coaching me throughout the project.

Last but not least, my sincere gratitude to the following, as the project will not be possible without them:

- Dr. William (Bill) Collins, Dr John W. Barnwell and Ms JoAnn Sullivan from the Centers for Disease Control and Prevention, USA, for their generosity in providing the simian *Plasmodium* controls.

- Dr Jeffery Cutter from the Communicable Diseases Division, Ministry of Health (Singapore), for granting the use and publication of the *P. knowlesi circumsporozoite* protein gene sequence of the two imported human knowlesi cases.

- Dr Kevin Tan from National University of Singapore for the provision of the *P. malariae* and *P. ovale* blood spots.

- Mr Patrick Lam from the Singapore Armed Forces for the provision of entomological surveillance data.

- Our collaborators – the Singapore Armed Forces, National Parks Board and the Agri-Food and Veterinary Authority.

- **TABLE OF CONTENTS**

SUMMARY

Plasmodium knowlesi is a simian malaria parasite currently recognized as the fifth cause of human malaria. Singapore reported its first local human knowlesi infection in 2007 and epidemiological investigations revealed that long-tailed macaques were the reservoir host of this blood parasite. Apart from *P. knowlesi*, long-tailed macaques are also natural host to *P. coatneyi, P. fieldi, P. cynomolgi* and *P. inui*, of which the latter two were also found to be infectious to humans under laboratory conditions. As there was no previous study of simian malaria parasites in Singapore's macaques, this study aims to determine their prevalence for the risk assessment of zoonotic transmission of simian malaria parasites to the general human population. Detection and accurate identification of simian malaria parasites through microscopy is typically challenged by low parasitemia, mixed species infection in the natural hosts and overlapping morphological characteristics among the different simian *Plasmodium* species. A sensitive *Plasmodium* parasite screening polymerase chain reaction (PCR) assay and a simian malaria species-specific nested PCR assay were thus developed. The PCR primers for *Plasmodium* parasites screening were designed against the conserved regions in the small subunit ribosomal RNA (SSU rRNA) genes. These primers were able to detect the four human and five simian *Plasmodium* species parasites, and could be used in both conventional and real-time PCR. The simian *Plasmodium* species-specific nested PCR assay, on the other hand, was developed using the *Plasmodium* circumsporozoite protein (*csp*) gene. *Plasmodium* screening on 65 peri-domestic and 92 wild macaques revealed that the former group was uninfected, while 71.7% of the sampled wild macaques were infected. Peri-domestic macaques were found in areas near human habitations while wild macaques were caught in military

forest where access is restricted to the general public. All five simian *Plasmodium* species were detected, with *P. knowlesi* having the highest prevalence (68.2%), followed by *P. cynomolgi* (60.6%), *P. fieldi* (16.7%), *P. coatneyi* (3.0%) and *P. inui* (1.5%). Co-infection with multiple species of *Plasmodium* parasites was also observed; double infection was detected in 23 (34.8%) macaques while five (7.6%) were infected with three *Plasmodium* species. Phylogenetic analysis of the non-repeat region of the *Plasmodium csp* gene from 15 infected macaques revealed high genotypic diversity of the parasites, reflecting a high intensity of malaria transmission among the macaques in the forest. On the other hand, all four local knowlesi cases had single *P. knowlesi* genotype which was identical to the *P. knowlesi* isolates of some macaques, suggesting that macaques were the reservoir hosts of the knowlesi malaria. Identical *Plasmodium csp* sequences shared by macaques caught at different timepoint also illustrates an ongoing sylvatic transmission. Despite these findings, the risk of zoonotic transmission of simian malaria parasites to the general population is assessed to be low as malaria parasites were absent among peri-domestic macaques, and all human knowlesi cases reported in Singapore were thus far occupational or travel related. However, to enable continuous risk assessment and surveillance, more studies will be required to determine the identity and distribution of the mosquito vector/s and the spatial distribution of the wild macaques.

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVATIONS

An	*Anopheles*
bp	base pair
CDC	Centre for Disease Control
csp	circumsporozoite protein
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
dH2O	deionised water
EDTA	ethylenediaminetetraacetic acid
$MgCl_2$	magnesium chloride
min	minute
ml	millilitre
mM	millimolar
ML	maximum likelihood
NJ	neighbour-joining
nm	nanometre
Pct	*Plasmodium coatneyi*
Pcy	*Plasmodium cynomologi*
Pf	*Plasmodium falciparum*
Pfi	*Plasmodium fieldi*
Pin	*Plasmodium inui*
Pk	*Plasmodium knowlesi*
Pm	*Plasmodium malariae*
Po	*Plasmdoium ovale*
Pv	*Plasmodium vivax*
PCR	polymerase chain reaction
rpm	round per minute
SSU rRNA	small sub-unit ribosomal ribonucleic acid
sec	second
WHO	World Health Organization
µl	microliter
µM	micromolar
°C	degree Celsius

CHAPTER ONE

General Introduction

1.1 Malaria

Malaria is an ancient disease, first described by ancient Egyptians in 1500B.C [1]. Despite years of intensive research, no successful vaccine for this disease has yet been developed, and it remains a serious public health problem in many tropical countries. According to World Health Organization (WHO), 225 million cases of malaria were reported in 2009, with a mortality of 781,000 [2]. In 2009, an estimated 1.3 billion people or 76% of the total population in Southeast Asian region were at risk of malaria [3].

Malaria is caused by protozoan parasites of the genus *Plasmodium*, family Plasmodiidae, suborder Haemosporidiidae, order Coccidia. Approximately 170 species of *Plasmodium* parasites, capable of infecting rodents, primates, reptiles and birds, have been discovered thus far [1, 4]. Five species of parasites, namely *P. falciparium, P. vivax, P. ovale, P. malariae* and *P. knowlesi* have been reported to cause disease in humans. *Plasmodium vivax* is the most widely distributed human malaria, while infection by *P. falciparium* is usually the most fatal. *Plasmodium knowlesi,* a simian malaria parasite originating from the Old World macaques, was recently incriminated as the fifth malaria species that infects humans.

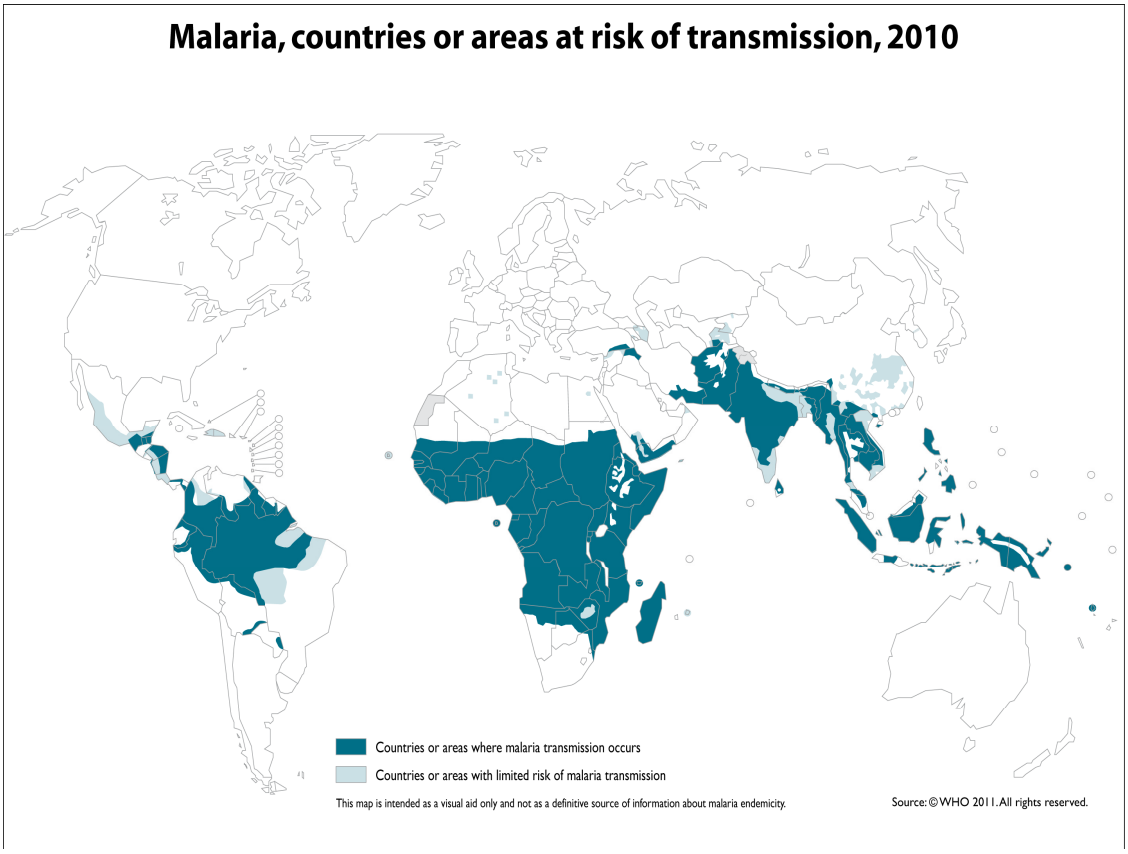

Figure 1.1: Global malaria situation, 2010 [5]

The classic clinical symptoms of malaria infection include intermittent fever, shivering, joint pains, headaches and repeated vomiting. If treatment is delayed, it can lead to severe complications such as renal failure, hypoglycemia, anemia, pulmonary edema, shock and coma, and eventually death [6].

1.1.1 Life cycle of malaria parasites

All malaria parasites require two hosts to complete their life cycle; the definitive invertebrate hosts and the intermediate vertebrate hosts. Most *Plasmodium* parasites are transmitted by mosquitoes, and those infecting human and non-human primates are transmitted exclusively by anopheline mosquitoes [4, 7].

Vertebrate hosts are infected through the bite of an infective mosquito when sporozoites are inoculated into the bloodstream during feeding (Figure 1.2). These sporozoites migrate to the liver and invade the hepatocytes, where they undergo an extensive replication known as primary schizogony, to produce exoerythrocytic schizonts (exoerythrocytic phase). Some species of *Plasmodium* parasites, such as *P. vivax*, *P. ovale*, *P. cynomolgi*, *P.fieldi* and *P. simiovale*, can produce a latent hepatic stage known as hypnozoites, which lay dormant in the liver for a period of time before invading the blood cells again [4, 7-10].

Each exoerythrocytic schizonts may contain 30,000 to 50,000 merozoites, which are released into the bloodstream where they invade the red blood cells (erythrocytic phase). In the erythrocytes, the merozoites undergo asexual development, forming

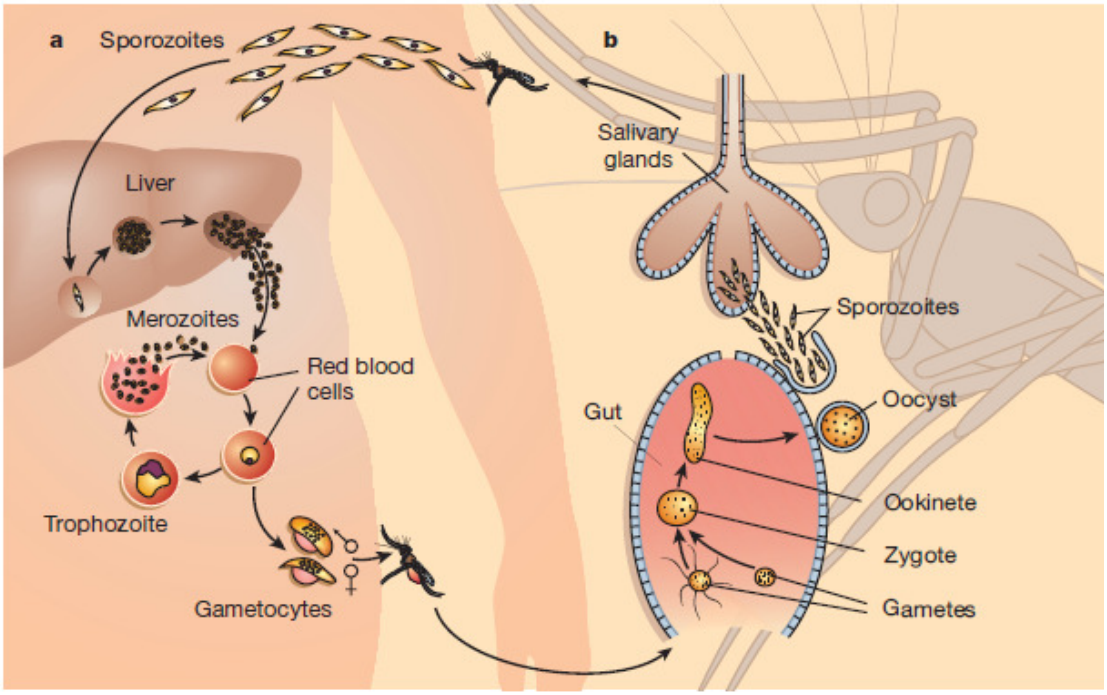

Figure 1.2: The life cycle of malaria parasite [11]

ring forms or early trophozoites, which will develop into mature trophozoites. These trophozoites then undergo schizogony, producing schizonts. The infected erythrocytes eventually lyze and merozoites are released into the blood stream. Some merozoites invade other erythrocytes and reinitiate another asexual erythrocytic cycle, while others differentiate into the microgametocytes (male) and macrogametocytes (female). The release of cellular contents from the ruptured erythrocytes triggers the host's immune system, resulting in clinical symptoms of fever and chills. Depending on the species of malaria parasite, the periodicity (time required to complete an erythrocytic cycle) ranges from 24 hours (quotidian periodicity) to 48 hours (tertian periodicity) or 72 hours (quartan periodicity) [7, 9].

The infection cycle in invertebrate hosts begins when it ingests both gametocyctes during its blood meal. The fall in temperature and presence of xanthurenic acid in the mosquito's gut trigger the development of the gametocytes to gametes. In the mosquito's midgut, the microgametes fuse with the macrogametes to form a zygote. Within 24 hours, the zygote differentiates into a motile and elongated ookinete, which then penetrates through the midgut epithelium and develops into an oocyst. Oocysts undergo sporogony (asexual multiplication in mosquito) and produce thousands of sporozoites. Eventually, the oocysts rupture, releasing the sporozoites which enter the haemolymph and subsequently migrate to the salivary gland. Inoculation of the sporozoites during blood feeding into a new vertebrate host perpetuates the malaria parasite's life cycle.

1.2 Non-human primate malarias

More than 20 species of simian malarial parasites that infect monkeys, apes and lemurs have been described (Table 1.1) [1, 7, 12]. These parasites, together with their natural hosts, can be found in the Asian, African, Central and South American region. Most of these parasites can be grouped with the four human malaria parasites based on the similarity of their erythrocytic cycle periodicity and morphology [7]. The distribution of simian malaria parasites affecting macaques in Southeast Asia was reported to follow the distribution of the *Anopheles leucosphyrus* group of mosquitoes (Figure 1.3) [12].

1.3 Simian malaria infections in man

Several studies had been conducted to test the infectivity of simian malaria parasites in man. The first experiment was carried out by Blacklock and Adler in 1922, using *P. reichenowi*, the simian form of *P. falciparium* [13]. However, the transfer of this simian malaria parasite species from chimpanzee to human volunteer using blood passage failed. The first reported successful experimental transmission was performed a decade later by Knowles and Das Gupta, who transmitted *P. knowlesi* to three human volunteers using blood inoculation [14]. The clinical symptoms observed ranged from mild, intermittent to severe fever. Unlike other human malaria infections, the fever of this simian malaria infection was observed to be of a daily remittent type. With the knowledge of *P. knowlesi* capable of inducing fever, this parasite was later used as a pyretic agent to treat patients with neuro-syphilis [15]. Other than *P. knowlesi*, the same author also successfully infected human volunteers with *P. inui* using blood passages in 1938 [16].

Table 1.1: List of non-human primate *Plasmodium* species, their periodicity, distribution and natural hosts [1, 7, 9, 12, 17]

***Plasmodium* species**	**Periodicity**	**Distribution**	**Natural Hosts**
P. knowlesi	Quotidian	Southeast Asia	Old world monkeys
P. cynomolgi*	Tertian	Southeast Asia, India, Sri Lanka	
P. coatneyi*	Tertian	Southeast Asia	
P. fieldi**	Tertian	Southeast Asia	
P. inui**	Quartan	Southeast Asia, India, Sri Lanka, Taiwan	
P. fragile*	Tertian	India, Sri Lanka	
P. simiovale**	Tertian	Sri Lanka	
P. shortii	Quartan	India, Sri Lanka	
P. gonderi*	Tertian	Africa	
P. petersi	Unknown	Africa	
P. georgsi	Unknown	Africa	
P. brasilianum**	Quartan	South America	New world monkeys
P. simium*	Tertian	Brazil	
P. eylesi*	Tertian	Southeast Asia	Gibbons
P. hylobati*	Tertian	Southeast Asia	
P. jefferyi*	Tertian	Southeast Asia	
P. youngi*	Tertian	Southeast Asia	
P. pitheci*	Tertian	Southeast Asia	Orang utans
P. silvaticum	Tertian	Southeast Asia	
P. schwetzi*	Tertian	Africa	Gorrillas, Chimpanzees
P. reichenowi*	Tertian	Africa	
P. rodhaini**	Quartan	Africa	
P. girardi	Unknown	Madagascar	Lemurs
P. foleyi	Unknown	Madagascar	
P. coulangesi	Unknown	Madagascar	
P. percygarnhami	Unknown	Madagascar	
P. uilenbergi	Unknown	Madagascar	
P. bucki	Unknown	Madagascar	
P. lemuris	Unknown	Madagascar	

"*", "**', "***", "****" indicates malaria parasites grouped under the *falciparum*-, *vivax*-, *malariae*- and *ovale*-type family, respectively [7]

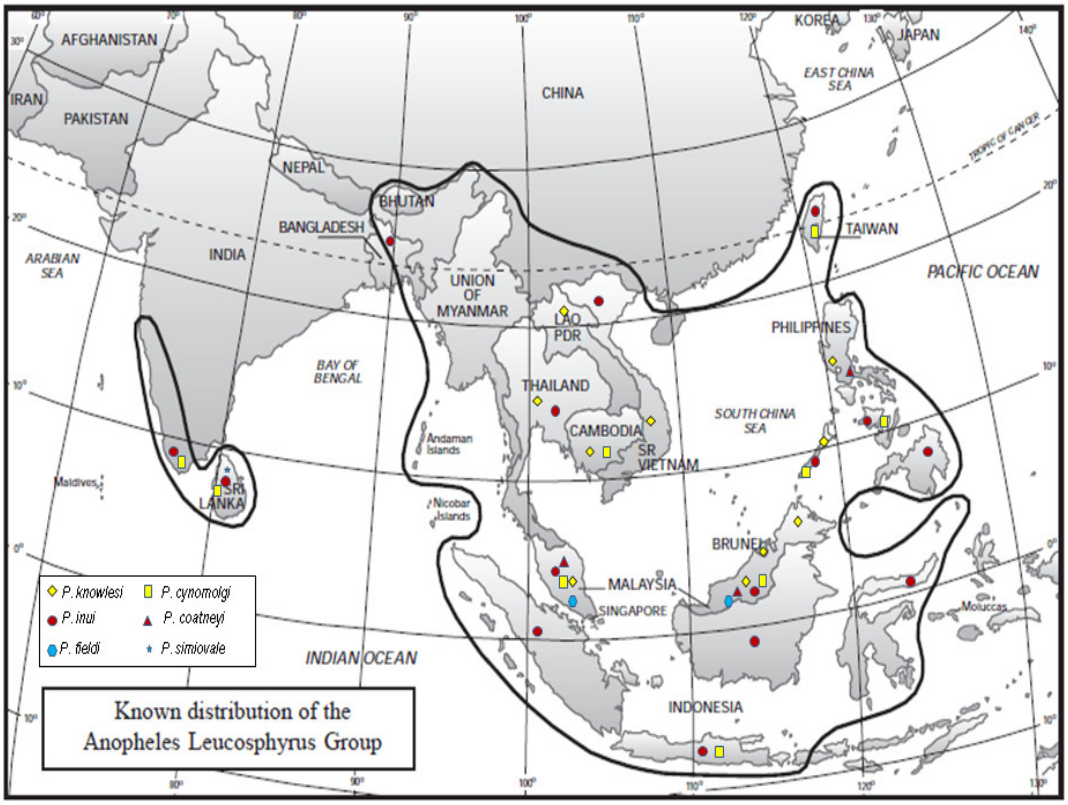

Figure 1.3: Distribution of simian malaria parasites in macaques [7, 12, 18-28] and the known limit of distribution of the *Anopheles leucosphyrus* sp. group of mosquitoes [29]

On the other hand, attempts to infect human with simian malaria parasites through mosquitoes were not successful [30, 31]. Hence, there was a general consensus that transmission of simian malaria parasites to humans was not possible. As such, non-human primate malaria was not taken into consideration during the strategic planning of malaria eradication during the World Health Assembly in 1955 [32]. In 1960, this dogma was proven wrong when reports of accidental human infection of *P. cynomolgi* by *An. freeborni* surfaced in two separate laboratories in the United States [33, 34]. These sparked the re-initiation of experimental mosquito transmission of simian malaria to man, and revealed the transmissibility of *P. knowlesi, P. inui* and *P. cynomolgi* from monkey to man, and from man to man through infectious mosquito bites under laboratory setting [35-39]. Likewise, other simian malaria parasites originating from apes and New World monkeys (*P. schwetzi*, *P. brasilianum, P. simium and P. eylesi*) were also proven to be transmissible to man [38, 40-42].

Substantial proof of natural infection of simian malaria parasites in man was only demonstrated in 1965 when Chin and co-workers reported a natural *P. knowlesi* infection in an American man who had spent nights working in a jungle in Pahang, peninsular Malaysia [43]. Surveillance studies in that locality revealed the presence of *P. knowlesi* in a sample of the wild macaque population there. However, when blood samples from residents in the area were pooled and injected into rhesus monkeys, a monkey species that typically does not survive *P. knowlesi* infections, none of these rhesus monkeys were infected. This large scale surveillance study concluded that human *P. knowlesi* infection was extremely rare. A few years later in 1971, another

presumptive case of natural human *P. knowlesi* infection was also reported in Johore, peninsular Malaysia [44].

The belief of human *P. knowlesi* infection being a rare incidence was overturned in 2004 when a large focus of human knowlesi infection was detected in the Kapit division of Sarawak, East Malaysia [45]. These cases were initially misdiagnosed as *P. malariae* using microscopy, although the symptoms were atypical of *P. malariae* infection and nested PCR failed to detect its DNA. Using molecular methods, 106 (51%) malaria cases in Kapit were attributable solely to *P. knowlesi* infection and 14 (7%) were co-infections of *P. knowlesi* and other human *Plasmodium* species. In contrast to the rare and sporadic reports of human *P. knowlesi* infection in the 1960s, this is the first report of a large focus of naturally acquired simian malaria infection in man. As *P. knowlesi* is morphologically similar to *P. falciparum* and *P. malariae* during the early ring stages and late trophozoites respectively, it is not possible to identify *P. knowlesi* parasites using microscopic observation of the thin blood film. Hence, Singh and co-workers designed a nested PCR assay for detection of *P. knowlesi* [45]. With the diagnostic test made available and an increased awareness of *P. knowlesi* as a possible cause of malaria in human, reports of naturally acquired human knowlesi cases surfaced in other parts of Southeast Asia: peninsular Malaysia [24], Singapore [46], Indonesian Borneo [47, 48], Sabah [49], Philippines [50], Thailand [51-53], Myanmar [54, 55], Vietnam [56, 57] and Cambodia[58]. The impact of *P. knowlesi* on travel medicine has also been recognised as non-endemic regions of the world, such as Europe [59-61], New Zealand[62], Australia [47] and the United States [63], reported importation of *P. knowlesi* cases from the Southeast

Asia region. Most significantly, fatalities due to *P. knowlesi* infections have been reported [64, 65]. The increased incidence of *P. knowlesi* infection and its associated fatalities prompted a synchronous echo from the public health community to relook into the impact of knowlesi malaria, and a classification of *P. knowlesi* as the fifth human malaria parasite [66-70]. However, unlike the other four human malaria parasite, *P. knowlesi* infection remains a zoonotic disease as there has been no evidence to suggest the occurrence of human-to-human transmission [25].

1.4 Detection and identification of simian malaria parasites

1.4.1 Microscopic observations

Microscopic examination of the Giemsa stained thin blood film is a universally accepted gold standard for primary identification of malaria parasites. It is also the main method for identification of *Plasmodium* parasites in non-human primates since the early 1900s [7, 9, 12, 18, 22, 23, 28, 71, 72]. However, there is an inherent difficulty in the accurate identification of simian malaria parasites due to overlapping morphological characteristics among these parasites [1]. Besides, individual macaques are often co-infected with two or more species of malaria parasites; and coupled with a low parasitaemia, microscopic identification of simian malaria parasites became confusing, inaccurate and insensitive [1, 7].

There are also shared morphological characteristics between simian malaria parasites and human malaria parasites. As a result, a significant proportion of the *P. knowlesi* cases in Kapit, Sarawak, were previously misdiagnosed as *P. falciparum* or *P.*

malariae using microscopy [45]. In addition, the morphology of *P. cynomolgi* and *P. fieldi* resembles that of *P. viva*x and *P. ovale*, respectively [27, 33], , and *P. inui* is reminiscent of *P. malariae* [7]. Hence, the accuracy of the identification using microscopy is greatly dependent on the experience of the microscopists.

1.4.2 Polymerase Chain Reaction (PCR) assays

Although microscopic examination of blood film remains the gold standard for malaria diagnostics, there is an increasing trend in using PCR to confirm the presence of malaria infection. As PCR can provide discriminatory power that could circumvent the limitations of identifying malaria parasites using microscopy, this method is frequently used when epidemiological and clinical findings do not match the microscopy results.

The nested PCR assay is a widely used method to detect the four human malaria parasites [73]. The nest one amplification reaction uses the *Plasmodium* genus-specific PCR primers, which amplifies all *Plasmodium* species' small subunit ribosomal RNA (SSU rRNA) gene. To determine the species of *Plasmodium* parasites present, the products of this nest one PCR reaction are subjected to four separate nest two amplification reactions, using primers specific for each human malaria parasite species. This assay is reported to have higher sensitivity than the conventional microscopy method [74].

The high sensitivity of the malaria-specific nested PCR assay allows the detection of malaria sporozoites in mosquitoes [75-77], and dried blood spots on filter papers [74, 78], making it useful for epidemiological investigation of malaria outbreaks and the detection of low-grade parasitaemia in high malaria endemicity areas [79]. This method has also been verified by the US CDC researchers to be the method of choice for detection of mixed malaria infections and sub-clinical infections [80].

Due to similarities in morphology between simian malaria parasites and that of humans, it is difficult to ascertain the occurrence of zoonosis through microscopy. Hence, cases of naturally-acquired human infection of simian malaria parasite may be overlooked. Nested PCR using *P. knowlesi*-specific primers played an important role in the discovery of a large focus of human knowlesi malaria, previously diagnosed as either *P. malariae* and/or *P. falciparum* cases. In the 1940s, Field illustrated an infection which he considered as an aberrant form of *P. vivax* in two patients from Malaysia [81]. Twenty years later, Sandosham *et al*. presented a slide of *P. cynomolgi bastianellii,* which had identical features to what Field had described [27]. Due to the close morphological similarity between these parasites, *P.cynomolgi* could be transmitted unknowingly to humans in nature. The development of simian malaria species specific PCR assay will hence aid in the confirmation of such zoonoses.

1.5 Malaria in Singapore

1.5.1 The historical perspective

Singapore attained the malaria-free status from World Health Organization (WHO) on 22 Nov 1982. However, the route to attaining the stature of malaria eradication is not without its labours. Singapore, like its neighbouring countries in Southeast Asia, was also once plagued with malaria.

Malaria was rampant in the early British colonial ruling days. In 1908, it was the second leading cause of death after tuberculosis. At the peak of an outbreak in 1911, about 20 deaths due to malaria were reported in a day. Hence, to bring the malaria epidemics under control, a comprehensive anti-malaria drainage system and oiling programme was introduced [82]. In 1966, malaria became a notifiable disease and all notified cases were investigated for epidemiological and entomological information. Legislation to control the breeding of *Anopheles* vectors was also tightened in 1968 [83].

However, rapid urbanization in the 1970s exacerbated the malaria problem in Singapore as land developments created favourable breeding grounds for the *Anopheles* vectors, and construction workers were mostly recruited from malaria-endemic countries. Despite precautionary measures to prevent *Anopheles* breeding and efforts to screen foreign workers for malaria parasites, malaria outbreaks still occurred. A revolutionary change in the strategy of malaria control in Singapore took place in 1975 when more aggressive efforts were taken to break the transmission

cycle. Vector surveillance and control was stepped up and maps of malaria sensitive areas were updated bi-yearly. Oiling programme was also extended to previously uncontrolled areas and areas with vector breeding were oiled frequently. Foreign workers' dormitories were also routinely sprayed with insecticide. This highly structured vector surveillance and control program nearly eradicated the malaria vectors. Whenever a malaria transmission is suspected, active case detection and mass blood surveys ensued until the reservoir of infection has been detected and treated. Vector control efforts such as larvicidal measures and residual spraying were also intensified. With this control strategy, the number of local malaria cases began to decline [83] (refer to Figure 1.4).

1.5.2 The current situation

Since attaining the malaria-free status, Singapore has maintained the standing for years without major local transmissions. Although malaria cases have been reported, more than 90% of these cases were contracted in Southeast Asia and the Indian subcontinent as most Singaporeans travelled to malaria endemic countries without taking adequate personal precautionary measures and chemoprophylaxis. Apart from local residents, work permit holders, student pass holders, foreigners seeking medical treatment in Singapore and tourists made up the rest of the overseas-acquired malaria cases. Most of these infections were caused by *P. vivax* (66%-78.4%), followed by *P. falciparum* (19.2%-31%)[84].

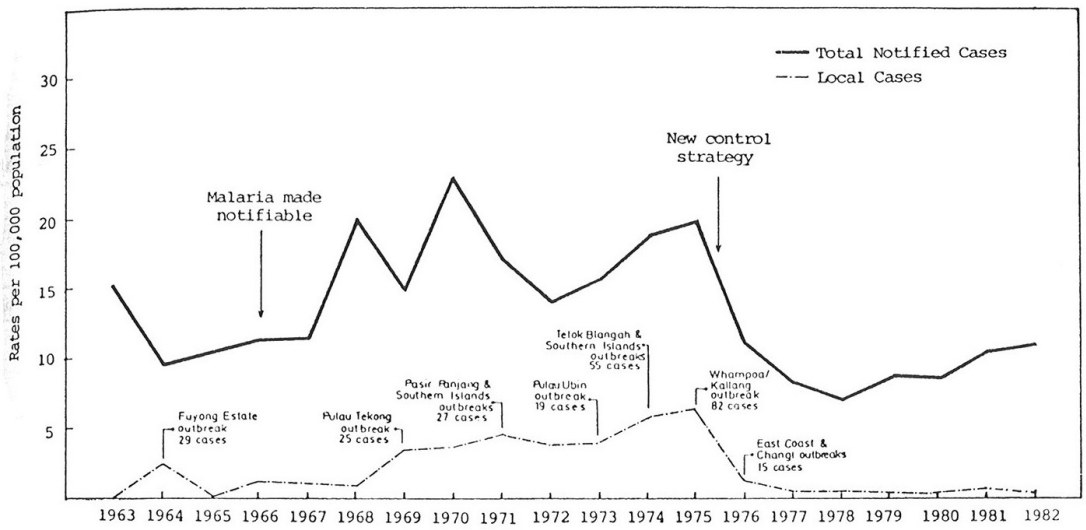

Figure 1.4: Malaria trend in Singapore, 1963 – 1982 [83]. The inception of new control strategy in 1975 reduced the number of local cases significantly.

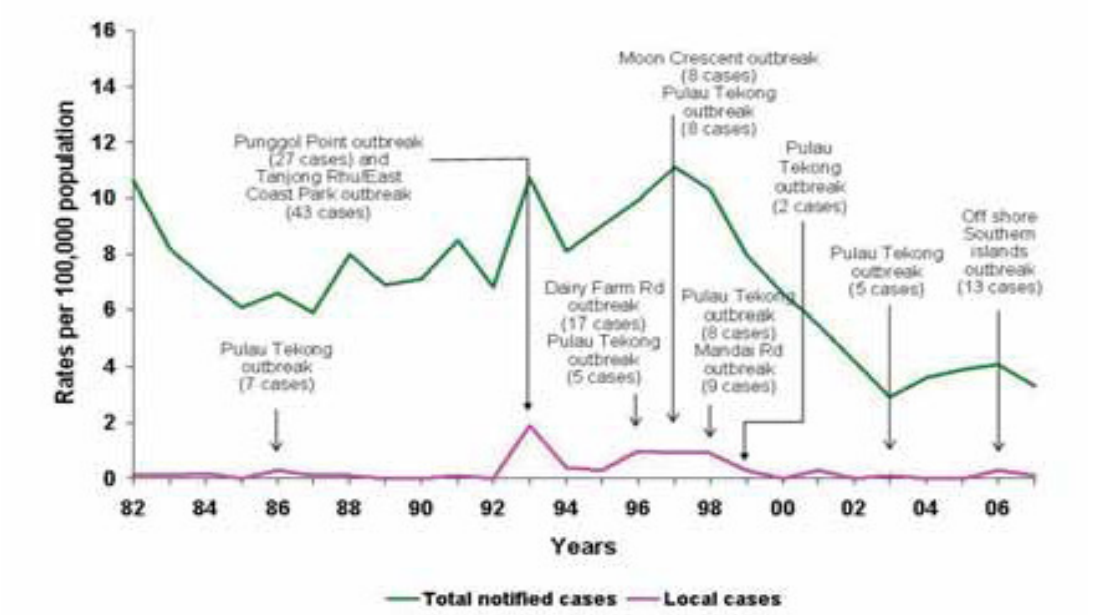

Figure 1.5: Malaria trend in Singapore, 1982-2006 [84]. Most cases were imported and only small localised outbreaks were reported.

With the influx of foreign labor from neighboring malaria-endemic countries and presence of pockets of *Anopheles* vectors, Singapore has not been spared from the occasional localised outbreaks of malaria. Between 1983 and 2009, 30 localised outbreaks involving a total of 220 cases were reported. These outbreaks include those that occurred in Punggol point, Tanjong Rhu/ East Coast Park, Dairy Farm, Mandai-Sungei Kadut, Jurong Island, Sembawang and Lim Chu Kang [84, 85]. All were eliminated through intensive epidemiological surveillance and vector control operations.

Apart from human malaria transmission, malaria parasites from the monkey reservoir too pose a threat to Singapore's malaria-free status. Singapore reported its first naturally-acquired human knowlesi malaria in 2007 [46]. The index case was a soldier who contracted *P. knowlesi* infection after a period of training in a forested area inhabited by the long-tailed macaque (*Macaca fascicularis*) in Lim Chu Kang, north-western Singapore. This prompted a fever monitoring and surveillance for soldiers who had visited the affected forest, which detected an additional five cases - four cases in 2007 and one in 2008 [20, 86]. All were military personnel who had no travel history, but had visited this restricted access forest prior to the onset of symptoms [20].

As long-tailed macaque, the natural host of *P. knowlesi*, is an inhabitant in this affected forest and various public nature parks, a joint operation was carried out by the Singapore Armed Forces, the National Parks Board and the National Environment Agency (NEA) to evaluate the risk of *P. knowlesi* infection in Singapore. Three long-

tailed macaques were sampled from the heart of the restricted-access forest and ten were sampled from a public nature reserve park. All three macaques from the restricted forest were infected with *P. knowlesi* while those from the nature reserve park were free from malaria infection. Phylogenetic analysis of the non-repeat region of the *P. knowlesi* circumsporozoite protein gene revealed shared genotypes between the human cases and the infected macaques, indicating that the cases had acquired the infection in the vicinity where these monkeys were found [20].

The finding of *P. knowlesi* in Singapore is of no surprise as this parasite was first discovered in India in 1931, from a long-tailed macaque imported from Singapore [14, 87]. The re-discovery of *P. knowlesi* parasites from long-tailed macaques 80 years later demonstrated the continuous and ongoing sylvatic transmission of *P. knowlesi* among the local long-tailed macaque population. The long-tailed macaque is the most predominant non-human primate in Singapore. Apart from *P. knowlesi*, this species of macaques is also known to harbor *P. cynomolgi*, *P. inui, P. fieldi* and *P. coatneyi* [7]. However to-date, there has been no reports on the prevalence of malaria in Singapore's macaques. Surveillance studies of natural incidence of simian malaria parasites in wild macaques had been conducted in Malaysia, Thailand, Indonesia, Cambodia, Philippines, Taiwan, Pakistan and Bangladesh [12, 23, 24, 28].

Detection and identification of simian malaria parasites by microscopic observation of the thin blood film has been stricken with difficulties and limitations, as previously described. Correct identification can be achieved with PCR assays using primers specific for each simian malaria parasite. These assays will also be useful in detecting

zoonoses in humans, which may be overlooked using microscopy, due to close morphology between simian and human malaria parasites.

1.6 Objectives of the study

The report of the locally-acquired knowlesi cases and the subsequent detection of *P. knowlesi* parasites in a sample of local wild macaques demonstrate a potential risk of zoonotic transmission of *P. knowlesi* in Singapore. However, as only a small sample of macaques was tested for *P. knowlesi* previously, there is a need to screen for simian malaria parasites in a larger population of macaques, preferably from different geographical locations, for a better understanding on the prevalence rate of malaria infection in local macaques. This is to enable a risk evaluation of zoonotic transmission of simian malaria parasites to the general human population.

The overall objective of this project is to identify the simian malaria parasites in Singapore's long-tailed macaques. Specifically, the study aims to:

1. Develop a simian malaria species-specific PCR assay to identify *P. knowlesi*, *P. cynomolgi*, *P. inui, P. fieldi* and *P. coatneyi* infections in long-tailed macaques,
2. Determine the prevalence of simian malaria parasites in Singapore's long-tailed macaque population,
3. Characterize the circumsporozoite protein *(csp)* genes of simian malaria parasites found in long-tailed macaques, and
4. Determine the molecular epidemiological linkage between the *P. knowlesi* isolated from Singapore's human cases and those isolated from local long-tailed macaques.

The information gathered from this study will not only constitue the first report of the prevalence of malaria infection in Singapore's macaques, but also help in expanding our current understanding on the epidemiology of *P. knowlesi* in Singapore.

CHAPTER TWO

Development of PCR assays for screening of simian malaria parasites

2.1 Introduction

Polymerase chain reaction assays are often used in malaria surveillance studies due to its ability to process large sample numbers and its higher sensitivity as compared to the microscopic examination of blood smears [74]. Although microscopy has been the gold standard for malaria diagnosis, it is time consuming to screen large number of samples using this method due to the preparation and interpretation of individual slides [88]. In addition, false negative results may occur while screening samples with low parasitemia [79, 89]. In view of this, a *Plasmodium* genus-specific nested PCR assay was developed by Singh and co-workers [74]. However, due to the need to perform two separate PCR reactions to confirm malaria infection, this assay can be time consuming, expensive and prone to PCR product carry-over contamination. To overcome this limitation, a sensitive *Plasmodium* genus-specific PCR assay (conventional and real-time format) using a single pair of primer was developed in this study.

A simian malaria species-specific PCR assay will also be developed for the identification of the five simian malaria parasites (*P. knowlesi*, *P. cynomolgi*, *P. inui, P. fieldi* and *P. coatneyi*) which long-tailed macaques are natural host to. As these five parasites have overlapping morphological characteristics at different life stages, their identification and differentiation using microscopy is impossible. On top of surveying simian malaria parasites in monkeys, this assay can also be used in the confirmation

of *P. knowlesi* infection in humans. The published *P. knowlesi*- specific PCR primers (Pmk8 and Pmkr9), was recently reported to exhibit stochastic cross amplification with *P. vivax* genomic DNA [90], resulting in misidentification of these two parasites in patients. Hence, the development of the simian malaria species-specific PCR assay will be useful in the differentiation of *P. knowlesi* and *P. vivax* infections in humans.

Apart from *P. knowlesi*, other simian malaria parasites, such as *P. cynomolgi* and *P. inui*, were also shown to be potentially infectious to humans [16, 33, 36, 37, 39, 91, 92]. The design of a simian malaria species-specific PCR assay will therefore be useful in the surveillance of these parasites in macaques for the risk assessment of potential zoonotic transmission of simian malaria parasites to the general human population. Moreover, it could also aid in the detection of naturally-acquired *P. knowlesi*, and possible *P. cynomolgi* and *P. inui* infections in humans.

2.2 Materials and methods

2.2.1 Source of *Plasmodium* DNA material for PCR assays development

Filter paper blood spots of *P. malariae* and *P. ovale* were acquired from the National Malaria Reference Centre, which was based in the Department of Microbiology, National University of Singapore prior to 2009. This centre is currently managed by the National Public Health Laboratory, Ministry of Health, Singapore. *Plasmodium falciparum, P. vivax and P. knowlesi* were obtained through the routine malaria diagnostic blood samples received by the Environmental Health Institute (EHI). Bioethics approval and informed consent from patients had been obtained for the use of these samples. Blood spots containing *P. coatneyi*, *P. cynomolgi, P. fieldi* and *P. inui* on the Isocode™ Stix (Krackeler Scientific, Inc., Albany, N.Y.) were obtained from the Laboratory Research and Development Unit (LRDU) of the Malaria branch, Division of Parasitic Diseases and Malaria, Centers for Disease Control & Prevention (CDC), Georgia, USA (Appendix A).

2.2.2 DNA extraction

2.2.2.1 Filter paper blood spots

DNA was extracted from dried filter paper blood spots using Instagene™ (Bio-Rad Laboratories, Hercules CA, USA) based on the method described by Cox-Singh *et al*. [78]. Two hundred microlitres of fully suspended Instagene™ matrix was added to a clean 1.5ml microcentrifuge tube using a large bore pipette tip. Two dried blood spots were clipped out using an ethanol flamed paper punch. The clippings were then added into the Instagene™ suspension. The tube was incubated at 56°C for 30min, with

vortexing for 10sec every 15min of incubation, before it was placed in a boiling water bath for 8min. It was then centrifuged at 12,000 rpm for 3min and the supernatant (containing the DNA) was decanted. The DNA template was stored at - 20°C until further use.

2.2.2.2 Blood spots on Isocode™ Stix

Extraction of DNA from blood spotted on Isocode™ Stix was carried out using protocol published by CDC's Division of Parasitic Diseases and Malaria [93]. One triangle of the dipstick was clipped off and transferred into a microcentrifuge tube and washed twice with 500µl of deionized sterile water (dH_2O) by vortexing three times for at least 5sec. After complete removal of dH_2O, the tube was briefly centrifuged and the residual water was pipetted off. Fifty microlitres of dH_2O were added and incubated at 95°C for 30min. Finally, the tube was gently tapped 20 times before the supernatant was transferred into a new microcentrifuge tube. The DNA template was then stored at - 20°C until further use.

2.2.2.3 Whole blood

DNA was extracted from 200 µl of whole blood (venous blood in EDTA anti-coagulant) using DNeasy® Blood and Tissue kit (QIAGEN, Hilden, Germany) according to the manufacturer's instructions. Briefly, 20 µl of Proteinase K was added into a 1.5ml microcentrifuge tube, followed by 200µl of the sample whole blood and 200µl of Buffer AL. The sample was vortexed before incubating at 56°C for 10min. Two hundred microlitres of molecular grade absolute ethanol was then added to the sample followed by vortexing. The entire mixture was pipetted into the DNeasy Mini spin column placed in collection tube. The column was centrifuged at 8,000rpm for a

minute. The flow-through and the collection tube were discarded. Five hundred microlitres of wash buffer AW1 was then added to the spin column coupled with a new collection tube, followed by centrifugation at 8,000rpm for a minute. The flow-through and collection tube were discarded and 500µl of the final wash buffer AW2 was added into the column, with a new collection tube. Final centrifugation at 14,000rpm at three minutes was applied to dry the membrane of the spin column. To elute the DNA, the column was transferred to a sterile 1.5ml microcentrifuge tube and 200µl of buffer AE was added directly onto the column membrane. The column was incubated at room temperature for a minute and finally spun at 8,000rpm for a minute to elute.

2.2.3 Development of *Plasmodium* genus-specific PCR assays

2.2.3.1 Design of *Plasmodium* genus-specific PCR primers

Sequences of the small subunit ribosomal RNA (SSU rRNA) genes of both sexual and asexual stages of human and simian *Plasmodium* species were retrieved from GenBank database. These sequences were aligned using the MegAlign software (DNASTAR, Lasergene, USA) and the *Plasmodium* genus-specific primers were designed based on the conserved regions of the gene. Figure 2.1 illustrates the alignment of the reference sequences and the selection of potential primer binding sites. All oligonucleotides (top-purified grade) were synthesized by a company specialized in oligonucleotide synthesis (AITbiotech Pte Ltd., Singapore). The oligonucleotide sequences are shown in Table 2.1. The theoretical melting temperature (T_m) for each primer was calculated using the basic Wallace rule [94]:

T_m (^{o}C) = 2°C(A+T) + 4°C(G+C)

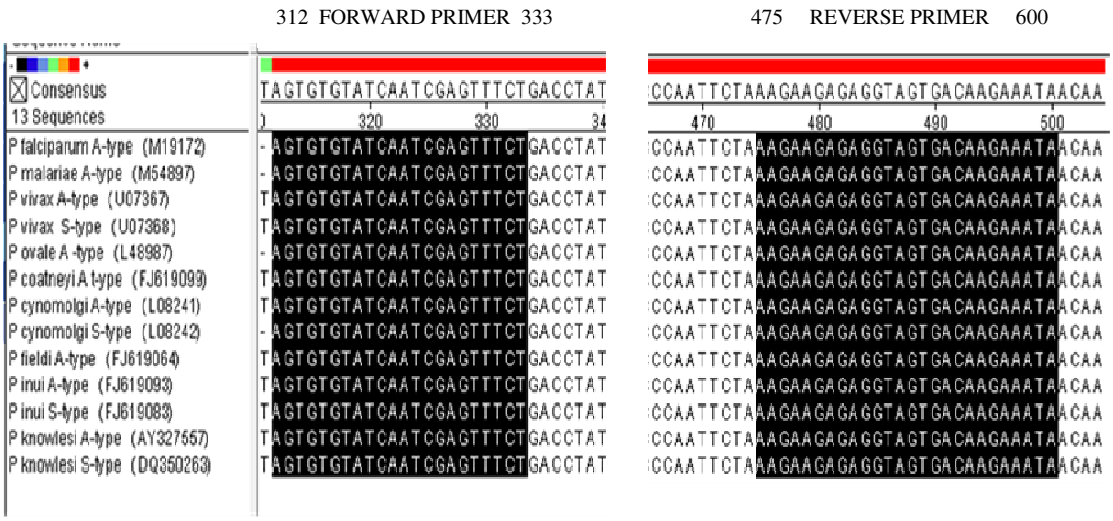

Figure 2.1: Alignment of SSU rRNA genes of the different *Plasmodium* species for design of the *Plasmodium* genus-specific primers.

Table 2.1: Oligonucleotide sequences of PCR primers designed for malaria parasite detection

Primer name	**Sequence**	**T_m (^oC)**	**Expected Product size (bp)**
PlasF	5'- AGTGTGTATCAATCGAGTTTCT -3'	44.9	188
PlasR	5'- CTTGTCACTACCTCTCTTCTTTAGA -3'	48.2	

2.2.3.2 Use of primers PlasF and PlasR for conventional PCR

To determine the optimum annealing temperature required for primers PlasF and PlasR, amplification was performed in a 50µl -reaction mixture (as illustrated in Table 2.2), using gradient conventional PCR (Veriti® Thermal Cycler, Applied Biostsyems, Foster City, CA USA). The temperature tested was between 51^{o}C to 61^{o}C, with an increment of 2^{o}C between each test temperature. The DNA templates used for testing were genomic DNA of *P. vivax*, with parasite count of 100, 0.3 and 0.06 parasite/µl. Separate reactions with genomic DNA of uninfected human and macaque samples were also included in the PCR optimization process. This was to test for any cross-reactivity of the primers with these DNA. The PCR parameters are listed in Table 2.3.

2.2.3.3 Comparison of sensitivity of detection with nested PCR assay

To determine and compare the sensitivity of the PCR assay, blood from *P. vivax* infected patient was used. The parasite density of this sample was determined by counting the number of parasites per 200 leukocytes. The parasite density was converted into parasites/µl, assuming a mean leukocyte count of 8000 [95]. Thereafter, the blood sample was diluted with malaria-free blood to obtain a theoretical parasite density of 100, 50, 25, 12.5, 6.25, 3.13, 1.56, 0.78, 0.39, 0.195, 0.0975, 0.0488, 0.024, 0.012, 0.006 and 0.003 parasites/µl of blood. The DNA of these serially diluted *P. vivax* blood samples were extracted according to the protocol

mentioned in Section 2.2.2.3, and tested with the primers PlasF and PlasR using the optimized annealing temperature.

Table 2.2: Components of "master-mix" for optimization of primers using conventional PCR

Components	**Final concentration**	**Volume (µl)**
RNase & DNase-free molecular grade water (Promega, Madison WI, USA)	-	18.75
5x reaction buffer, green (Promega)	1x	10.0
$MgCl_2$ 25mM (Promega)	2.5 mM	5.0
dNTP mix, 10mM each (Promega)	200 µM each	1.0
Forward primer (2.5 µM)	0.25 µM	5.0
Reverse primer (2.5 µM)	0.25 µM	5.0
GoTaq DNA polymerase, 5U/ µl (Promega)	1.25 U	0.25
DNA template	-	5.0
Total volume per reaction	**-**	**50.0**

Table 2.3: Cycling parameters for conventional PCR optimization

Steps	**Temperature/^{o}C**	**Time/s**	**No. cycles**
Initial denaturation	95	240	1
Denaturation	95	30	
Annealing	X	30	44
Extension	72	30	
Final extension	72	120	1
	20	~	

To compare the sensitivity of the optimized single round PCR assay with that of the published *Plasmodium* genus-specific nested PCR assay, the same panel of DNA was amplified using primers rPLU1 and rPLU5 followed by rPLU3 and rPLU4 as described by Singh *et al.* [74]. The PCR products were analyzed by electrophoresis in 2% agarose gel (1st Base Pte Ltd, Singapore), stained with Gel-Red™ (Biotium, Hayward, CA, USA), and observed under ultraviolet transillumination.

2.2.3.4 Use of primers PlasF and PlasR in real-time PCR assay

Real-time PCR using SYBR green method was carried out using LightCycler® 480 Instrument (Roche Diagnostics, Penzberg, Germany). The components of the real-time PCR mix and cycling parameters for the PCR program are listed in Table 2.4 and 2.5, respectively.

After PCR amplification, T_m curve analysis and melting temperature was performed using the LightCycler® 480 Melting Curve analysis software. The PCR products were heated to 95°C for 30sec and cooled to 60°C for 30 sec and then slowly heated back to 95°C at a rate of 2.2°C/sec. Obtained fluorescence signals are continuously monitored during the slow heating process. Plotting the fluorescence (F) versus temperature (T) generates the melting curve chart. Melting temperature was determined using the LightCycler® 480 Basic Software T_m calling analysis module by plotting a derivative melting curve (-dF/dT) where the center of a melting peak corresponds to the point of inflection. Amplification graphs were checked for the cross-point (Cp) value of the PCR product. The Cp value represented the cycle by which the fluorescence of a sample increased to a level higher than the background fluorescence in the amplification cycle.

Table 2.4: Components of "master-mix" for real-time PCR assay

Components	Final concentration	Volume (µl)
RNase & DNase-free molecular grade water (QIAGEN, Hilden, Germany)	-	3
2x Quantitect SYBR Green PCR Master Mix (QIAGEN)	1x	10.0
PlasF (10 µM)	0.50 µM	1.0
PlasR (10 µM)	0.50 µM	1.0
DNA template	-	5.0
Total volume per reaction	**-**	**20.0**

Table 2.5: Real-time PCR program for malaria screening using LightCycler® 480 Instrument

Program	Temperature/ ºC	Time/s	Cycle	Slope (ºC/sec)	Acquisition mode
Denaturation	95	900	1	4.4	None
Amplification	94	15		2.2	None
	50	30	50	2.2	None
	72	30		4.4	Single
Melting	95	30		4.4	None
	60	30	1	4.4	None
	95	0		2.2	Continuous
Cooling	40	10	1	2.2	None

2.2.3.5 Sensitivity and specificity of real-time PCR assay using primers PlasF and PlasR

The sensitivity of the primers PlasF and PlasR in real time PCR assay was determined based on both parasite density (Section 2.2.3.3) and parasite's SSU rRNA copy numbers (Section 2.2.3.6.5).

The specificity of the real-time PCR assay in detecting *Plasmodium* parasites was determined by comparing the amplification results obtained using the DNA of four human and five simian malaria parasites, with that of DNA from the non-infected human and macaque samples.

2.2.3.6 Preparation of plasmid standards for quantitative real-time PCR assay

2.2.3.6.1 Amplification of gene insert for plasmid standards

The gene insert for the control plasmid was a segment of the SSU rRNA gene, which encompassed the region amplified by the designed primers PlasF and PlasR. The primers used in the amplification of this gene insert are given in Table 2.6. Using protocol described in Section 2.2.3.2, PCR optimization for this set of cloning primers was conducted. Product from this amplification was subsequently cloned into a TOPO plasmid vector.

Table 2.6: Oligonucleotide sequences of PCR primers for amplifying the gene insert in control plasmids

Primer name	**Sequence**	**T_m (^oC)**	**Expected Product size (bp)**
CloningF	5' TATTAACTTAAGGAATTATAACAAAGAAG 3'	48.5	370
CloningR	5' ATACGCTATTGGAGCTGGAATTACCG 3'	59.7	

2.2.3.6.2 Cloning of PCR product

TOPO TA Cloning® Kit (Invitrogen, Carlsbad CA, USA) was used and performed according to the manufacturer's instructions. Briefly, the ligation reaction was carried out in a six microlitres reaction volume containing four microlitres of the PCR product, one microlitre of salt solution and one microlitre of TOPO® vector. The reaction mix was incubated at room temperature for 30min.

The chemically competent cells used for transformation were One Shot® TOP10 *E.coli* provided in the kit (Invitrogen, Carlsbad CA, USA). Two microlitres of the ligation reaction mix was added to the vial of competent cells and incubated on ice for 30min. For the heat shock procedure, the whole set-up was placed in a 42^{o}C waterbath for 30sec without shaking, and thereafter immediately put on ice for two minutes. To revive the cells, 250µl of SOC medium (Invitrogen, Carlsbad CA, USA) was added to the transformants and incubated at 37^{o}C with horizontal shaking at 200rpm. Two volumes of 50µl transformant culture were then spread on Luria-Bertani (LB) agar containing 2.5% (w/v) of LB broth, Miller (Amresco, USA) and 1.5% (w/v) of nutrient agar (Pronadisa, Spain), supplemented with 50 µg/ml Kanamycin (Invitrogen, Carlsbad CA, USA). The plates were incubated overnight at 37°C for bacterial growth.

Colony PCR was conducted using *Plasmodium* genus-specific primers PlasF and PlasR to screen the *E. coli* transformants for the gene insert. A single colony of the *E.coli* transformant was picked using a sterile 10 µl pipette tip and dipped into 20 µl PCR reaction mix containing 0.5 µM of each primer, 200 µM dNTP (Promega, WI, USA), 3mM $MgCl_2$, 1x reaction buffer and 0.5 units *Taq* DNA polymerase (Promega, WI, USA). Colony PCR was carried out with an initial denaturation at 95^{o}C for

10min, followed by 35 amplification cycles of 95^{o}C for 30sec, 57 ^{o}C for 30sec and 72 ^{o}C for 30sec. The final elongation step was 72 ^{o}C for four minutes. At the end of the last cycle, the temperature was reduced to 20°C. The PCR products were analyzed by electrophoresis in 2% agarose gel, stained with Gel-Red™ (Biotium, Hayward, CA, USA), and observed under ultraviolet transillumination.

2.2.3.6.3 Preparation of glycerol stocks

Escherichia coli colonies, which contained the plasmid construct with the gene of interest, were each inoculated into five millilitres of LB broth, containing 2.5% of LB broth, Miller (Amresco, USA), supplemented with 50µg/ml of Kanamycin (Invitrogen, Carlsbad CA, USA). The culture was grown at 37°C in a shaker incubator at 200rpm for at least eight hours for the subsequent preparation of glycerol stocks and plasmid extraction.

Glycerol stocks were prepared for long-term storage of the individual bacterial cultures at -80^{o}C. They were prepared by mixing 0.85ml of culture with 0.15ml of sterile glycerol (BDH, UK) in a 1.5ml microcentrifuge tube, followed by subsequent storage at -80°C.

2.2.3.6.4 Extraction of plasmid DNA

Extraction of plasmid DNA was carried out using PureLink™ Quick Plasmid Miniprep (Invitrogen, Carlsbad CA, USA) according to the manufacturer's protocol. An overnight broth culture of the transformants was pelleted in a 1.5ml microcentrifuge tube. It was then resuspended by vortexing in 250µl of Resuspension

Buffer with RNase A. To lyze the cells, 250µl lysis buffer was added and mixed gently by inverting the tubes, followed by incubation at room temperature for three minutes. To precipitate the lyzed bacterial cells, 350µl of Precipitation Buffer was added and mixed immediately by inverting the tubes until the solution became homogenous. The mixture was then centrifuged at 12,000rpm for 10min to clarify the lysate from lysis debris. The entire solution was then transferred onto the PureLink ™ Quick Plasmid Miniprep Spin Column and centrifuged at 12,000rpm for one minute. The flow through was discarded and the collection tube was re-used. Seven hundred microlitres of Wash Buffer was added onto the column and it was thereafter centrifuged at 12,000rpm for one minute. The flow through was discarded and the collection tube was re-used. To dry the column membrane, the column/collection tube was centrifuged at 12,000rpm for one minute. The collection tube was discarded and the column was transferred to a new 1.5ml microcentrifuge tube. Seventy-five microlitres of TE buffer was added directly to the membrane and the column/microcentrifuge tube was incubated at room temperature for one minute. To elute the plasmid DNA, the column/microcentrifuge tube was centrifuged at 12,000rpm for two minutes. The eluted plasmid DNA was sent for sequencing at 1st BASE Pte Ltd (Singapore), using BigDye Terminator Cycle Sequencing kit (Applied Biosystems, USA). The remaining plasmid was stored at -20°C for later use.

2.2.3.6.5 Dilution of stock plasmid for qPCR standards

The stock plasmid DNA concentration was determined by measuring the absorbance at 260nm (*A*260) in a spectrophotometer (GeneQuant Pro, GE Healthcare, UK). For reliable DNA quantification, *A*260 readings should lie between 0.1 and 1.0. An

absorbance of 1 unit at 260nm corresponds to 50µg plasmid DNA per ml. Serial dilution of the stock plasmids was performed to obtain plasmid standards containing 300,000, 30,000, 3,000, 300, 30, 3, 0.3, 0.03 and 0.003 copies of the gene insert. Firstly, the mass of single plasmid molecule was calculated using the formula:

Mass = plasmid size (bp) x $1.096e^{-21}$

Thereafter, the mass of plasmid DNA containing the required copy number of insert was calculated by multiplying the mass of single plasmid molecule with the required copy number. This calculated mass was divided by five microlitre (the volume transferred into each PCR reaction) to obtain the concentration of plasmid DNA needed to achieve the copy number of interest. With the concentration of the stock plasmid known after measurement by the spectrophotometer, serial dilution was conducted using TE buffer (QIAGEN, Hilden, Germany) to obtain the required plasmid DNA concentration of each copy number of interest, using the formular $C_1V_1 = C_2V_2$. The nine quantification standards (equivalent from 0.003 to 30,000 genome copies per µl) were run in triplicates using real-time PCR for generation of standard curve. Using LightCycler® 480 Basic Software, a standard curve and the PCR efficiency was automatically calculated and displayed.

2.2.4 Development of simian malaria species-specific nested PCR assay

2.2.4.1 Optimization of annealing temperature for nest one *Plasmodium* genus-specific primers

The simian malaria species-specific nested PCR assay was designed based on the *Plasmodium* circumsporozoite protein (*csp*) gene. The nest one PCR assay involves the amplification of the full *Plasmodium csp* gene using oligonucleotide primers

PkCSP-F [45] and the PKCSPR2 [20] (refer to Table 2.7). The amplified products have an approximate size between 1,000bp to 1,200bp. PCR amplification was carried out using high fidelity DNA polymerase in a 50µl reaction volume. The reaction mix contained five microlitre of DNA template, 200 µM of dNTP (Promega, Madison WI, USA), 1x Phusion® Flash PCR Master Mix (Finnzymes, Espoo, Finland) and 0.5µM of each primers. PCR optimization was performed with annealing temperature of 51°C to 61°C, with 2°C increment, using Veriti® Thermal Cycler (Applied Biostsyems, Foster City, CA USA). The cycling parameters were as followed: initial denaturation at 98°C for 10sec, followed by 44 cycles of 98°C for one second, different annealing temperatures for five seconds, and extension at 72 °C for 20 sec. Final elongation was at 72°C for two minutes.

DNA extracted from blood of malaria-free human and macaque samples were used as negative controls. PCR products were visualized in 2% agarose gel and the optimum annealing temperature was determined.

2.2.4.2 Nest two simian *Plasmodium* species-specific PCR assay

2.2.4.2.1 Cloning and sequencing of the simian malaria parasites' *csp* genes

Cloning and sequencing of the *csp* genes of the five simian malaria parasites were conducted to obtain the complete gene sequence for the design of simian malaria species-specific primers.

Cloning was conducted using Zero Blunt® PCR Cloning Kit (Invitrogen, Carlsbad CA, USA) and the cloning procedures described in Section 2.2.3.6.2 were used. For colony PCR of the *E. coli* transformants, PCR reaction mix and cycling parameters

Table 2.7: Oligonucleotide sequences of PCR primers used for amplifying the *csp* gene

Primer name	**Sequence**
PkCSP-F	5’ TCCTCCACATACTTAATACAAGA 3’
PKCSPR2	5’ TCAGCTACTTAATTGAATAATGC 3’

described in Section 2.2.3.6.2 were used, with PkCSP-F and PKCSPR2 as primers, and an annealing temperature of 55^{o}C and extension of one minute.

Positive clones were inoculated into five millitres of LB broth, containing 2.5% of LB broth, Miller (Amresco, USA), supplemented with 50µg/ml of Kanamycin (Invitrogen, Carlsbad CA, USA). The culture was grown at 37°C in a shaker incubator at 200rpm for at least eight hours. Plasmids of the individual bacterial culture were extracted using PureLink™ Quick Plasmid Miniprep (Invitrogen, Carlsbad CA, USA) as described in Section 2.2.3.6.4.

Sequencing of the *csp* gene was conducted by a commercial company, 1st BASE Pte Ltd (Singapore), using BigDye Terminator Cycle Sequencing kit (Applied Biosystems, USA), using primers M13F(-20) (5'-GTAAAACGACGGCCAGT-3') and M13R(-24) (5' GGAAACAGCTATGACCATG 3').

2.2.4.2.2 Circumsporozoite protein gene sequence analysis

The consensus sequence of *csp* genes from each simian *Plasmodium* species was obtained by assembling a contiguous sequence from the raw sequencing data using Seqman program (Lasergene, DNASTAR, USA). The 5' and 3' untranslated regions of the *csp* gene were removed to obtain the full coding sequence. As *csp* gene's internal repeat region is not useful for simian malaria species-specific primer design, only the 456 nucleotide residues encoding the non-repeat N-terminal (first 195 nucleotides of coding sequence) and C-terminal (261 nucleotides of coding sequence) of the gene [96] were used for sequence alignment and primer design.

2.2.4.2.3 Simian *Plasmodium* species-specific primer design

In addition to the sequences generated, the *csp* gene sequences of other simian malaria parasites published in GenBank were also used in primer design. All sequences were aligned using MegAlign software (Lasergene, DNASTAR, USA). The species-specific PCR primers were designed based on the highly variable regions of the gene. The primers for the five simian *Plasmodium* species are listed in Table 2.8. The primer binding sites of each primer are illustrated in Appendix B.

2.2.4.2.4 Optimization of nest two species-specific PCR assay

To determine the optimum annealing temperature for the species-specific primers, nest two PCR assays were carried out using gradient PCR, with annealing temperatures listed in Table 2.8. These primers were tested against the four human and five simian *Plasmodium* species to ensure their species specificity. Table 2.9 and 2.10 listed the nest two PCR reaction mix and cycling conditions used. All PCR products were analyzed by electrophoresis in 2% agarose gel, stained with Gel-Red™ (Biotium, Hayward, CA, USA), and observed under ultraviolet transillumination.

Table 2.8: Oligonucleotide sequences of primers and the range of annealing temperatures used for PCR optimization

***Plasmodium* species**	**Primer name**	**Primer sequence**	**Expected product size (bp)**	**Annealing temperature (^{o}C)**
P. coatneyi	CspCOAT-F1 CspCOAT-R1	5' – TTACCTACAGAAAATTAGATCTAC – 3' 5' – GCCCTAATGAATTACTCACAAA – 3'	238	58, 60, 62, 64
P. cynomolgi	CspCYNO-F2.1 CspCYNO-R2	5'–TCTACCATT(A/G)GC(G/A)(C/T)CGAGTGGAG – 3' 5' – AGGACTAACAATATGACTAGC – 3'	203	58, 60, 62, 64, 66
P. fieldi	CspFIELDI-F2a PKCSPR2	5' – GGTGACAAAAAACCAGATA – 3' 5'-TCAGCTACTTAATTGAATAATGC-3'	141	59, 61, 63, 65
P. inui	CspINUI-F2 CspINUI-R1	5' –CTTACCACCGAATGGAGTG – 3' 5'-AATAATGCTA(G/T)GACTA(G/A)CAATAT(T/G)ACTAC-3'	206	58, 60, 62, 64, 66
P. knowlesi	CspKnowlesiF PKCSPR2	5'- ACCTTGA(G/A)GTGGAAGCTTGTGT-3' 5'-TCAGCTACTTAATTGAATAATGC-3'	107	59, 61, 63, 65

Table 2.9: Components of "master-mix" for nest two PCR optimization

Components	Final concentration	Volume (µl)
RNase & DNase-free molecular grade water (Promega, Madison WI, USA)	-	7.9
5x reaction buffer, green (Promega)	1x	4.0
$MgCl_2$ 25 mM (Promega)	2mM	1.6
dNTP mix, 10 mM each (Promega)	200µM	0.4
Forward primer (2.5 mM)	0.25 µM	2.0
Reverse primer (2.5 mM)	0.25 µM	2.0
GoTaq DNA polymerase, 5U/ µl (Promega)	1.25U	0.1
DNA template	-	2
Total volume per reaction	**-**	**20.0**

Table 2.10: Cycling parameters for nest two PCR

Steps	Temperature/^{o}C	Time/s	No. Cycles
Initial denaturation	95	240	1
Denaturation	95	30	
Annealing	X	30	44
Extension	72	30	
Final extension	72	120	1
	20	~	

2.3 Results

2.3.1 Use of primers PlasF and PlasR for conventional PCR

Primer set PlasF and PlasR was able to detect DNA extracted from *P. vivax* of different parasite load, at all the tested annealing temperature. No amplification was observed with DNA from non-infected samples. All PCR reactions with *P. vivax* DNA yielded the expected product size of 180bp (Figure 2.2). Although stochastic cross-reaction was seen with malaria-negative macaque DNA at lower annealing temperature, the product size of this non-specific PCR reaction was incorrect, and this stochastic cross-reaction was abolished at higher annealing temperature.

2.3.2 Comparison of sensitivity with nested PCR

The *Plasmodium* genus-specific nested PCR assay designed by Singh and colleagues was reported to have a sensitivity of at least six parasites/µl of blood using DNA extracted from bloodspot [74]. Using DNA of *P. vivax* diluted to different parasitemia, the nested PCR assay produced a constantly high intensity of specific product band (as observed on gel electrophoresis) regardless of the parasite count in the samples, while the single run PCR showed a concentration-dependent intensity of the PCR band (Figure 2.3). The nested PCR assay was able to detect up to 0.006 parasites/µl whereas the single conventional PCR assay using the newly designed primers PlasF and PlasR was able to detect up to 0.003 parasites/µl (Figure 2.3).

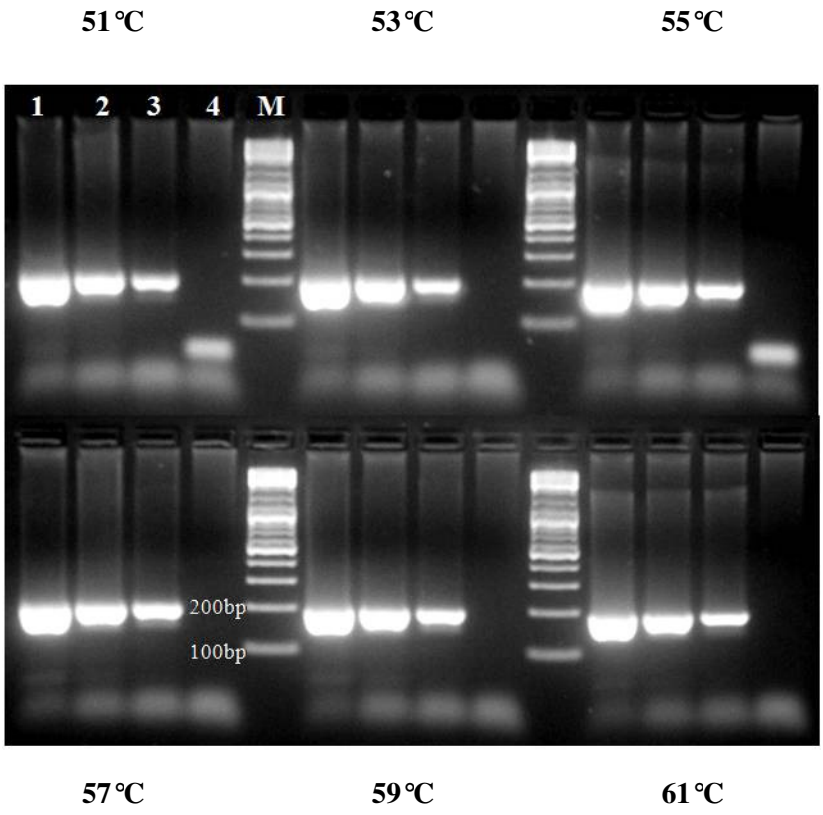

Figure 2.2: PCR optimization of primer set PlasF and PlasR. Lane numbers (1 to 4) represent *P. vivax* of parasitemia 100p/µl, 0.3 p/µl, 0.06 p/µl, and malaria-negative macaque sample, respectively. Molecular size markers (100-basepair ladder) are marked in lane M.

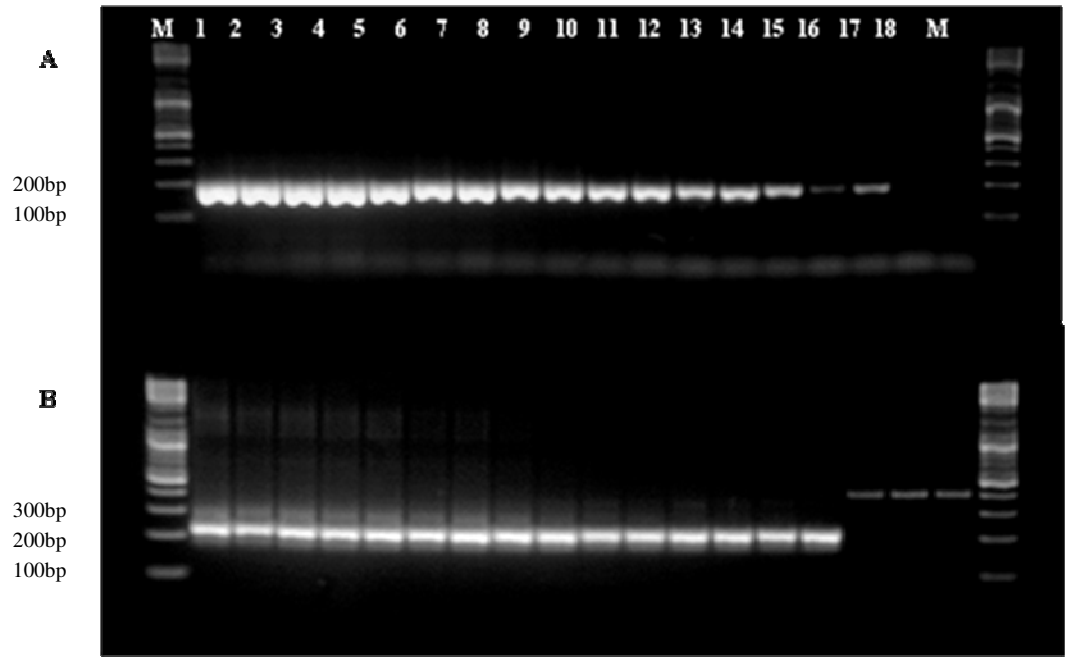

Figure 2.3: Comparison of sensitivity between single conventional PCR run using PlasF and PlasR (A) and the published nested PCR (B) in *Plasmodium* parasite detection. Molecular size markers are in lane M. Lane 1 to 18 are PCR amplification products using *P. vivax* of the following parasitemia (parasites/ µl): 100 (lane 1), 50 (lane 2), 25 (lane 3), 12.5 (lane 4), 6.25 (lane 5), 3.13 (lane 6), 1.56 (lane 7), 0.78 (lane 8), 0.39 (lane 9), 0.195 (lane 10), 0.0975 (lane 11), 0.0488 (lane 12), 0.024 (lane 13), 0.012 (lane 14), 0.006 (lane 15), 0.003 (lane 16), malaria-negative human sample (lane 17), and malaria-negative macaque sample (lane 18).

2.3.3 Sensitivity of real-time PCR assay

Real-time PCR using DNA of *P. vivax* with parastemia ranging from 0.003 to 100 parasites/µl (Figure 2.4) was conducted to compare the sensitivity with the single conventional PCR run. Both real-time and conventional PCR assays were observed to have comparable sensitivity of detecting at least 0.003 parasites/µl of blood (Table 2.11). In terms of sensitivity level in copy numbers, the real-time PCR assay was able to detect up to 0.3 copies/µl (Figure 2.5 and Table 2.12). However, only one out of the triplicates was found to be positive. No amplification was detected using 0.03 and 0.003 copies/µl.

The slope of the standard curve describes the kinetics of the PCR amplification i.e. how fast the target DNA can increase with the amplification cycles (an indication of PCR efficiency). A perfect amplification will produce a standard curve with efficiency value of "two", denoting that the amount of product doubles with each PCR cycle. The standard curve generated by the real-time PCR assay using primers PlasF and PlasR had an efficiency of 1.854 (Figure 2.6). This translates to an efficiency of 92.7%, which is within the acceptable range of 90% to 100% [97].

The error value (mean squared error of the single data points fit to the regression line), is a measure of the accuracy of the quantification result based on the standard curve. The error value of the standard curve produced is 0.013, which was within the acceptable range of less than 0.2 [98].

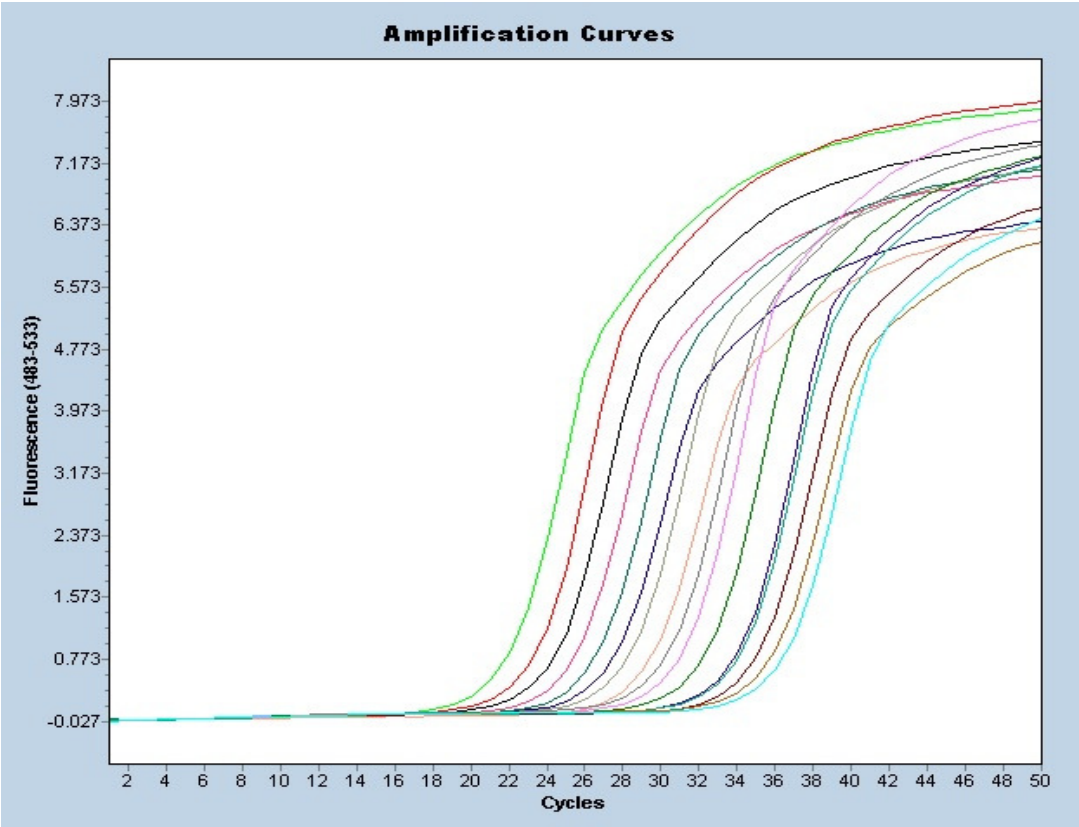

Figure 2.4: Amplification curve of *P. vivax* with parasitemia of 0.003 to 100parasites/µl. The graph was generated using LightCycler® 480 software.

Table 2.11: Sensitivity of real-time PCR based on parasitemia

Parasites/µl	CP value
100	21.46
50	22.89
25	23.91
12.5	24.91
6.25	26.05
3.125	26.90
1.56	27.89
0.78	28.84
0.39	30.00
0.195	30.82
0.0975	31.97
0.04875	33.68
0.024	34.61
0.012	33.82
0.006	35.33
0.003	36.05

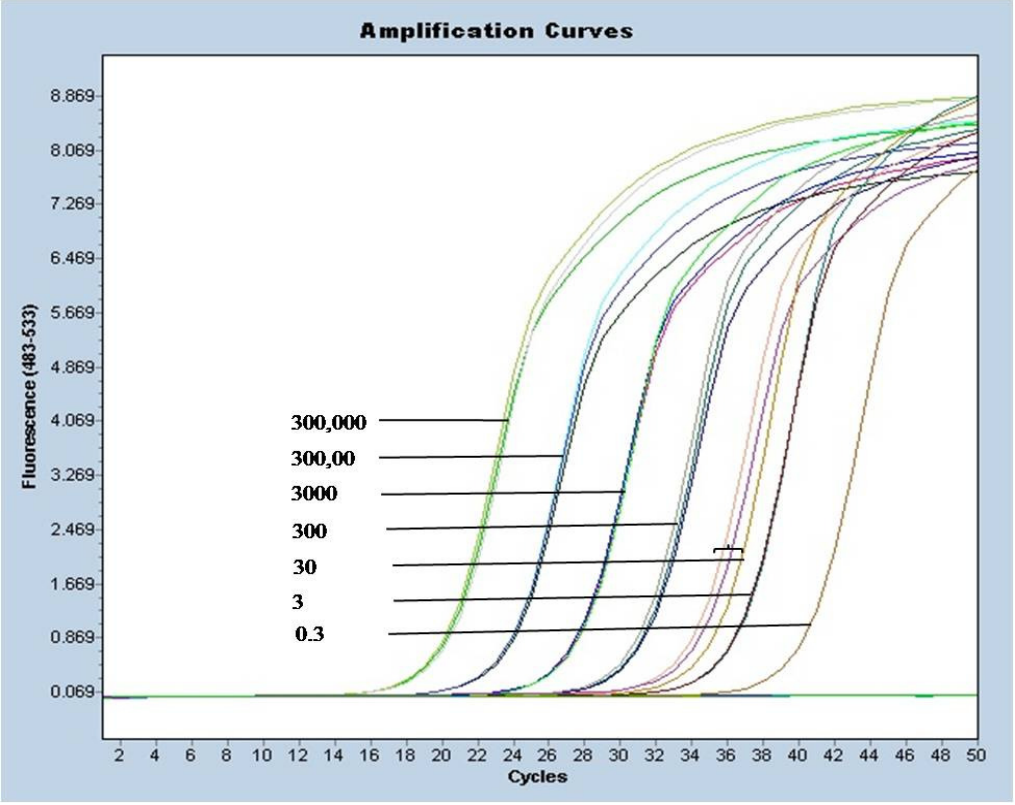

Figure 2.5: Amplification curve using plasmid controls of 0.003 to 300,000 copies. Genome concentration at 0.03 and 0.003 copies/µl were too low to be detected hence no amplification curve was observed. The graph was generated using LightCycler® 480 software.

Table 2.12: Sensitivity of real-time PCR based on copy numbers

DNA copy numbers/ ul	CP
300,000	19.75
300,000	19.71
300,000	19.97
30,000	23.50
30,000	23.47
30,000	23.45
3,000	27.20
3,000	27.12
3,000	27.42
300	30.94
300	30.96
300	30.74
30	33.85
30	34.17
30	35.07
3	36.44
3	36.18
3	36.16
0.3	-
0.3	-
0.3	40.09
0.03	-
0.03	-
0.03	-
0.003	-
0.003	-
0.003	-

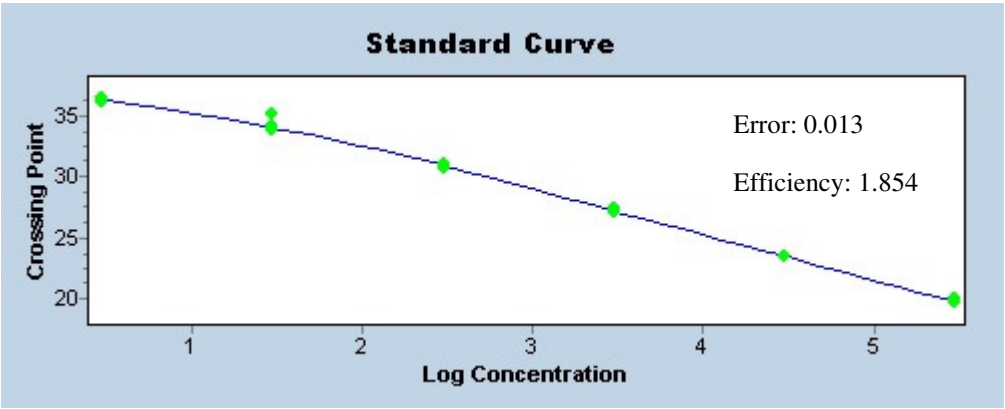

Figure 2.6: Standard curve generated from the amplification profile of the SYBR green-based quantitative PCR of known genome copy numbers (3 to 300,000 copies/ µl) using the PlasF and PlasR primers. This standard curve was generated using LightCycler® 480 software.

2.3.4 Specificity of primers in detecting *Plasmodium* parasites

To determine the *Plasmodium* genus specificity of primers PlasF and PlasR, real-time PCR assays were run against the four human and five simian malaria parasites' and malaria-negative human and macaques' DNA. Melting peak analysis revealed the clear detection of all the *Plasmodium* species controls, with an average melting temperature (T_m) of 80.32°C (Figure 2.7), though *P. cynomologi* showed a slightly higher T_m of 81.10ºC. None of the amplification reaction with malaria-negative human and macaque DNA produced a product with a melting temperature of 80ºC (Table 2.13).

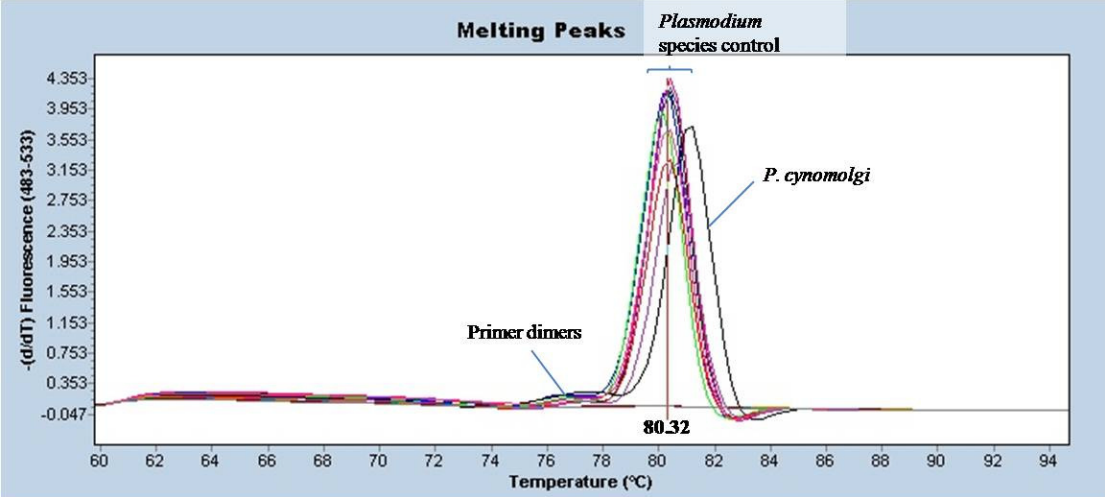

Figure 2.7: Melting curve analysis with nine *Plasmodium* species controls and four malaria-negative human and macaques samples. Strong T_m peaks were seen using *Plasmodium* controls but not malaria-negative samples. The graph was generated using LightCycler® 480 software.

Table 2.13: T_m values of PCR producted generated with each *Plasmodium* species.

Samples	T_m/^{0}**C**
P. falciparum	80.38
P. malariae	80.42
P. ovale	80.40
P. vivax	80.23
P. knowlesi	80.09
P. coatneyi	80.32
P. cynomolgi	81.10
P. fieldi	80.32
P. inui	80.60
Negative human blood 1	-
Negative human blood 2	-
Negative monkey blood 1	-
Negative monkey blood2	-

2.3.5 Development of simian malaria species-specific nested PCR assay

2.3.5.1 Optimization of annealing temperature for nest one *Plasmodium* genus-specific primers

Primer set PkCSP-F and PkCSPR2 was able to detect the positive control (*P. knowlesi* DNA) at all the annealing temperatures tested (Figure 2.8). Although non-specific amplification of human and macaque DNA was observed at lower annealing temperature of 51^{o}C and 53^{o}C, these bands disappeared at higher temperature. Hence, 55^{o}C was used as the annealing temperature for the nest one amplification of the full *csp* gene. When tested against the four human and five simian malaria parasites' DNA, these primers were found to be specific against *P. knowlesi, P.coatneyi, P. cynomolgi, P. fieldi, P. inui* only (Figure 2.9).

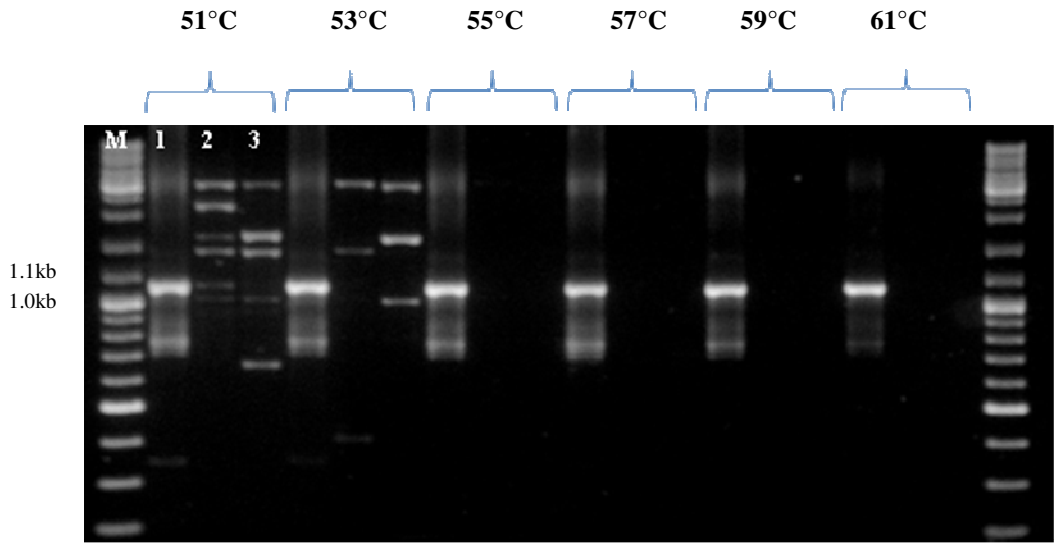

Figure 2.8: Determination of optimum annealing temperature for nest one PCR assay. Lane numbers (1 to 3) represent *P. knowlesi* DNA, malaria-negative human and macaque DNA, respectively. Annealing temperatures tested were 51°C, 53°C, 55°C, 57°C, 59°C and 61°C. Molecular size markers (1000-basepair ladder) are marked in lane M.

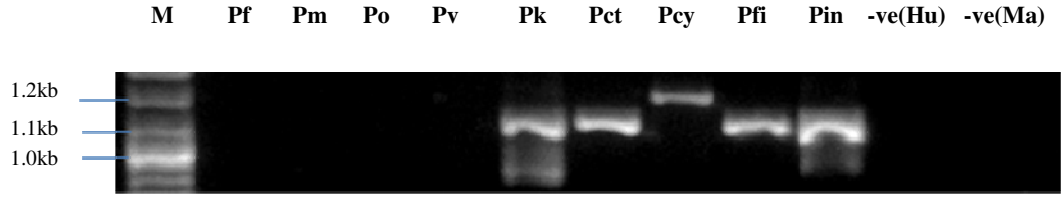

Figure 2.9: Nest one PCR on the four human and five simian malaria parasite controls. The DNA templates used were as follows: *P. falciparum* (Pf), *P. malariae* (Pm), *P. ovale* (Po), *P. vivax* (Pv), *P. knowlesi* (Pk), *P. coatneyi* (Pct), *P. cynomolgi* (Pcy), *P. fieldi* (Pfi), *P. inui* (Pin), and malaria-negative human (-veHu) and macaque (-veMa). Molecular size markers (1000-basepair ladder) are marked in lane M.

2.3.5.2 Determination of optimum annealing temperature and specificity of nest two species-specific primers

Each set of species-specific primers were tested against the nest one products of the four human and five simian *Plasmodium* species parasites, using annealing temperature listed in Table 2.8. The optimized annealing temperature for CspCOAT-F1 and CspCOAT-R1, CspCYNO-F2.1 and CspCYNOR2, CspFIELDI-F2a and PKCSPR2, CspINUI-F2 and CspINUI-R1, and CspKnowlesiF and PKCSPR2 are 62 ºC, 66ºC, 61ºC, 64ºC and 63ºC, respectively (Table 2.14). Using the optimized annealing temperature, the respective primer sets were highly specific towards the *Plasmodium* species they were designed for (see Figure 2.10). No amplification was observed for all four human *Plasmodium* species.

Table 2.14: Specificity of each primer pair in detecting the five simian *Plasmodium* parasites' DNA at various annealing temperatures. The optimum temperature selected for each species-specific primer set is labelled with *.

Primer Pair	***Plasmodium* species tested**	**Annealing temperature(^{o}C) Presence of amplified products**				
		58	60	62*	64	66
	P. coatneyi	+	+	+	+/-	-
	P. cynomolgi	-	-	-	-	-
CspCOAT-F1 CspCOAT-R1	*P. fieldi*	-	-	-	-	-
	P. knowlesi	+	+/-	-	-	-
	P.inui	-	-	-	-	-
		58	60	62	64	66*
	P. coatneyi	-	-	-	-	-
	P. cynomolgi	+	+	+	+	+
CspCYNO-F2.1 CspCYNOR2	*P. fieldi*	+	+	+	+/-	-
	P. knowlesi	-	-	-	-	-
	P.inui	-	-	-	-	-
			59	61*	63	65
	P. coatneyi		-	-	-	-
	P. cynomolgi		-	-	-	-
CspFIELDI-F2a PKCSPR2	*P. fieldi*		+	+	+/-	+/-
	P. knowlesi		-	-	-	-
	P.inui		-	-	-	-
		58	60	62	64*	66
	P. coatneyi	+	+	+/-	-	-
	P. cynomolgi	-	-	-	-	-
CspINUI-F2 CspINUI-R1	*P. fieldi*	-	-	-	-	-
	P. knowlesi	-	-	-	-	-
	P.inui	+	+	+	+	+/-
			59	61	63*	65
	P. coatneyi		+	+/-	-	-
	P. cynomolgi		-	-	-	-
CspKnowlesiF PKCSPR2	*P. fieldi*		+/-	-	-	-
	P. knowlesi		+	+	+	+/-
	P.inui		-	-	-	-

"+" indicates strong band
"+/-" indicates faint band
"-" indicates no amplification product

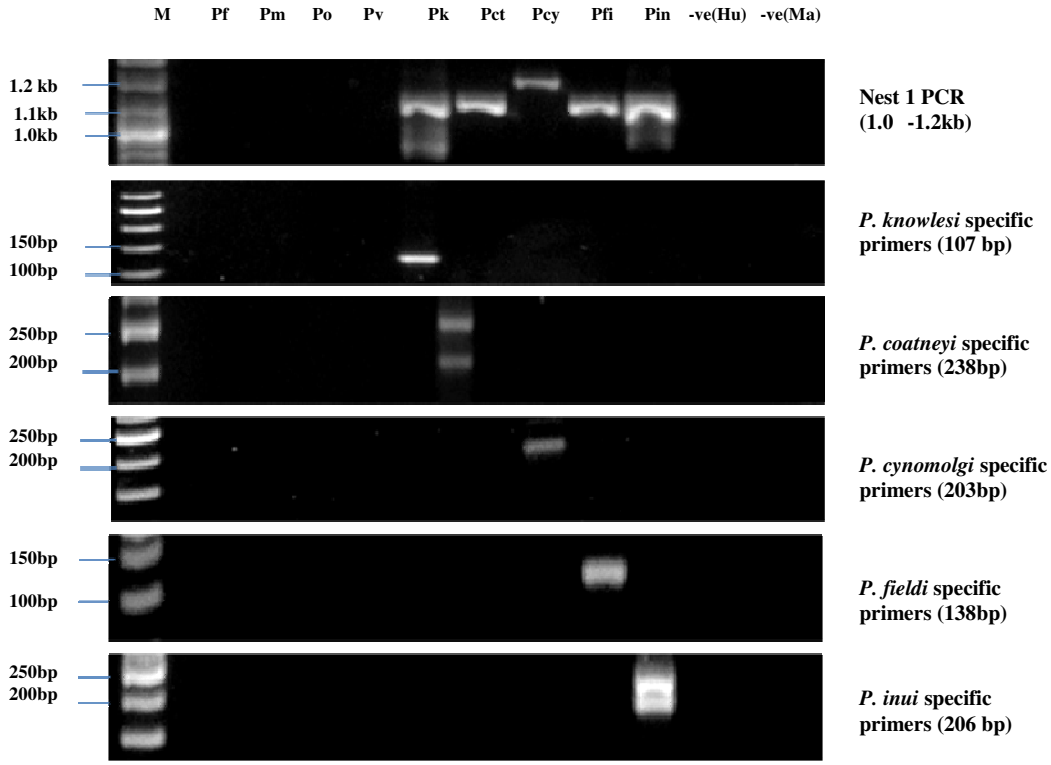

Figure 2.10: Specificity of each primer set in detecting the five simian *Plasmodium* species. The DNA templates used were *P. falciparum* (Pf), *P. malariae* (Pm), *P. ovale* (Po), *P. vivax* (Pv), *P. knowlesi* (Pk), *P. coatneyi* (Pct), *P. cynomolgi* (Pcy), *P. fieldi* (Pfi), *P. inui* (Pin), and malaria-negative human (-veHu) and macaque (-veMa). Molecular size markers (50-basepair ladder) are marked in lane M.

2.4 Discussion

The newly designed PlasF and PlasR are efficient and versatile primers that can be used for both conventional and real-time PCR methods. SYBR green-based and probe-based assays are the two commonly used real-time PCR methods to detect malaria parasites in blood samples. Between these methods, SYBR green assay is less expensive and precludes the use of complex light-sensitive probes. Furthermore, probe-based method is sensitive to nucleotide base mismatch which may result in false negative results. This is probable if an assay is designed to encompass multiple species within a genus, such as malaria. With these in mind, the SYBR green-based assay was chosen instead.

Primers PlasF and PlasR were designed based on the conserved region of both asexual and sexual stages of the *Plasmodium* SSU rRNA. When used in both PCR formats, the assay can detect at least 0.003parasites/µl, which is slightly higher than that of the published nested PCR assay. Furthermore, the ability of the SYBR green quantitative real-time PCR assay to detect less than three gene copies/µl also translates to a detection limit of less than one parasite/µl, as SSU rRNA is a multi-copy gene in *Plasmodium* parasites. *Plasmodium* species possess around four to eight SSU rRNA gene copies, with different copies expressed at different developmental stage of the parasite [99]. Hence, the sensitivity of the primers is most likely due to the multi-copy nature of the target gene and its short fragment length of 180bp [100].

The quantitative feature of real-time PCR is an additional advantage over the conventional PCR assays and microscopy. Counting blood-stage parasites can be

time-consuming and the accuracy of quantification is affected by factors, such as number of fields observed under the microscopy and the experience of the microscopist. Hence, due to these limitations, the standard curve for quantitative PCR was generated using plasmid standards rather than the counted parasitemia. In addition, as different *Plasmodium* species have different copy number of SSU rRNA genes, a standard curve generated using known parasitemia of one *Plasmodium* species will not be useful for quantification of another. Therefore, real-time PCR based on known copy numbers of SSU rRNA provides an unbiased and objective method in determining parasite load.

Aside from its high sensitivity, the primers were able to detect all human and five simian malaria parasites with no cross-reactivity with human or long-tailed macaque DNA. From the results obtained, the T_m averaged at 80.32 ºC with the exception of *P. cynomolgi* at 81.10ºC (Table 2.13). The difference in the T_m values between *P. cynomolgi* and the rest of the *Plasmodium* species is likely a result of sequence polymorphism in the region between the two primers' binding sides, resulting in a difference in the GC content. Nonetheless, primers PlasF and PlasR are found to be highly specific for *Plasmodium* parasites, since no amplification reaction with malaria-negative human and macaque DNA produced a product with a melting temperature of 80ºC and above (Figure 2.7).

When the current primers were used in both methods, complete amplification reaction was achieved in less than 1.5 hours, whereas, the published *Plasmodium* genus-specific nested PCR assays can only be completed in around four hours. Moreover, as

this assay requires only a single round of amplification, the problem of cross contamination of PCR products (commonly associated with nested PCR assay), was significantly reduced. The ability of the primers, PlasF and PlasR, to detect all human malaria parasites coupled with its high sensitivity and short amplification time can be useful in the *Plasmodium* screening of large sample numbers. In comparison to the *Plasmodium* genus-specific nested PCR assay, disease surveillance using a single run PCR method will cost less especially when conducted in a hypoendemic area where malaria prevalence is low. The ability of the primers to be used in both conventional and real time PCR methods will also allow areas with different resources to have comparable malaria screening results.

In contrast to the *Plasmodium* parasites screening PCR assays, the simian malaria species-specific nested PCR assay was designed based on the *Plasmodium csp* gene instead of the SSU rRNA gene. The SSU rRNA gene is a multi-copy gene; each *Plasmodium* parasite species possesses several structurally distinct sets of SSU rRNA gene [99]. Hence, the design of five sets of simian *Plasmodium* species-specific primers using this gene may be challenging due to potential shared primer binding site between different species of malaria parasites. An example is the cross reaction of the published *P. knowlesi* specific primers (Pmk8 and Pmkr9) with the SSU rRNA gene segment of *P. vivax* [90], resulting in inaccurate identification of these parasites in human patients. Aside from being a single copy gene, the *Plasmodium csp* gene is a good target for the design of simian malaria species specific primers as *csp* sequences of several malaria parasites are readily available in the GenBank.

Overall, the newly developed nested PCR assay was found to be specific in distinguishing *P. knowlesi*, *P. coatneyi*, *P. cynomolgi*, *P. fieldi* and *P. inui* in macaque samples. Furthermore, primers CspKnowlesiF and PKCSPR2 were able to differentiate *P. knowlesi* from *P. vivax* in human infections. However, the sensitivity of this nested PCR assay could not be determined since the parasitemia of the simian *Plasmodium* DNA controls were not known.

CHAPTER THREE

Prevalence of simian malaria parasites in Singapore's macaques

3.1 Introduction

A surveillance study of malaria parasites in Malayan monkeys revealed that long-tailed macaques (*Macaca fascicularis*) have the greatest prevalence of malaria as compared to other macaque species [22]. With an estimated population of 1400, long-tailed macaque is the predominant macaque species in Singapore [101, 102]. They can be found in forest and near human habitations. In a previous study, this species of macaque was incriminated as the natural host of *P. knowlesi* infection for the human cases in Singapore [20]. As *P. knowlesi* infection in humans can be severe and potentially fatal [64, 65, 69], and other simian malaria parasites such as *P. cynomolgi* and *P. inui*, have been shown to be potentially infectious to humans, there is a need to determine the prevalence of simian malaria parasites in local macaques so that the risk of potential malaria zoonosis can be established

3.2 Materials and methods

3.2.1 Macaques' blood samples

Macaques in this surveillance study can be categorized into two groups – the wild (denoted with code WM) and the peri-domestic (denoded with code PM). Wild macaques were caught in the forests used for military training, an area which the general public had no access to. Hence, wild macaques had limited interaction with the general human population in Singapore. These macaques were caught under an operational surveillance program approved by the Singapore military's joint medical committee and the DSO National Laboratory's Institutional Animal Care and Use Committee. Two macaques were sampled from the mainland forest in November 2007 upon notification of the first human knowlesi cases. Another 91 macaques were sampled from both the mainland (n=84) and offshore military forest (n=7) from April 2009 to May 2011. On the other hand, peri-domestic macaques are found near human habitations and have closer interactions with the general public. Ten macaques were sampled from the central nature reserve park in January 2008. Another 55 were sampled from various parts of Singapore as part of a routine population control effort by the Agri-Food and Veterinary Authority of Singapore (AVA).

All macaques caught were sent to AVA for age, sex and species characterization (see Appendix C and D for details of peri-domestic and wild macaques respectively). The age of macaques was estimated by dentition analysis [103]. Macaques age three years and below were classified as juveniles while those estimated to be age three years and above were classified as adults (Elvira Menguita, personal communication,1st Nov 2011). Blood was collected in accordance with the ethical practices of AVA, and

EDTA blood samples were sent to EHI for analysis. Figure 3.1 illustrates the locations where macaques were sampled.

3.2.2 DNA extraction and screening of macaques' blood samples for simian malaria parasites

DNA was extracted from the EDTA whole blood using protocol described in Section 2.2.2.3. A rapid screening of malaria parasites was conducted using the real-time PCR described in Chapter Two. For samples positive for malaria parasites, species-specific PCR assay were subsequently used to identify the species of malaria parasites. PCR products were analyzed by electrophoresis in 2% agarose gel, stained with Gel-Red^{TM} (Biotium, Hayward, CA, USA), and observed under ultraviolet transillumination.

Figure 3.1: Geographical representation of locations where macaques in this study were sampled. Area highlighted in green is the restricted access forest where wild long-tailed macaques were caught. Seven wild macaques were also sampled in Pulau Tekong (represented by a green point). Peridomestic macaques, on the other hand, were indicated with blue points. The size of the blue points was relative to the number of macaques sampled. Area highlighted in blue is the Central Nature Reserve, where ten peridomesticated macaques were sampled in 2008.
Map was plotted using www.map.gov.sg.

3.3 Results

3.3.1 Screening of macaques for *Plasmodium* parasites

A total of 157 out of the 158 long-tailed macaques (*Macaca fascicularis*) sampled were screened for malaria parasites using real-time PCR assay. Of these, 92 and 65 were wild and peri-domestic macaques, respectively. Of the 92 wild macaques, 71.7% (*n*=66) were infected with malaria parasites. Among these infected ones, 36.4% (*n*=24) were juveniles. This corresponds to a high infection rate of 80% (out of total 30) among the juveniles macaques. Comparatively, the infection rate among the adult macaques is only 67.7%. None of the peri-domesticated macaques were found to be infected (Table 3.1).

All the malaria positive samples were subsequently screened with the nested PCR to determine the species of *Plasmodium* parasites present. All five simian malaria parasites were detected, with *P. knowlesi* being the most prevalent (68.2%), followed by *P. cynomolgi* (60.6%), *P. fieldi* (16.7%), *P. coatneyi* (3.0%) and *P. inui* (1.5%). Furthermore, 62.5% (*n*=15) of the infected juvenile macaques harboured *P. knowlesi*. *Plasmodium inui* was only detected from the single malaria-positive macaque trapped in the military off-shore island.

In addition to the high malaria prevalence among the wild macaques, co-infection with multiple species of *Plasmodium* parasites was also observed. Dual infection was detected in 23 (34.8%) macaques, of which five were juveniles. Five (7.6%) macaques were infected with three *Plasmodium* species, of which two were juveniles Out of these 28 macaques with multiple infections, 25 (89.3%) had *P. knowlesi*

infection in them. Table 3.2 summarizes the malaria infections in the infected macaques. The screening results of individual macaques are listed in Appendix D.

Table 3.1: Summary of malaria infections in macaques sampled in this study

Macaque population		Infection rate of macaques screened				
		N	Infected		Not infected	
			n	Percentage/%	n	Percentage/%
Wild	Adult	62	42	67.7	20	32.3
	Juvenile	30	24	80	6	20
	Sub total	92	66	71.7	26	28.3
Peri-domestic	Adult	50	0		50	100
	Juvenile	15			15	
	Sub total	65			65	
Total *Plasmodium* positive: 66						
Total *Plasmodium* negative: 91 (26 wild, 65 peri-domestic)						
Total screened: 157						

Table 3.2: Breakdown of malaria infections in infected macaques

		Macaque type		
Infection	***Plasmodium* species**	**Adult**	**Juveniles**	**Total**
Single	Pk	11	9	20
	Pcy	10	8	18
Double	Pk, Pcy	11	4	15
	Pk, Pfi	3	0	3
	Pk, Pct	1	0	1
	Pk, Pin	1	0	1
	Pcy, Pfi	2	1	3
Triple	Pk, Pcy, Pfi	3	1	4
	Pk, Pfi, Pct	0	1	1
Total *Plasmodium*-positive				**66**

Pct, Pcy, Pfi, Pin and Pk denodes *Plasmodium coatneyi*, *P. cynomolgi*, *P. fieldi*, *P. inui* and *P. knowlesi*, respectively

3.4 Discussion

The screening of 157 macaques sampled from different localities in Singapore revealed that all malaria-infected macaques were found in the restricted forest, an area with very limited access to the general public. On the other hand, none of the peri-domestic macaques was found to be infected with malaria parasites. Similar observations were reported in Thailand and Malaysia [24, 28]. In both countries, all macaques caught in the urban areas were negative for malaria infection while those caught from the forest had a high infection rate. The reason for this observed difference was hypothesized to be the lack of competent vectors for simian malaria transmission in the urban areas.

Overall infection rate was high, with 71.7% of the population sampled infected. Of these 57.6%, 34.8% and 7.6% were due to single, dual or triple infections, respectively. Among the infected wild macaques, *Plasmodium knowlesi* has the highest prevalence rate (68.2%). Other than *P. knowlesi*, these wild macaques also harboured *P. coatneyi, P. cynomolgi, P. fieldi* and *P. inui*. In addition, 80% of the juvenile macaques were found to be infected with malaria parasites. The high infection rate, especially among the juveniles, suggests a high intensity of transmission occurring in the forest. Although transplacental transmission in simian malaria parasites has not been reported, it cannot be ruled out that the high infection rates among juvenile macaques can be due to this. An understanding of the genotypic diversity of *Plasmodium* parasites found in these macaques might be able to shed further insights regarding these hypotheses.

Plasmodium cynomolgi is also a common malaria parasite found in local macaque population, with a prevalence rate of 60.6%. Together with *P. inui*, these malaria parasites were shown to be infectious to humans under laboratory conditions [35-39]. To date, there has been no report of naturally-acquired human infections with these parasites. One possible reason could be misdiagnosis due to their overlapping morphological characteristics with the human malaria parasites. Therefore, the simian malaria species-specific PCR assay will be useful in distinguishing these zoonotic malaria infections.

Although none of the peri-domesticated macaques were infected with malaria, we can only conclude that the risk of a zoonotic transmission is low but not totally diminished. It is possible for the wild macaques in the restricted-access forest to migrate out to areas with human habitations, although they generally inhabited in the restricted-access forest, an area restricted to the general public. However, the actual risk of simian malaria transmission to humans could not be determined since the vector involved in the transmission in Singapore is currently not known. In addition, the sampling of macaques in this study was carried out under the operations of the military and the national veterinary authority. Hence, no random selection of macaques sampling sites was performed and the areas where our macaques were caught from might not constitute the entire range of Singapore's macaques' distribution. Hence, it is possible that there may be other areas (not covered in this study), with simian malaria transmission occurring. From the current study, *P. inui* was only detected in one of the wild macaques sampled from an offshore military island. Although *P. inui* was not detected among the macaques from mainland

Singapore, we cannot conclude that this species of malaria parasite is absent as the sample of wild macaques obtained may not be representative of the total population.

This study constitutes the first report of surveillance of simian malaria parasites in long-tailed macaques in Singapore. Although previous study had illustrated that wild long-tailed macaques were the reservoir hosts of human *P. knowlesi* infections [20], the current results showed that local long-tailed macaques also harbour *P. coatneyi, P. cynomolgi, P. fieldi* and *P. inui*.

CHAPTER FOUR

Characterization of the *circumsporozoite* protein genes of *Plasmodium* species from Singapore's macaques

4.1 Introduction

Sixty-six wild long-tailed macaques were screened positive for malaria parasites using real-time PCR (Chapter 3). With the use of nested PCR assay developed in this study, *P. knowlesi, P. cynomolgi, P. fieldi P. coatneyi* and *P. inui* were detected. To confirm the nested PCR results, at least one gene of the parasite is to be characterized. The *circumsporozoite* protein *(csp)* gene is a gene which has been proven useful for the phylogenetic inferences of *Plasmodium* parasites. It produces a phylogenetic tree with topology similar to one constructed using the SSU rRNA gene sequence [96]. As it is an attractive candidate for malaria vaccine development due to its high immunogenic nature, there are more nucleotide sequence information on this gene in the GenBank as compared with other antigen-coding *Plasmodium* gene [104]. It is thus particularly useful for phylogenetic inferences.

The *csp* gene is a single copy gene which encodes for an antigenic protein that covers the entire surface of the sporozoites [96, 105]. It is composed of a variable central region of repeats, flanked by two conserved motifs known as region I and region II-plus, located at the amino- and carboxyl-terminal ends of the gene, respectively [96, 106, 107]. Region I is based on the short amino acid motif KLKQP and this motif can be found in almost all mammalian *Plasmodium* parasites described thus far. On the other hand, region II-plus is a 20-amino acid motif EWSXCXVTCGXG(V/I)XXRX(K/R), which is homologous to the type 1 repeat of

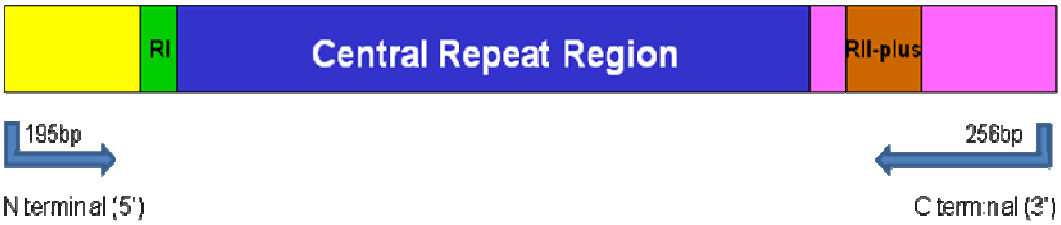

Figure 4.1: A schematic diagram illustrating the anatomy of the *Plasmodium csp* gene. The variable central repeat region is flanked by two conserved motifs (region I and region II-plus) at the amino- and carboxyl-terminal ends, respectively. Region I and Region II-plus is denoted by RI and RII-plus. For phylogenetic analysis, only the first 195 from the non-repeat N-terminal and the last 256 nucleotides from the C-terminal were used.

human thromospondin [106]. Both regions have been described to be important for the sporozoites' invasion into the mammalian host's hepatocyte cells [96, 108]. Subsequent studies had also revealed the potential role of region II-plus in sporozoite motility and invasion into mosquito salivary glands [106, 109].

Conversely, the variable central repeat region is made up of short amino acid residues which are tandemly repeated, making the circumsporozoite protein highly immunogenic. The length of a repeat unit and the number of repeats differ across and within species, making the size of a full *csp* gene highly variable [105]. As such, only the non-repeat regions at the N- and C-terminal were used for sequence alignment and phylogenetic analysis (Figure 4.1). Although the exact function of the repeat region is unknown, it has been hypothesized that the peptide repeats may act as a "smoke screen" to evade host immune system during the sporozoites' invasion [105].

In this study, the *csp* gene was used to determine the phylogenetic relationship of malaria parasites detected from the long-tailed macaques. This gene has been used in phylogeny studies to understand the evolutionary history and relatedness of different primate malaria parasites [25, 104, 110]. Since *csp* is a single copy gene, each *csp* sequence is representative of a single *Plasmodium* parasite isolate. Hence, characterization of the *csp* gene could provide an indication on the genetic diversity of the *Plasmodium* parasites in our macaque samples. Moreover, this gene has also been used in the molecular epidemiological investigation of human knowlesi cases [20, 24, 25]. Characterization of *Plasmodium csp* gene from humans and macaques will aid also in the understanding on the transmission dynamics of simian malaria parasites in Singapore.

4.2 Materials and methods

4.2.1 Isolates used for *csp* gene characterization

The *csp* genes of malaria parasites from 15 wild long-tailed macaques, denoted as SG/EHI/WM01/Y07, SG/EHI/WM02/Y07, SG/EHI/WM04/Y09, SG/EHI/WM05/Y09, SG/EHI/WM11/Y09, SG/EHI/WM15/Y09, SG/EHI/WM16/Y09, SG/EHI/WM17/Y09, SG/EHI/WM18/Y09, SG/EHI/WM26/Y09, SG/EHI/WM33/Y09, SG/EHI/WM35/Y09, SG/EHI/WM42/Y10, SG/EHI/WM44/Y10 and SG/EHI/WM91/Y11, and together with those obtained from the LRDU of the Malaria branch, Division of Parasitic Diseases and Malaria, CDC, USA were characterized. The *P. knowlesi csp* gene sequences obtained from the human cases isolated in 2007 (SG/EHI/H1/Y07, SG/EHI/H2/Y07, SG/EHI/H7/Y07) and 2008 (SG/EHI/H24/Y08), were analyzed. In addition, two local knowlesi cases imported from peninsular Malaysia reported in 2009 were also included in the analysis. The full *csp* gene sequences of these isolates, together with their GenBank accession numbers were listed in Appendix E.

4.2.2 Cloning of the *Plasmodium csp* genes

Amplification of the *Plasmodium csp* gene was carried out using the protocol described in Section 2.3.5. For cloning of blunt-end PCR products, the Zero Blunt® PCR Cloning Kit (Invitrogen, Carlsbad CA, USA) was used, with reference to the methods described in Section 2.2.3.6.2. At least 80 *E. coli* transformants were screened by colony PCR using PKCSP-F and PKCSPR2 primers, as described in Section 2.2.4.2.1.

For macaque samples co-infected with two or more species of malaria parasites, a second round PCR was performed on the products of the colony PCR to screen for the relevant simian malaria species in each monkey. The amplification parameters and reaction mix followed that described in Section 2.2.4.1.4, using the appropriate set of simian malaria species-specific primers and its optimized annealing temperature. Upon screening, at least ten transformants of each *Plasmodium* species from each human or macaque samples were selected for sequencing.

4.2.3 Preparation of glycerol stocks and plasmid DNA extraction

Glycerol stocks of transformants containing the correct *csp* inserts were prepared using methods described in Section 2.2.3.6.3.

Extraction of plasmid DNA was carried out using PureLink™ Quick Plasmid Miniprep (Invitrogen, Carlsbad CA, USA) according to the manufacturer's protocol (Section 2.2.3.6.4).

4.2.4 Sequencing of the *csp* gene

Sequencing of the *csp* genes was conducted, by a commercial company (1st BASE Pte Ltd., Singapore), according to the BigDye Terminator Cycle Sequencing kit (Applied Biosystems, USA) protocol. The primers used for the sequencing of the complete *csp* genes are listed in Table 4.1.

Table 4.1: Oligonucleotide sequences of primers used in the sequencing of the *csp* gene

Primers	**Direction**	**Oligonucleotide sequences**
M13F (-20)	Forward	5' GTAAAACGACGGCCAGT 3'
M13R (-24)	Reverse	5' GGAAACAGCTATGACCATG 3'
CSP Internal Repeat F	Forward	5’ CGAGGCAGAGGACTTGGTGA 3’
CSP Internal Repeat R	Reverse	5’ CCACAGGTTACACTGCAT 3’

4.2.5 DNA sequence analysis

The consensus sequence of *csp* gene from each *Plasmodium* species was obtained by assembling a contiguous sequence from the raw sequencing data using Seqman program (Lasergene, DNASTAR, USA). The untranslated regions of the *csp* gene were removed to obtain the full coding gene sequence. Only the sequence encoding the non-repeat N-terminal (first 195 nucleotides of coding sequence) and C-terminal (261 nucleotides of coding sequence) of the *csp* gene was aligned [45, 96] using MegAlign software (Lasergene, DNASTAR, USA). Transformants representative of each sequence polymorphism were selected for subsequent phylogenetic and full *csp* gene analysis. Complete *csp* sequences from the GenBank database (Table 4.2) were also retrieved to compare and analyse with the sequences obtained from this study.

To analyse the repeat region, the *csp* gene sequences were translated into amino acids using the EditSeq software (Lasergene, DNASTAR, USA). The entire repeat region of the *csp* gene was determined based on the maximum number of amino acid that formed a tandem repeat motif. To determine the polymorphisms within the repeat region of each *Plasmodium* species, unique amino acid motif sequence were assigned with an alphabet and arranged to reflect the actual amino acid sequence in the repeat region.

Table 4.2: List of GenBank *csp* sequences used in the phylogenetic analysis

***Plasmodium* species**	**Accession number**	**Geographical origin**	**Host**	**Reference number**
P. falciparum	M83164	Thailand	*Homo sapiens*	[111]
P. vivax	M34697	Thailand	*H. sapiens*	[112]
P. malariae	U09766	China	*H. sapiens*	[113]
P.knowlesi, H	K00822	Peninsular Malaysia	*H. sapiens*	[114]
P.knowlesi, Nuri	M11031	Peninsular Malaysia	*Macaca fascicularis*	[115]
P. knowlesi, MPRK13	EU687469	Peninsular Malaysia	*H. sapiens*	[24]
P. knowlesi, M197	EU821336	Peninsular Malaysia	*M. fascicularis*	[24]
P.knowlesi, KH 100	GU002488	Sarawak	*H. sapiens*	[25]
P.knowlesi, LT48-B11	GU002510	Sarawak	*M. fascicularis*	[25]
P. cynomolgi, Ceylon	M15103	Sri Lanka	*M. nemestrina*	[116]
P. cynomolgi, Berok	M15104	Peninsular Malaysia	*M. nemestrina*	[116]
P. coatneyi	GU002522	Sarawak	*M. fascicularis*	[25]
P. coatneyi, Hackeri	AY135360	Peninsular Malaysia	*M. fascicularis*	[110]
P. fieldi	GU002521	Sarawak	*M. fascicularis*	[25]
P. inui	GU002523	Sarawak	*M. fascicularis*	[25]
P. inui	FJ009512	Peninsular Malaysia	*M. fascicularis*	[24]
P. inui, strain Taiwan II	FN597613	Taiwan	*M. cyclopis*	[117]
P. inui, strain Taiwan I	FN597612	Taiwan	*M. cyclopis*	[117]
P. simium	L05069	Brazil	*Alouatta fuscus*	[118]
P. simiovale	U09765	Sri Lanka	*M. sinica*	[119]
P. berghei, ANKA	X17606	Zaire	*Grammomys surdaster*	[120]
P. yoelii	J02695	Central Africa	*Thamnomys rutilans*	[120]
P. gallinaceum	U65959	Sri Lanka	*Gallus gallus*	[96]

4.2.6 Phylogenetic analysis

The neighbour-joining (NJ) [121] and maximum-likelihood (ML) [122] method were used to analyze the *csp* gene sequences of all *Plasmodium* species obtained from long-tailed macaques and knowlesi patients. Phylogenetic analysis was carried out using MEGA version 5.0 ([123]; **http://www.megasotfware.net**). For both NJ and ML method, Kimura-parameter model was used in all analyses, including transition and transversion. Internal node reliability was measured by the bootstrap method after 1000 replicates [110].

4.3 Results

4.3.1 Cloning and sequencing of *Plasmodium* species *csp* genes

Plasmodium csp gene from 15 long-tailed macaques, four locally-acquired and two imported human knowlesi cases were amplified and cloned. The presence of *Plasmodium csp* genes further confirmed the presence of simian malaria parasites in these samples. The number of transformant screened by PCR for each sample and those chosen for complete sequencing are listed Table 4.3. At least 80 and 100 transformants for samples with single and mixed infection, respectively, were randomly screened by PCR using primers PKCSP-F and PKCSPR2. However, for SG/EHI/WM15/Y09 and SG/EHI/WM91/Y11, despite screening more than 500 transformants, none was found to contain inserts containing *P. coatneyi* and *P. knowlesi csp* genes respectively.

Table 4.3: Summary of number of *E.coli* transformants of each isolate analyzed by colony PCR, and the code of transformants selected for complete *csp* gene analysis and phylogenetic inferences

Isolate	**No. transformants analyzed by PCR**	***Plasmodium* species**	**Code of transformants for phylogenetic and full sequence analysis**
SG/EHI/H01/Y07	90	Pk	SG/EHI/H1/Y07-15
SG/EHI/H02/Y07	85	Pk	SG/EHI/H2/Y07-12
SG/EHI/H07/Y07	80	Pk	SG/EHI/H7/Y07-01
SG/EHI/H24/Y08	80	Pk	SG/EHI/H24/Y08-10
SG/EHI/H1-im/Y09	102	Pk	SG/EHI/H1-im/Y09 -17, 25, 80, 95, 102
SG/EHI/H2-im/Y09	105	Pk	SG/EHI/H2-im/Y09 -3,7, 11, 12, 18
SG/EHI/WM01/Y07	150	Pk Pfi	SG/EHI/WM01/Y07-6 SG/EHI/WM01/Y07-23
SG/EHI/WM02/Y07	120	Pk Pcy	SG/EHI/WM02/Y07-1, 39 SG/EHI/WM02/Y07-110
SG/EHI/WM04/Y09	80	Pk	SG/EHI/WM04/Y09-8, 9, 12, 13, 14, 15
SG/EHI/WM05/Y09	160	Pk Pcy Pfi	SG/EHI/WM05/Y09-70, 79 SG/EHI/WM05/Y09-65 SG/EHI/WM05/Y09-68
SG/EHI/WM11/Y09	80	Pk	SG/EHI/WM11/Y09-74
SG/EHI/WM15/Y09	450	Pk Pfi Pct	SG/EHI/WM15/Y09-149 SG/EHI/WM15/Y09-163 -
SG/EHI/WM16/Y09	100	Pk Pcy	SG/EHI/WM16/Y09-85 SG/EHI/WM16/Y09-2, 34
SG/EHI/WM17/Y09	80	Pk	SG/EHI/WM17/Y09-4, 30
SG/EHI/WM18/Y09	100	Pcy Pfi	SG/EHI/WM18/Y09-24 SG/EHI/WM18/Y09-92
SG/EHI/WM26/Y09	80	Pk	SG/EHI/WM26/Y09-1, 13, 47, 60, 98, 123
SG/EHI/WM33/Y09	87	Pk Pcy	SG/EHI/WM33/Y09-39 SG/EHI/WM33/Y09-47
SG/EHI/WM35/Y09	80	Pk	SG/EHI/WM35/Y09-38, 74
SG/EHI/WM42/Y10	80	Pcy	SG/EHI/WM42/Y10-1
SG/EHI/WM44/Y10	105	Pcy Pfi	SG/EHI/WM44/Y10-3, 30 SG/EHI/WM44/Y10-64
SG/EHI/WM91/Y11	320	Pin Pk	SG/EHI/WM91/Y11-61, 73 -

Pct, Pcy, Pfi, Pin and Pk denodes *Plasmodium coatneyi*, *P. cynomolgi*, *P. fieldi*, *P. inui* and *P. knowlesi*, respectively.

4.3.2 Phylogenetic analyses of the *csp* genes

The phylogenetic trees constructed using NJ and ML methods are shown in Figure 4.2 and 4.3, respectively. Both methods produced phylogenetic trees of similar topology, and demonstrated that simian malaria parasites isolated from Singapore samples can be clustered into four major clades, namely *P. knowlesi, P. cynomolgi, P. fieldi* and *P. inui*.

Circumsporozoite protein gene sequences from transformants derived from 11 monkeys and the six human knowlesi cases formed a cluster with five subclades within the *P. knowlesi* clade (Figure 4.4). The *csp* gene sequences derived from the four locally-acquired human knowlesi cases were found to be identical to those isolated from some of the long-tailed macaques caught in the restricted forest. On the other hand, the *csp* gene sequences derived from two human knowlesi cases, which were epidemiologically classified as "imported", were found to form a distinct subclade.

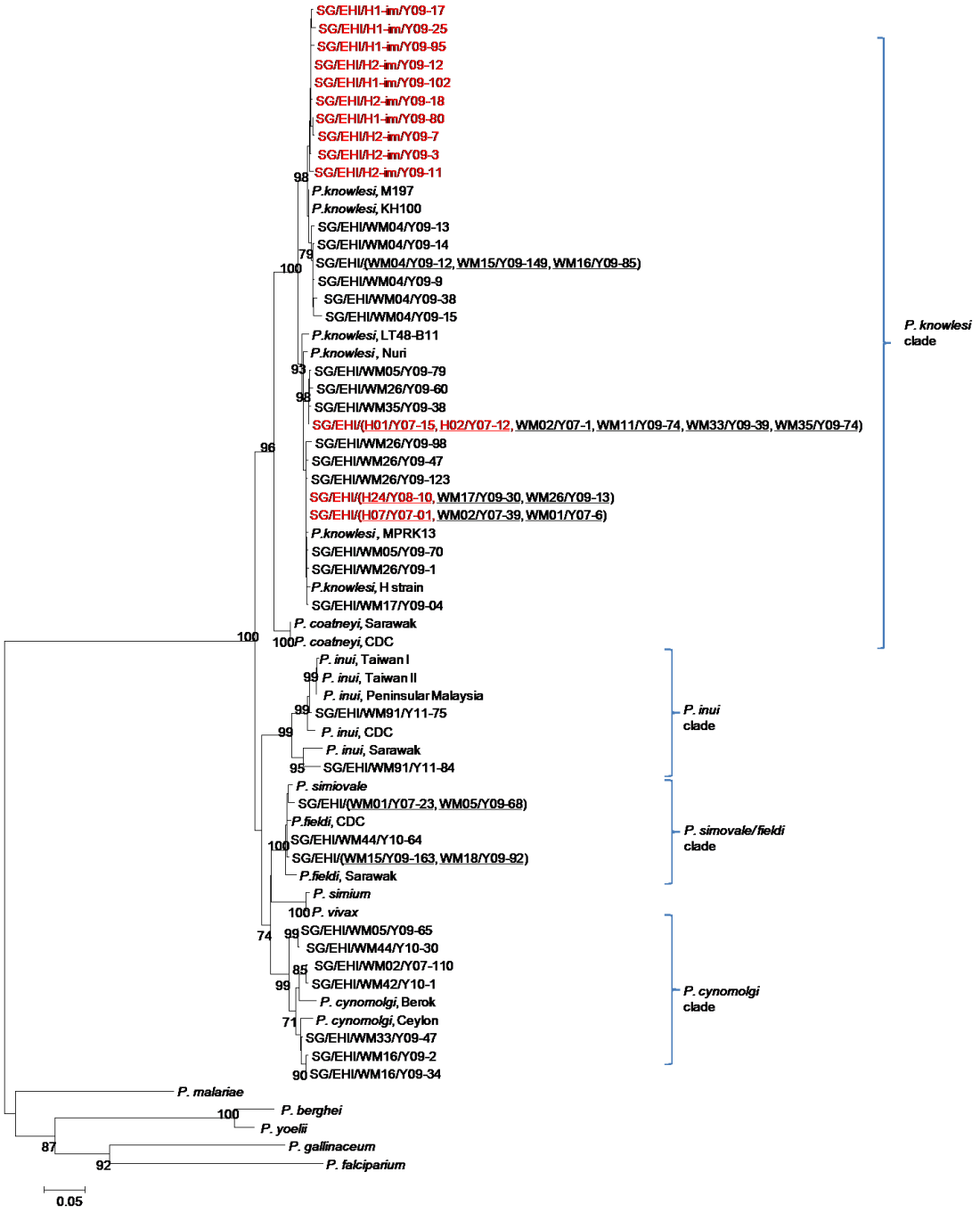

Figure 4.2: Phylogenetic tree of the non-repeat region of the *Plasmodium* species *csp* genes, constructed using the neighbour-joining method. Clones colored red are isolates from human samples. Clones underlined had shared genotype. Figures on the branches are bootstrap percentages based on 1000 replicates, and only bootstrap percentages above 70% are shown.

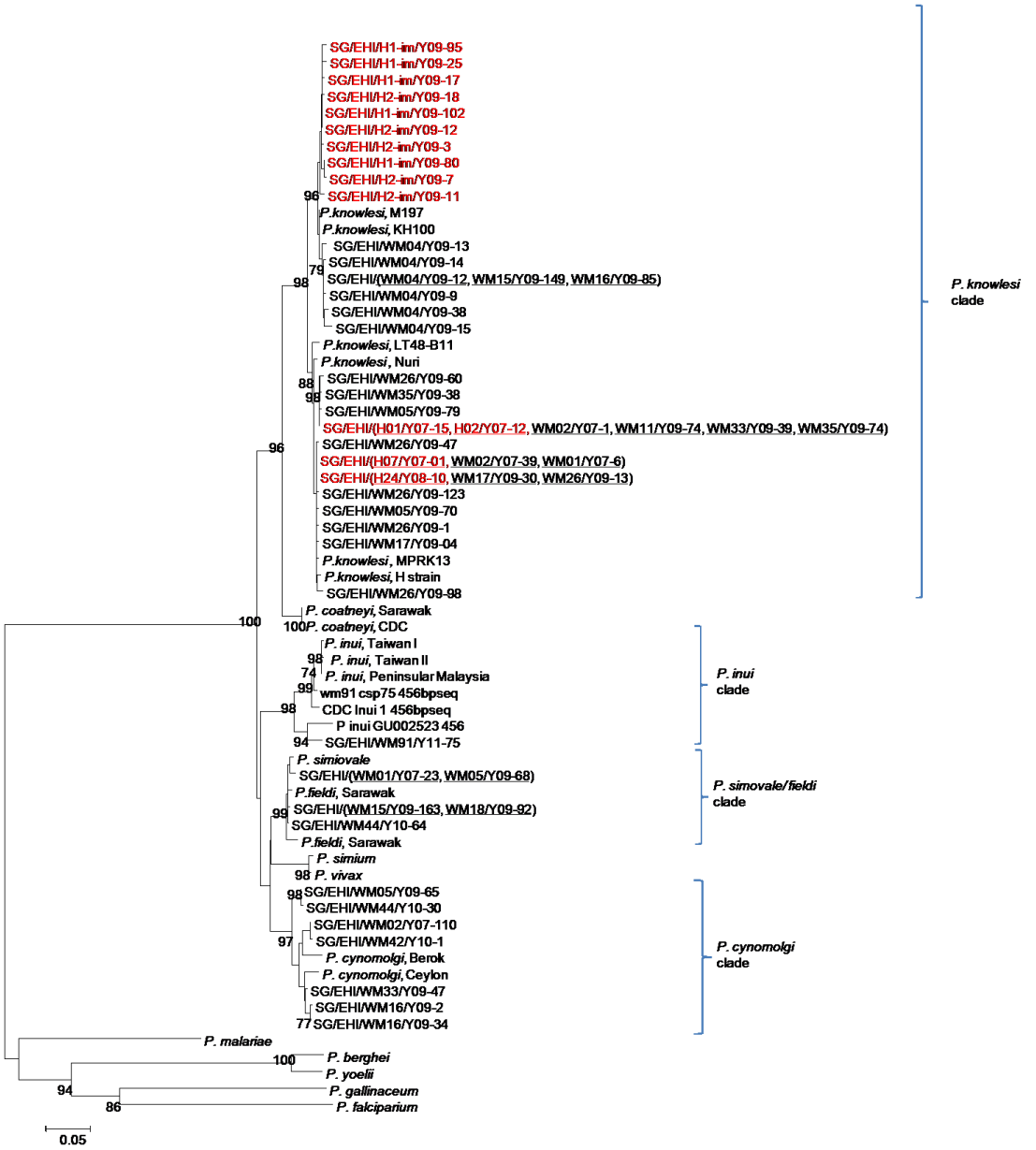

Figure 4.3: Phylogenetic tree of the non-repeat region of the *Plasmodium* species *csp* genes, constructed using the maximum-likelihood method. Clones colored red are isolates from human samples. Clones underlined had shared genotype. Figures on the branches are bootstrap percentages based on 1000 replicates, and only bootstrap percentages above 70% are shown.

The sequences of the non-repeat region of transformants SG/EHI/H7/Y07-1 were found to be identical to SG/EHI/WM2/Y07-39 and SG/EHI/WM1/Y07-6, while SG/EHI/H24/Y08-10 was found to be identical to SG/EHI/WM17/Y09-30 and SG/EHI/WM26/Y09-13. Similarly, the non-repeat region of transformants SG/EHI/H1/Y07-15, SG/EHI/H2/Y07-12, SG/EHI/WM02/Y07-1, SG/EHI/WM11/Y09-74, SG/EHI/WM33/Y09-39 and SG/EHI/WM35/Y09-74 were also found to be identical. SG/EHI/WM15/Y09-149, SG/EHI/WM4/Y09-12 and SG/EHI/WM16/Y09-85 were also found to have indistinguishable non-repeat regions. Interestingly, the non-repeat *P. knowlesi csp* sequence of human and macaque samples were identical although they were collected across years; human samples in 2007 (SG/EHI/H01/Y07, SG/EHI/H02/Y07 and SG/EHI/H07/Y07) and 2008 (SG/EHI/H24/Y08), while most of the macaques in this study were surveyed in 2009 (Appendix D).

Nine *csp* gene sequences from seven long-tailed macaques (SG/EHI/WM02/Y07, SG/EHI/WM05/Y09, SG/EHI/WM16/Y09, SG/EHI/WM18/Y09, SG/EHI/WM33/Y09, SG/EHI/WM42/Y10, and SG/EHI/WM44/Y10) were found to cluster in the *P. cynomolgi* clade. Within the *P. cynomologi* cluster, three distinct sub-clades with high bootstrap value were observed (Figure 4.2 and 4.3). Transformants derived from SG/EHI/WM05/Y09 and SG/EHI/WM44/Y10 were found to cluster in a subclade within *P. cynomolgi* clade, while transformants derived from SG/EHI/WM16/Y09 and SG/EHI/WM33/Y09, and SG/EHI/WM02/Y07, SG/EHI/WM18/Y09 and SG/EHI/WM42/10, formed two distinct subclades.

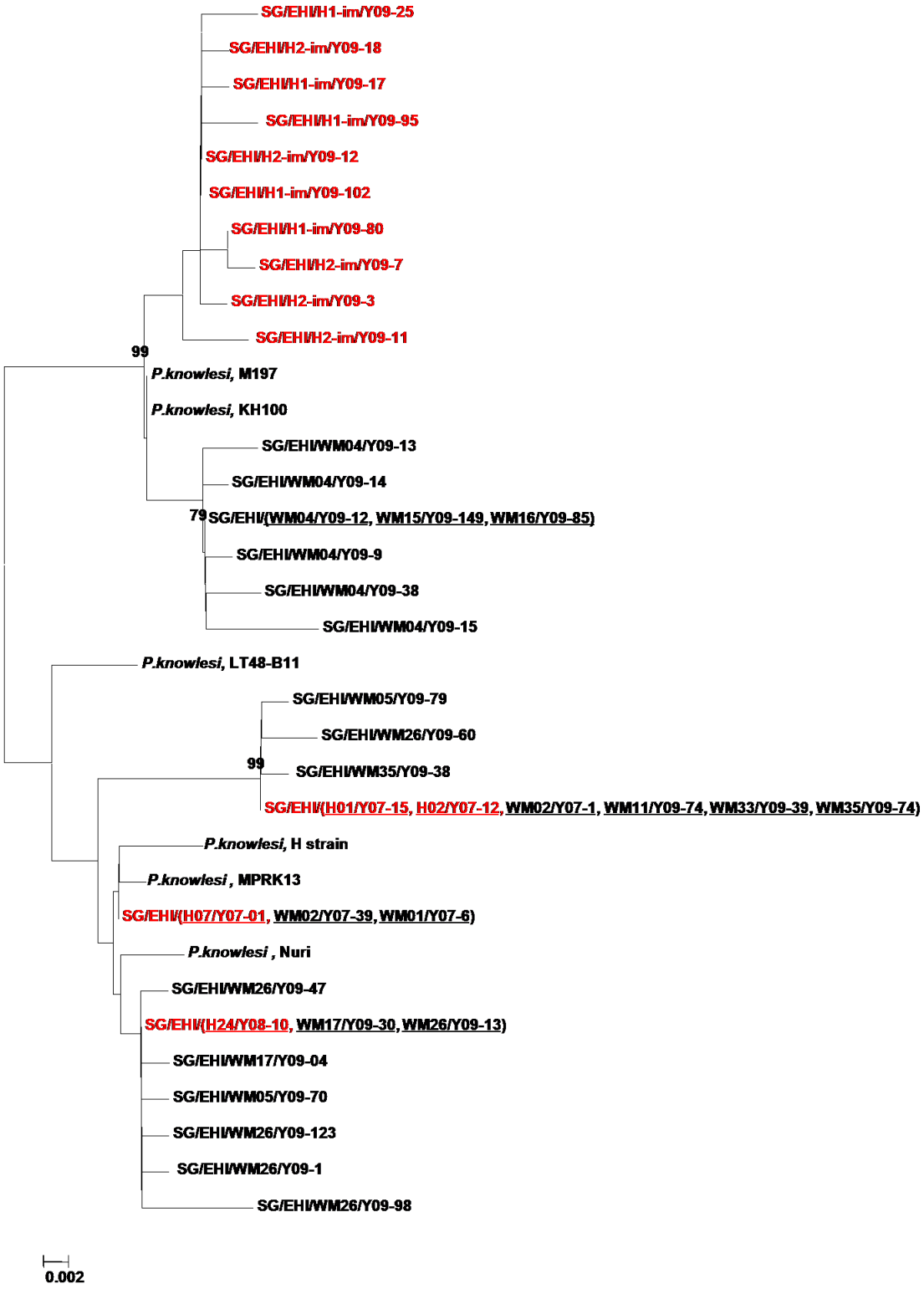

Figure 4.4: Phylogenetic tree of the non-repeat region of the *P. knowlesi csp* genes, constructed using the neighbour-joining method. Clones colored red are isolates from human samples. Figures on the branches are bootstrap percentages based on 1000 replicates, and only bootstrap percentages above 70% are shown.

The *P. fieldi* clade consisted of one transformant each derived from SG/EHI/WM01/Y07, SG/EHI/WM05/Y07, SG/EHI/WM15/Y09, SG/EHI/WM18/Y09 and SG/EHI/WM44/Y10. Transformants that shared identical nucleotide sequences of the non-repeat region (SG/EHI/WM01/Y07-23 with SG/EHI/WM05/Y09-68, and SG/EHI/WM15/Y09-163 with SG/EHI/WM18/Y09-92) were clustered into one subclade (Figure 4.2 and 4.3). On the other hand, the *P. inui* clade consisted of two transformants derived from SG/EHI/WM91/Y11 macaque. Each transformant formed different subclade with high bootstrap values (Figure 4.2 and 4.3).

4.3.3 Polymorphisms of the non-repeat regions of the *Plasmodium* species *csp* gene

4.3.3.1 *P. knowlesi* transformants

The 456 nucleotides sequence coding the non-repeat regions of the *csp* gene from Singapore isolates were aligned with the *P. knowlesi* H strain as reference (Table 4.4). Only single *P. knowlesi csp* genotypes were detected in each local human case, while the number of genotypes presents in each monkey varied from one to six. SG/EHI/WM04/Y09 and SG/EHI/WM26/Y09 harbored six different genotypes each, while two genotypes were detected for monkey SG/EHI/WM35/Y09 and SG/EHI/WM17/Y09. Single genotypes were found in the rest of the monkeys. Interestingly, five genotypes were detected from each of the imported human knowlesi cases. Comparison of Singapore's *P. knowlesi* isolates with the reference H strain showed 53 polymorphic sites. Of these, 24 were due to synonymous mutations

Table 4.4: Gene polymorphisms based on the 456 nucleotide residues encoding the non-repeat region of the *csp* gene of *P. knowlesi* malaria parasites from Singapore's human and long-tailed macaques (in bold). Nucleotide positions are numbered vertically above the polymorphic sites. Dots indicate identical nucleotide residues. Highlighted areas denote non-synonymous mutations.

Strain/ clone	**Nucleotide position**																				
	5	**13**	**49**	**61**	**72**	**75**	**85**	**93**	**94**	**111**	**115**	**122**	**125**	**126**	**148**	**149**	**150**	**152**	**156**	**162**
P. knowlesi H	**A**	**A**	**T**	**C**	**T**	**T**	**T**	**C**	**A**	**A**	**T**	**A**	**T**	**A**	**C**	**A**	**G**	**A**	**G**	**A**
P. knowlesi Nuri	.	.	.	.	.	.	.	A	.	.	.	.	.	.	G	C	A	.	A	.
SG/EHI/H1/Y07-15	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	A	.	.	.
SG/EHI/H2/Y07-12	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	A	.	.	.
SG/EHI/H7/Y07-1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	A	.	.	.
SG/EHI/H24/Y08-10	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	A	.	.	.
SG/EHI/H1-im/Y09-17	G	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	A	.	A	.
SG/EHI/H1-im/Y09-25	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	A	.	A	.
SG/EHI/H1-im/Y09-80	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	A	.	A	.
SG/EHI/H1-im/Y09-95	.	.	.	.	.	.	.	.	.	.	C	.	.	.	G	C	A	.	A	.
SG/EHI/H1-im/Y09-102	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	A	.	A	.
SG/EHI/H2-im/Y09-3	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	A	.	A	.
SG/EHI/H2-im/Y09-7	G	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	A	.	A	.
SG/EHI/H2-im/Y09-11	.	.	.	.	.	.	.	.	.	.	.	.	.	G	G	C	A	.	A	.
SG/EHI/H2-im/Y09-12	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	A	.	A	.
SG/EHI/H2-im/Y09-18	.	.	.	.	.	.	A	.	.	.	.	.	.	.	G	C	A	.	A	.
SG/EHI/WM01/Y07-6	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	A	.	.	.
SG/EHI/WM02/Y07-1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	A	.	.	.
SG/EHI/WM02/Y07-39	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	A	.	.	.
SG/EHI/WM04/Y09-8	.	.	.	T	.	.	.	.	.	G	.	G	.	.	G	C	A	.	A	.
SG/EHI/WM04/Y09-9	.	.	C	T	.	.	.	.	.	.	.	G	.	.	G	C	A	.	A	.
SG/EHI/WM04/Y09-12	.	.	.	T	.	.	.	.	.	.	.	G	.	.	G	C	A	.	A	.
SG/EHI/WM04/Y09-13	.	.	.	T	.	.	.	.	.	.	.	G	.	.	G	C	A	.	A	G
SG/EHI/WM04/Y09-14	.	.	.	T	.	.	.	.	.	.	.	G	.	.	G	C	A	.	A	.
SG/EHI/WM04/Y09-15	.	.	.	T	.	.	.	.	G	.	.	G	.	.	G	C	A	.	A	.
SG/EHI/WM05/Y09-70	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	A	.	.	.
SG/EHI/WM05/Y09-79	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	A	.	.	.
SG/EHI/WM11/Y09-74	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	A	.	.	.
SG/EHI/WM15/Y09-149	.	.	.	T	.	.	.	.	.	.	.	G	.	.	G	C	A	.	A	.
SG/EHI/WM16/Y09-85	.	.	.	T	.	.	.	.	.	.	.	G	.	.	G	C	A	.	A	.
SG/EHI/WM17/Y09-4	.	.	.	.	.	.	.	.	.	.	.	.	C	.	G	C	A	.	.	.
SG/EHI/WM17/Y09-30	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	A	.	.	.
SG/EHI/WM26/Y09-1	.	.	.	.	.	C	.	.	.	.	.	.	.	.	G	C	A	.	.	.
SG/EHI/WM26/Y09-13	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	A	.	.	.
SG/EHI/WM26/Y09-47	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	A	.	.	.
SG/EHI/WM26/Y09-60	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	A	G	.	.
SG/EHI/WM26/Y09-98	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	A	.	.	.
SG/EHI/WM26/Y09-123	.	.	.	.	C	.	.	.	.	.	.	.	.	.	G	C	A	.	.	.
SG/EHI/WM33/Y09-39	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	A	.	.	.
SG/EHI/WM35/Y09-38	.	T	.	.	.	.	.	.	.	.	.	.	.	.	G	C	A	.	.	.
SG/EHI/WM35/Y09-74	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	A	.	.	.

Table 4.4 continued.

Strain/ clone	Nucleotide position																			
	194	197	198	200	201	202	213	219	228	251	276	297	303	305	306	307	308	309	310	315
P. knowlesi H	A	C	A	A	T	G	G	T	C	C	C	T	A	A	A	G	C	T	C	A
P. knowlesi Nuri	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	G	.	.	.
SG/EHI/H1/Y07-15	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	C	A	G	A	T
SG/EHI/H2/Y07-12	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	C	A	G	A	T
SG/EHI/H7/Y07-1	.	.	.	.	.	.	.	.	.	.	.	.	.	.		.	.	.	.	.
SG/EHI/H24/Y08-10	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.
SG/EHI/H1-im/Y09-17	.	.	.	.	.	.	A	.	G	T	.	.	.	G	.	.	.	.	.	.
SG/EHI/H1-im/Y09-25	.	.	.	.	.	.	A	C	G	T	T	.	.	G	.	.	.	.	.	.
SG/EHI/H1-im/Y09-80	.	.	.	.	.	.	A	.	G	T	.	.	.	G	.	.	.	.	.	.
SG/EHI/H1-im/Y09-95	.	.	.	.	.	.	A	.	G	T	.	.	.	G	.	.	.	.	.	.
SG/EHI/H1-im/Y09-102	.	.	.	.	.	.	A	.	G	T	.	.	.	G	.	.	.	.	.	.
SG/EHI/H2-im/Y09-3	.	.	.	.	.	.	A	.	G	T	.	.	.	G	.	.	.	.	.	.
SG/EHI/H2-im/Y09-7	.	.	.	.	.	.	A	.	G	T	.	.	.	G	.	.	.	.	.	.
SG/EHI/H2-im/Y09-11	.	.	.	.	.	.	A	.	G	T	.	.	.	G	.	.	.	.	.	.
SG/EHI/H2-im/Y09-12	.	.	.	.	.	.	A	.	G	T	.	.	.	G	.	.	.	.	.	.
SG/EHI/H2-im/Y09-18	.	.	.	.	.	.	A	.	G	T	.	.	.	G	.	.	.	.	.	.
SG/EHI/WM01/Y07-6	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SG/EHI/WM02/Y07-1	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	C	A	G	A	T
SG/EHI/WM02/Y07-39	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SG/EHI/WM04/Y09-8	.	.	.	.	.	.	A	.	G	.	.	.	.	G	.	.	.	.	.	.
SG/EHI/WM04/Y09-9	.	.	.	.	.	.	A	.	G	.	.	.	.	G	.	.	.	.	.	.
SG/EHI/WM04/Y09-12	.	.	.	.	.	.	A	.	G	.	.	.	.	G	.	.	.	.	.	.
SG/EHI/WM04/Y09-13	.	.	.	.	.	.	A	.	G	.	.	.	.	G	.	.	.	.	.	.
SG/EHI/WM04/Y09-14	.	.	.	.	.	.	A	.	G	.	.	.	.	G	.	.	.	.	.	G
SG/EHI/WM04/Y09-15	.	.	.	.	.	.	A	.	G	.	.	.	G	G	.	.	.	.	.	.
SG/EHI/WM05/Y09-70	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	G	.	.	.
SG/EHI/WM05/Y09-79	G	.	.	.	.	.	.	.	.	.	.	.	.	G	.	C	A	G	A	T
SG/EHI/WM11/Y09-74	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	C	A	G	A	T
SG/EHI/WM15/Y09-149	.	.	.	.	.	.	A	.	G	.	.	.	.	G	.	.	.	.	.	.
SG/EHI/WM16/Y09-85	.	.	.	.	.	.	A	.	G	.	.	.	.	G	.	.	.	.	.	.
SG/EHI/WM17/Y09-4	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.
SG/EHI/WM17/Y09-30	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.
SG/EHI/WM26/Y09-1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.
SG/EHI/WM26/Y09-13	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.
SG/EHI/WM26/Y09-47	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.
SG/EHI/WM26/Y09-60	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	A	.	.	T
SG/EHI/WM26/Y09-98	.	A		T	G	A	.	.	.	.	.	.	.	.	.	.	G	.	.	.
SG/EHI/WM26/Y09-123	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.
SG/EHI/WM33/Y09-39	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	C	A	G	A	T
SG/EHI/WM35/Y09-38	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	C	A	G	A	T
SG/EHI/WM35/Y09-74	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	C	A	G	A	T

Table 4.4 continued.

Strain/ clone	Nucleotide position															
	3 1 7	**3 2 1**	**3 2 6**	**3 3 2**	**3 4 3**	**3 4 8**	**3 5 7**	**3 8 0**	**3 8 4**	**4 0 1**	**4 1 2**	**4 2 0**	**4 2 6**	**4 3 1**	**4 3 5**	**4 3 9**
P. knowlesi H	**G**	**T**	**A**	**A**	**A**	**T**	**G**	**A**	**C**	**T**	**T**	**C**	**C**	**T**	**A**	**C**
P. knowlesi Nuri	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SG/EHI/H1/Y07-15	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T
SG/EHI/H2/Y07-12	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T
SG/EHI/H7/Y07-1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SG/EHI/H24/Y08-10	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SG/EHI/H1-im/Y09-17	A	G	.	.	.	.	A	.	T	.	.	G	A	.	.	.
SG/EHI/H1-im/Y09-25	A	G	.	.	.	.	A	.	T	.	.	G	A	.	.	.
SG/EHI/H1-im/Y09-80	A	G	.	.	.	.	A	.	T	.	.	G	A	.	.	.
SG/EHI/H1-im/Y09-95	A	G	.	.	.	.	A	.	T	.	.	G	A	.	.	.
SG/EHI/H1-im/Y09-102	A	G	.	.	.	.	A	G	T	.	.	G	A	.	G	.
SG/EHI/H2-im/Y09-3	A	G	.	.	G	.	A	.	T	.	.	G	A	.	.	.
SG/EHI/H2-im/Y09-7	A	G	.	.	.	.	A	.	T	.	A	G	A	.	.	.
SG/EHI/H2-im/Y09-11	A	G	G	.	.	.	.	.	T	.	.	G	A	.	.	.
SG/EHI/H2-im/Y09-12	A	G	.	.	.	.	A	.	T	.	.	G	A	.	.	.
SG/EHI/H2-im/Y09-18	A	G	.	.	.	.	A	.	T	.	.	G	A	.	.	.
SG/EHI/WM01/Y07-6	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SG/EHI/WM02/Y07-1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T
SG/EHI/WM02/Y07-39	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SG/EHI/WM04/Y09-8	A	G	.	.	.	.	A	.	T	.	.	G	.	C	.	.
SG/EHI/WM04/Y09-9	A	G	.	.	.	.	A	.	T	.	.	G	.	.	.	.
SG/EHI/WM04/Y09-12	A	G	.	.	.	.	A	.	T	.	.	G	.	.	.	.
SG/EHI/WM04/Y09-13	A	G	.	.	G	.	A	.	T	.	.	G	.	.	.	.
SG/EHI/WM04/Y09-14	A	G	.	.	.	.	A	.	T	.	.	G	.	.	.	.
SG/EHI/WM04/Y09-15	A	G	.	G	.	.	A	.	T	C	.	G	.	.	.	.
SG/EHI/WM5/Y09-70	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SG/EHI/WM5/Y09-79	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T
SG/EHI/WM11/Y09-74	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T
SG/EHI/WM15/Y09-149	A	G	.	.	.	.	A	.	T	.	.	G	.	.	.	.
SG/EHI/WM16/Y09-85	A	G	.	.	.	.	A	.	T	.	.	G	.	.	.	.
SG/EHI/WM17/Y09-4	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SG/EHI/WM17/Y09-30	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SG/EHI/WM26/Y09-1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SG/EHI/WM26/Y09-13	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SG/EHI/WM26/Y09-47	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SG/EHI/WM26/Y09-60	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	T
SG/EHI/WM26/Y09-98	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SG/EHI/WM26/Y09-123	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
SG/EHI/WM33/Y09-39	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T
SG/EHI/WM35/Y09-38	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T
SG/EHI/WM35/Y09-74	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T

while the rest were due to non-synonymous mutations. Unique haplotypes CAGAT (at positions 307, 308, 309, 310 and 315) were detected in two human samples (SG/EHI/H1/07 and SG/EHI/H2/07) and five monkey samples (SG/EHI/WM02/Y07, SG/EHI/WM05/Y09, SG/EHI/WM11/Y09, SG/EHI/WM33/Y09 and SG/EHI/WM35/Y09). Both human samples and SG/EHI/WM02/Y07 were isolated in 2007 while SG/EHI/WM05/Y09, SG/EHI/WM11/Y09, SG/EHI/WM33/Y09 and SG/EHI/WM35/Y09 were all isolated in 2009. Unique nucleotide sequence at position 308 (G) were found in nine transformants from one human (SG/EHI/H24/Y08-10) and four monkey (SG/EHI/WM05/Y09-70, SG/EHI/WM17/Y09-4/30, SG/EHI/WM26/Y09-1/98/47/123/134) samples. Two unique nucleotide sequences at position 61 (T) and 122 (G) was detected in eight transformants from three monkeys (SG/EHI/WM04/Y09-8/9/12/13/14/15, SG/EHI/WM15/Y09-149 and SG/EHI/WM16/Y09-85). Similarly, two unique nucleotide sequences at position 251 (T) and 426 (A) were also detected in both imported human knowlesi case. No unique polymorphism can be found in transformant SG/EHI/H7/Y07-1 and SG/EHI/WM02/Y07-39 and SG/EHI/WM01/Y07-6. Interestingly, *P. knowlesi* transformants with shared unique haplotypes were found to cluster in the same subclade (Figure 4.4).

The pair-wise divergence of different *P. knowlesi* transformants from both monkeys and human cases ranged from 0.2% to 4.8% (Table 4.5).

Table 4.5: Percentage divergence of the non-repeat regions of the *P. knowlesi* clones calculated with the Kimura-2 parameter, using transitions and transversions.

	Isolate/ clone	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
1	*P. knowlesi* H	***																												
2	*P. knowlesi* Nuri	1.6	***																											
3	SG/EHI/H1/Y07-15, SG/EHI/H2/Y07-12, **SG/EHI/WM2/Y07-1**, **SG/EHI/WM11/Y09-74**, **SG/EHI/WM33/Y09-39**, **SG/EHI/WM35/Y09-74**	2.2	2.2	***																										
4	SG/EHI/H1-im/Y09-17	3.4	3.1	3.8	***																									
5	SG/EHI/H1-im/Y09-25	3.6	3.4	4.1	0.7	***																								
6	SG/EHI/H1-im/Y09-95	3.4	3.1	3.8	0.4	0.7	***																							
7	SG/EHI/H1-im/Y09-102	3.6	3.4	4.1	0.7	0.9	0.7	***																						
8	SG/EHI/H2-im/Y09-3	3.4	3.1	3.8	0.4	0.7	0.4	0.7	***																					
9	SG/EHI/H2-im/Y09-7	3.6	3.4	4.1	0.2	0.9	0.7	0.9	0.7	***																				
10	SG/EHI/H2-im/Y09-11	3.4	3.1	3.8	0.9	1.1	0.9	1.1	0.9	1.1	***																			
11	SG/EHI/H1-im/Y09-80, SG/EHI/H2-im/Y09-12	3.1	2.9	3.6	0.2	0.4	0.2	0.4	0.2	0.4	0.7	***																		
12	SG/EHI/H2-im/Y09-18	3.4	3.1	3.8	0.4	0.7	0.4	0.7	0.4	0.7	0.9	0.2	***																	
13	**SG/EHI/WM1/Y07-**6, SG/EHI/H7/Y07-1, **SG/EHI/WM2/Y07-39**	0.7	0.9	1.6	2.7	2.9	2.7	2.9	2.7	2.9	2.7	2.5	2.7	***																
14	**SG/EHI/WM4/Y09-8**	3.6	3.4	4.1	1.6	1.8	1.6	1.8	1.6	1.8	2	1.3	1.6	2.9	***															
15	**SG/EHI/WM4/Y09-9**	3.4	3.1	3.8	1.3	1.6	1.3	1.6	1.3	1.6	1.8	1.1	1.3	2.7	0.7	***														
16	**SG/EHI/WM4/Y09-13**	3.6	3.4	4.1	1.6	1.8	1.6	1.8	1.1	1.8	2	1.3	1.6	2.9	0.9	0.7	***													
17	**SG/EHI/WM4/Y09-14**	3.4	3.1	3.6	1.3	1.6	1.3	1.6	1.3	1.6	1.8	1.1	1.3	2.7	0.7	0.4	0.7	***												
18	**SG/EHI/WM4/Y09-15**	4.1	3.8	4.5	2	2.2	2	2.2	2	2.2	2.5	1.8	2	3.4	1.3	1.1	1.3	1.1	***											
19	**SG/EHI/WM5/Y09-70**	1.1	0.9	1.8	3.1	3.4	3.1	3.4	3.1	3.4	3.1	2.9	3.1	0.4	3.4	3.1	3.4	3.1	3.8	***										
20	**SG/EHI/WM5/Y09-79**	2.5	2.5	0.2	4.1	4.3	4.1	4.3	4.1	4.3	4.1	3.8	4.1	1.8	4.3	4.1	4.3	3.8	4.8	2	***									
21	**SG/EHI/WM15/Y09-149, SG/EHI/WM4/Y09-12, SG/EHI/WM16/Y09-85**	3.1	2.9	3.6	1.1	1.3	1.1	1.3	1.1	1.3	1.6	0.9	1.1	2.5	0.4	0.2	0.4	0.2	0.9	2.9	3.8	***								
22	**SG/EHI/WM17/Y09-4**	1.1	0.9	1.8	3.1	3.4	3.1	3.4	3.1	3.4	3.1	2.9	3.1	0.4	3.4	3.1	3.4	3.1	3.8	0.4	2	2.9	***							
23	**SG/EHI/H24/Y08-10, SG/EHI/WM17/Y09-30, SG/EHI/WM26/Y09-13**	0.9	0.7	1.6	2.9	3.1	2.9	3.1	2.9	3.1	2.9	2.7	2.9	0.2	3.1	2.9	3.1	2.9	3.6	0.2	1.8	2.7	0.2	***						
24	**SG/EHI/WM26/Y09-1**	1.1	0.9	1.8	3.1	3.4	3.1	3.4	3.1	3.4	3.1	2.9	3.1	0.4	3.4	3.1	3.4	3.1	3.8	0.4	2	2.9	0.4	0.2	***					
25	**SG/EHI/WM26/Y09-47**	1.1	0.9	1.8	3.1	3.4	3.1	3.4	3.1	3.4	3.1	2.9	3.1	0.4	3.4	3.1	3.4	3.1	3.8	0.4	2	2.9	0.4	0.2	0.4	***				
26	**SG/EHI/WM26/Y09-60**	2.7	2.7	0.4	4.3	4.5	4.3	4.5	4.3	4.5	4.3	4.1	4.3	2	4.5	4.3	4.5	4.1	5	2.2	0.7	4.1	2.2	2	2.2	2.2	***			
27	**SG/EHI/WM26/Y09-98**	1.8	1.6	2.5	3.8	4.1	3.8	4.1	3.8	4.1	3.8	3.6	3.8	1.1	4.1	3.8	4.1	3.8	4.5	1.1	2.7	3.6	1.1	0.9	1.1	1.1	2.9	***		
28	**SG/EHI/WM26/Y09-123**	1.1	0.9	1.8	3.1	3.4	3.1	3.4	3.1	3.4	3.1	2.9	3.1	0.4	3.4	3.1	3.4	3.1	3.8	0.4	2	2.9	0.4	0.2	0.4	0.4	2.2	1.1	***	
29	**SG/EHI/WM35/Y09-38**	2.5	2.5	0.2	4.1	4.3	4.1	4.3	4.1	4.3	4.1	3.8	4.1	1.8	4.3	4.1	4.3	3.8	4.8	2	0.4	3.8	2	1.8	2	2	0.7	2.7	2	***

4.3.3.2 *P. cynomolgi* transformants

Circumsporozoite protein gene sequence alignment between transformants from seven long-tailed macaques and *P. cynomolgic* Ceylon strain (as reference) disclosed 31 polymorphic sites. Twelve of these polymorphic sites were due to synonymous mutations while 19 were non-synonymous mutations (Table 4.6). Pair-wise sequence divergence between these transformants ranged from 0.2% to 4.5% (Table 4.7). After the sequence were aligned, there were two genotypes each for SG/EHI/WM16/Y09 and SG/EHI/WM44/10, while only one genotype each were observed for SG/EHI/WM02/Y07, SG/EHI/WM05/Y09, SG/EHI/WM18/Y09, SG/EHI/WM33/Y09 and SG/EHI/WM42/Y10. Unique haplotypes GCGG (at positions 21, 63, 79 and 119) were detected in SG/EHI/WM16/Y09 and SG/EHI/WM33/Y09, while haplotypes AACA (at positions 247, 250, 251 and 302) were detected in SG/EHI/WM05/Y09 and SG/EHI/WM44/Y10 (Table 4.6).

Table 4.6: Gene polymorphisms based on the 456 nucleotide residues encoding the non-repeat region of the *csp* gene of *P. cynomolgi* malaria parasites from Singapore's long-tailed macaques (in bold). Nucleotide positions are numbered vertically above the polymorphic sites. Dots indicate identical nucleotide residues. Highlighted areas denote non-synonymous mutations.

Strain/ clone	**Nucleotide position**																															
	1	**2**	**2**	**3**	**5**	**5**	**6**	**6**	**6**	**6**	**7**	**8**	**8**	**1**	**1**	**3**	**2**	**0**	**2**	**1**	**2**	**4**	**4**	**2**	**5**	**5**	**5**	**2**	**6**	**3**	**0**	**1**	**4**	**5**	**3**	**6**	**4**	**0**	**4**	**1**
	4	**1**	**3**	**9**	**4**	**6**	**1**	**3**	**6**	**9**	**9**	**7**	**9**	**9**	**8**	**3**	**0**	**4**	**4**	**7**	**0**	**1**	**2**	**6**	**2**	**1**	**4**	**8**	**0**	**8**	**8**
P. cynomolgi, Ceylon	**A**	**A**	**C**	**G**	**C**	**G**	**C**	**G**	**G**	**A**	**C**	**C**	**A**	**A**	**C**	**T**	**T**	**A**	**A**	**G**	**G**	**T**	**C**	**G**	**G**	**A**	**C**	**A**	**T**	**C**	**G**
P. cynomolgi, Berok	.	G	T	T	T	C	A	C	C	G	.	T	G	.	A	C	C	.	C	.	.	.	T	.	G	G	T	.	.	T	C
SG/EHI/WM02/Y07-110	T	.	.	.	T	C	A	C	C	G	G	.	G	.	.	.	.	C	.	.	.	.	.	.	.	G	.	.	.	T	.
SG/EHI/WM05/Y09 -65	T	.	.	.	T	C	A	C	C	G	G	.	G	.	A	.	.	C	.	A	A	C	.	.	A	.	T	G	A	T	.
SG/EHI/WM16/Y09-2	T	G	.	.	.	.	.	C	.	.	G	.	.	G	A	.	.	.	.	.	.	.	.	A	.	G	.	.	.	T	.
SG/EHI/WM16/Y09-34	T	G	.	.	.	.	.	C	.	.	G	.	.	G	A	.	.	.	.	.	.	.	.	.	.	G	.	.	.	T	.
SG/EHI/WM18/Y09-24	T	.	.	.	T	C	A	C	C	G	G	.	G	.	.	.	.	C	.	.	.	.	.	.	.	G	.	.	.	T	.
SG/EHI/WM33/Y09-47	T	G	.	.	.	.	.	C	.	.	G	.	.	G	A	.	.	.	.	.	.	.	.	.	.	G	T	G	A	T	.
SG/EHI/WM42/Y10-1	T	.	.	.	T	C	A	C	C	G	G	.	G	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.
SG/EHI/WM44/Y10-3	T	.	.	.	T	C	A	C	C	G	G	.	G	.	A	.	.	C	.	A	A	C	.	.	A	.	T	G	A	T	.
SG/EHI/WM44/Y10-30	T	.	.	.	T	C	A	C	T	G	G	.	G	.	A	.	.	C	.	A	A	C	.	.	A	.	T	G	A	T	.

Table 4.7: Percentage divergence of the non-repeat regions of the *P. cynomolgi* clones calculated with the Kimura-2 parameter, using transitions and transversions.

	Strain/ clone	1	2	3	4	5	6	7	8	9	10	11
1	*P. cynomolgi,* Ceylon	***	-	-	-	-	-	-	-	-	-	-
2	*P. cynomolgi,* Berok	4.5	***	-	-	-	-	-	-	-	-	-
3	**SG/EHI/WM02/Y07-110**	2.7	3.1	***	-	-	-	-	-	-	-	-
4	**SG/EHI/WM05/Y09 -65**	4.3	4.3	2	***	-	-	-	-	-	-	-
5	**SG/EHI/WM16/Y09-2**	2	4.3	2.5	4.1	***	-	-	-	-	-	-
6	**SG/EHI/WM16/Y09-34**	1.8	4.1	2.2	3.8	0.2	***	-	-	-	-	-
7	**SG/EHI/WM18/Y09-24**	2.7	3.1	0	2	2.5	2.2	***	-	-	-	-
8	**SG/EHI/WM33/Y09-47**	2.5	4.3	2.9	3.1	0.9	0.7	2.9	***	-	-	-
9	**SG/EHI/WM42/Y10-1**	2.2	3.6	0.4	2	2.9	2.7	0.4	3.4	***	-	-
10	**SG/EHI/WM44/Y10-3**	4.3	4.3	2	0	4.1	3.8	2	3.1	2	***	-
11	**SG/EHI/WM44/Y10-30**	4.3	4.5	2.2	0.2	4.1	3.8	2	3.1	2.2	2	***

4.3.3.3 *P. fieldi* transformants

The non-repeat regions of the *csp* gene from Singapore isolates were aligned and compared with *P. fieldi* reference strain obtained from CDC (Table 4.8). Single genotype was observed for each of the monkeys infected with *P. fieldi*. Comparison of *P. fieldi* isolated from Singapore monkeys with the reference strain showed seven polymorphic sites. Four of the polymorphic sites were synonymous while the other three were non-synonymous mutations. Unique haplotype AACCGG (at position 21, 92, 110, 222, 311 and 432) were detected in SG/EHI/WM01/Y07 and SG/EHI/WM05/Y09 that were caught in 2007 and 2009, respectively. SG/EHI/WM15/Y09 and SG/EHI/WM18/Y09, both caught in 2009, had identical genotype with a unique nucleotide at position 93 (A), 222 (C) and 432 (G). On the other hand, there were only two polymorphic sites found in SG/EHI/WM44/Y10-64 at nucleotide positions 222 (C) and 432 (G). The pair-wise divergence between different transformants of *P. fieldi* obtained from Singapore long-tailed macaques ranged from 0.2% to 1.3% (Table 4.9).

Table 4.8: Gene polymorphisms based on the 456 nucleotide residues encoding the non-repeat region of the *csp* gene of *P. fieldi* malaria parasites obtained from Singapore's long-tailed macaques (in bold). Nucleotide positions are numbered vertically above the polymorphic sites. Dots indicate identical nucleotide residues. High-lighted areas denote non-synonymous mutations.

	Nucleotide position						
				1	**2**	**3**	**4**
Strain/ clone	**2**	**9**	**9**	**1**	**2**	**1**	**3**
	1	**2**	**3**	**0**	**2**	**1**	**2**
P. fieldi (CDC)	**G**	**C**	**C**	**T**	**T**	**A**	**A**
SG/EHI/WM01/Y07-23	A	A	.	C	C	G	G
SG/EHI/WM05/Y09-68	A	A	.	C	C	G	G
SG/EHI/WM15/Y09-163	.	.	A	.	C	.	G
SG/EHI/WM18/Y09-92	.	.	A	.	C	.	G
SG/EHI/WM44/Y10-64			.	.	C	.	G

Table 4.9: Percentage divergence of the non-repeat regions of the *P. fieldi* clones calculated with the Kimura-2 parameter, using transitions and transversions.

	Strain/ clone	**1**	**2**	**3**	**4**	**5**	**6**
1	*P.fieldi* (CDC)	***	-	-	-	-	-
2	**SG/EHI/WM01/Y07-23**	1.3	***	-	-	-	-
3	**SG/EHI/WM05/Y09-68**	1.3	0	***	-	-	-
4	**SG/EHI/WM15/Y09-163**	0.7	1.1	1.1	***	-	-
5	**SG/EHI/WM18/Y09-92**	0.7	1.1	1.1	0	***	-
6	**SG/EHI/WM44/Y10-64**	0.4	0.9	0.9	0.2	0.2	***

4.3.3.4 *P. inui* transformants

Alignment of the non-repeat regions of *P. inui csp* gene sequence between transformants derived from SG/EHI/WM91/Y11 and *P. inui* obtained from CDC revealed a total of 29 polymorphic sites where 10 synonymous and 19 non-synonymous mutations were detected. When transformant SG/EHI/WM91/Y11-75 was compared to the CDC reference strain, there were only seven polymorphic sites and pair-wise sequence divergence of 1.6% (Table 4.10 and 4.11). However, when transformant SG/EHI/WM91/Y11-84 were compared to the CDC reference strain, there were 26 polymorphic sites and pair-wise sequence divergence was 5.9%. On the other hand, when SG/EHI/WM91/Y11-84 was compared to *P. inui* isolated from East Malaysia (GU002523), there were only 18 polymorphic sites observed with a pair-wise sequence divergence of 4.8%. In contrast, there were 33 polymorphic sites and a pair-wise sequence divergence of 7.6% when SG/EHI/WM91/Y11-75 was compared to the same East Malaysian *P. inui* strain.

Table 4.10: Gene polymorphisms based on the 456 nucleotide residues encoding the non-repeat region of the *csp* gene of *P. inui* malaria parasites obtained from Singapore's long-tailed macaques (in bold). Nucleotide positions are numbered vertically above the polymorphic sites. Dots indicate identical nucleotide residues. Highlighted areas denote non-synonymous mutations.

Strain/ clone	**Nucleotide position**																																									
	1	**4**	**5**	**7**	**8**	**9**	**0**	**4**	**4**	**5**	**6**	**7**	**7**	**7**	**9**	**9**	**0**	**2**	**3**	**3**	**3**	**4**	**6**	**7**	**8**	**8**	**8**	**1**	**1**	**4**	**4**	**5**	**5**	**5**	**7**	**1**	**2**	**3**	**3**	**3**
	8	**8**	**1**	**3**	**5**	**0**	**4**	**5**	**7**	**3**	**7**	**1**	**5**	**7**	**3**	**5**	**3**	**8**	**0**	**6**	**8**	**2**	**3**	**5**	**2**	**3**	**4**	**0**	**8**	**3**	**4**	**2**	**8**	**9**	**8**	**1**	**6**	**3**	**5**	**9**
P. inui (CDC)	**C**	**C**	**A**	**G**	**T**	**A**	**A**	**G**	**C**	**G**	**C**	**C**	**G**	**C**	**G**	**C**	**A**	**G**	**A**	**A**	**T**	**A**	**T**	**G**	**A**	**C**	**A**	**A**	**T**	**C**	**T**	**C**	**G**	**T**	**G**	**T**	**A**	**C**	**A**	**C**
P. inui (GU002523)	T	A	G	A	.	G	G	.	A	C	T	.	A	T	T	.	C	C	.	G	C	C	C	C	.	A	C	.	A	A	C	G	A	C	T	C	C	T	.	A
SG/EHI/WM91/Y11-75	.	.	.	A	.	G	.	.	.	.	.	.	.	.	A	A	.	C	G	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.
SG/EHI/WM91/Y11-84	.	.	.	A	A	G	.	C	.	.	A	A	.	.	A	.	C	C	.	.	C	C	C	C	G	A	C	G	A	.	C	.	.	C	T	C	C	T	G	A

Table 4.11: Percentage divergence of the non-repeat regions of the *P. inui* clones calculated with the Kimura-2 parameter, using transitions and transversions.

	Strain/ clone	**1**	**2**	**3**	**4**
1	*P. inui* (CDC)	***	-	-	-
2	*P. inui* (GU002523)	7.6	***	-	-
3	**SG/EHI/WM91/Y11-75**	1.6	7.6	***	-
4	**SG/EHI/WM91/Y11-84**	5.9	4.8	5.2	***

4.3.4 Polymorphisms within the Region I, Region II-plus and the central tandem repeat region of the *Plasmodium* species *csp* gene

4.3.4.1 *P. knowlesi* transformants

The region I of the *csp* genes (based on the short amino acid motif KLKQP), were found to be conserved in all *P. knowlesi* transformants obtained in this study (Table 4.12). The nucleotide sequence of region II-plus was also found to be conserved among all human and monkey samples, except at position 18, which was substituted with either arginine (R) or lysine (K) (Table 4.12). The central repeat regions *of P. knowlesi* transformants were highly variable with size ranging from 81 bp to 657 bp, and can be categorized into amino acid consensus groupings (Table 4.13). *Plasmodium knowlesi* transformants isolated from local human cases included the following amino acid groupings N(A/E)GQPQAQGD(G/R)A (H7-1), EQPA(A/P)(G/A)(A/P)(G/R)(G/R/A) (H1-15 and H2-12) and E(E/Q)PAPG(R/G)E(Q/E)PAP(G/A)(R/P) (H24-10). Interestingly, several transformants isolated from monkeys also showed similar amino acid motif groupings. To date, the amino acid motifs, (Q/E)GNGGAGQAQP and EGNREAPAQP were only found in monkeys. In contrast, the amino acid motifs (NAEGGAN(A/V)(G/R)QP and NAGGANAGQP) deduced from the two imported human cases were unique and not found among *P. knowlesi* transformants isolated from local monkeys and human cases.

Table 4.12: Comparison of amino acid sequences in the region I and region II-plus of the *P. knowlesi* H and Nuri strain, and isolates from the human and macaque samples. Dots represent identical amino acid and those underlined indicate substituted residue.

Strain/ clone	**Region I**	**Region II-plus**
P. knowlesi H	KLKQP	EWTPCSVTCGNGVRIRRK
P. knowlesi Nuri	. . . . .	. . . . .
SG/EHI/H1/Y07 - 15	. . . . .	. . . . . R
SG/EHI/H2/Y07 - 12	. . . . .	. . . . . R
SG/EHI/H7/Y07 - 1	. . . . .	. . . . .
SG/EHI/H24/Y08 - 10	. . . . .	. . . . .
SG/EHI/H1-im/Y09-17/18//80/95/100	. . . . .	. . . . . R
SG/EHI/H2-im/Y09-3/7/11/12/18	. . . . .	. . . . . R
SG/EHI/WM01/Y07 - 6	. . . . .	. . . . .
SG/EHI/WM02/Y07 - 1	. . . . .	. . . . . R
SG/EHI/WM02/Y07 - 39	. . . . .	. . . . . R
SG/EHI/WM04/Y09-3/6/8/9/12/13/14/15	. . . . .	. . . . . R
SG/EHI/WM05/Y09 - 70	. . . . .	. . . . .
SG/EHI/WM05/Y09 - 79	. . . . .	. . . . . R
SG/EHI/WM11/Y09 - 74	. . . . .	. . . . . R
SG/EHI/WM15/Y09 - 149	. . . . .	. . . . . R
SG/EHI/WM16/Y09 - 85	. . . . .	. . . . . R
SG/EHI/WM17/Y09 - 4/18/30	. . . . .	. . . . .
SG/EHI/WM26/Y09 - 1/13/47/98/123	. . . . .	. . . . .
SG/EHI/WM26 /Y09- 60	. . . . .	. . . . . R
SG/EHI/WM33 /Y09- 39	. . . . .	. . . . . R
SG/EHI/WM35/Y09 - 38/74	. . . . .	. . . . . R
SG/EHI/WM91/Y11 - 24	. . . . .	. . . . . R

Table 4.13: Comparison of amino acid motifs and the sequence size of the tandem repeat region and full *csp* gene for *P. knowlesi* H and Nuri strain, and isolates from human and macaque samples

Isolate/ clone/ strain	**Amino acid motifs of tandem repeat region (single alphabet code for each motif)**	**No. repeats**	**Sequence of tandem repeats**	**Tandem repeat region size (bp)**	**Full length of *csp* gene (bp)**
P. knowlesi H	NEGQPQAQGDGA (A) NAGQPQAQGDGA (B)	1 11	ABBBBBBBBBBB	432	1092
P. knowlesi Nuri	EQPAAGAGG (C) EQPAAGARG (D) EQPAPAPRR (E)	11 3 1	CCCDCCCCCDCCD CE	405	1056
H1-15*, H2-12*, WM11-74*, WM26 - 60, WM33 - 39*, WM35 - 38, WM35 - 74*	EQPAAGAGG (C) EQPAPAPRR (E) EQPAPGAGA (F)	12 2 1	CCCCCCCCCCCEC EF	405	1026
H7-1	NEGQPQAQGDGA (G) NAGQPQAQGDGA (H) NAGQPQAQGDRA (I)	1 10 1	GHHHHHHHHHIH	432	1092
H24-10*, WM5 - 70, WM17 - 4, WM17-18*, WM17-30*, WM26 -1, WM26 - 13*, WM26 - 47, WM26 - 123	EEPAPGREQPAPGR (J) EQPAPGREEPAPGR (K) EQPAPGREQPAPGR (L) EQPAPGGEQPAPGR (M) EQPAPGGEQPAPAP (N)	4 2 1 1 1	JJJJKKLMN	378	1020
H1-im-17, H1-im-25, H1-im-100, H2-im-3, H2-im-7, H2-im-11, H2-im-18	NAEGGANAGQP (O) NAGGANAGQP (P) NAEGGANARQP (Q)	10 1 4	OOOPOOQQOOOQ OOQ	492	1278
H1-im-80	NAEGGANAGQP (O) NAEGGANARQP (Q)	4 2	OOQOOQ	198	951
H1-im-95	NAEGGANAGQP (O) NAGGANAGQP (P) NAEGGANVGQP (R) NAEGGANARQP (Q)	9 1 1 4	OOOPORQQOOOQ OOQ	492	1242

Table 4.13 continued

Strain/ clone	Amino acid motifs of tandem repeat region (single alphabet code for each motif)	No. repeats	Sequence of tandem repeats	Tandem repeat region size (bp)	Full length of *csp* gene (bp)
H2-im-12	NAEGGANAGQP (O) NAGGANAGQP (P) NAEGGANARQP (Q)	14 1 5	OOOPOOQQOOOQ OOOOQOOQ	657	1443
WM04 - 3	QGNGGAGQAQP (S) EGNGGAGQAQP (T)	6 2	STTSSSSS	264	879
WM04 - 6*, WM04-12*, WM04 - 13, WM04 - 14, WM16-85*, WM15 - 149 WM91 - 24*	QGNGGAGQAQP (S) EGNGGAGQAQP (T) EGNREAPAQP (U)	5 9 1	STTTTTTTTTSSSUS	492	1107
WM04 - 8	QGNGGAGQAQP (S) EGNREAPAQP (U)	3 1	SSUS	129	744
WM04 - 9	QGNGGAGQAQP (S) EGNGGAGQAQP (T)	2 12	STTTTTTTTTTS	462	1077
WM04-15	QGNGGAGQAQP (S) EGNGGAGQAQP (T) EGNREAPAQP (U)	7 8 1	STTTTTTTSSSSTSU S	525	1140
WM05-79	EQPAAGAGG (C) EQPAPAPRR (E) EQPAPGAGA (F)	1 1 1	CEF	81	702
WM26 - 98	EEPAPGREQPAPGR (J) EQPAPGREEPAPGR (K) EQPAPGREEPAPAP (V)	4 1 1	JJJJKV	252	894

* Identical csp non-repeat sequences (within same group)

Prefix SG/EHI/- had been omitted in this table

4.3.4.2 *P. cynomolgi* transformants

The region I and region II-plus of *P. cynomolgi* transformants derived from all monkey isolates were found to be identical with the exception of SG/EHI/WM16/09-2 and SG/EHI/WM44/10-3/30 (Table 4.14). The region I was based on the short amino acid motif KLKQP, while region II-plus was based on EWSPCSVTCGKGVRMRRK. In SG/EHI/WM16/09-2, position five of the region II-plus was substituted with a tyrosine (Y) instead of cysteine (C), while in SG/EHI/WM44/10-3/30, the position 17 of the region II-plus was substituted with a lysine (K) instead of arginine (R). The central repeat regions of *P. cynomolgi* transformants were found to be variable, with sizes ranging from 369 bp to 585 bp and composed of the following amino acid motif sequences: DGNNAA, DGGVQPPA(G/A)GGN(N/R)A, PAAADGA, PAAGGN, QAG(A/G)Q(A/P)G(G/A)(G/N)(N/A), QAGGDAGNA, QAGGA, AAANAGDGQP, AANAGGA and QAAGGA (Table 4.15). The complete *csp* gene sequences of transformants were identical between SG/EHI/WM02/Y07-110 and SG/EHI/WM18/Y09-24, and between SG/EHI/WM05/Y09-65 and SG/EHI/WM44/Y10-3 (Appendix E).

Table 4.14: Comparison of amino acid sequences in the region I and region II-plus of the *P. cynomolgi* Ceylon and Berok strain, and isolates from the macaque samples. Dots represent identical amino acid and those underlined indicate substituted residue.

Strain/clone	**Region I**	**Region II-plus**
P. cynomolgi Ceylon	KLKQP	EWSPCSVTCGKGVRMRRK
P. cynomolgi Berok	. . . . .	. . . . .
SG/EHI/WM02/Y07 - 110	. . . . .	. . . . .
SG/EHI/WM05/Y09 - 65	. . . . .	. . . . .
SG/EHI/WM16/Y09 - 2	. . . . .	. . . . Y. . . . .
SG/EHI/WM16/Y09 - 34	. . . . .	. . . . .
SG/EHI/WM18/Y09 - 24	. . . . .	. . . . .
SG/EHI/WM33/Y09 - 47	. . . . .	. . . . .
SG/EHI/WM42/Y10 - 1	. . . . .	. . . . .
SG/EHI/WM44/Y10 - 3/ 30	. . . . .	. . . . . K .

Table 4.15: Comparison of amino acid motifs and the sequence size of the tandem repeat region and full *csp* gene for *P. cynomolgi* Ceylon and Berok strain, and isolates from macaque samples

Isolate/ clone/ strain	Amino acid motifs of tandem reoeat region (single alphabet code for each motif)	No. repeats	Sequence of tandem repeats	Tandem repeat region size (bp)	Full length of *csp* gene (bp)
P. cynomolgi Ceylon	AGNNAAAGE (A)	13	AAAAAABABAC ABABBADADAE	585	1197
	AGNNAAGGA(B)	5			
	AGNNAAGGE (C)	1			
	AGAGGAGR (D)	2			
	AGAGGAGG (E)	1			
P. cynomolgi Berok	PAGDGA (F)	1	FGHHHHHHHHH HIJKJKJJLM	537	1137
	PEGDGA (G)	1			
	PAAPAGDGA (H)	10			
	PAGNR (I)	1			
	AGGQPAAGGNQ (J)	3			
	AGGNR (K)	2			
	AGAQAGGNQ (L)	1			
	AGAQAGGAN (M)	1			
SG/EHI/WM2/Y07 - 110*, SG/EHI/WM18/Y09 -24*	DGNNAA (N)	1	NOOOOOOOOOO P	510	1128
	DGGVQPPAGGGNNA A (O)	10			
	DGGVQPPAAGGNRA (P)	1			
SG/EHI/WM5/Y09 - 65*, SG/EHI/WM44/Y10 - 3*, 30	PAAADGA (Q)	10	QQQQQQQQQQR SSSTTTSU	492	1095
	PAAGGN (R)	1			
	QAGAQAGAGGN (S)	4			
	QAGGQPGAGGN (T)	3			
	QAGGQAGGANA (U)	1			
SG/EHI/WM16/Y09 - 2, 34	QAGGDAGNA (V)	10	VVVWVWVWVW VWVWVWVW WWWWWWWW WW	540	1158
	QAGGA (W)	18			
SG/EHI/WM33/Y09-47	QAGGDAGNA (V)	10	VVVWVWVWVW VWVWVWV WWWWWWWW WW	525	1143
	QAGGA (W)	17			
SG/EHI/WM42/Y10-1	AAANAGDGQP (X)	10	XXXXXXXXXXY Zaa	369	978
	AANAGGA (Y)	1			
	QAAGGA (Z)	1			
	QAGGA (a)	2			

* Identical *csp* non-repeat sequences (within same group)

4.3.4.3 *P. fieldi* transformants

The region I and region II-plus of *P. fieldi* transformants derived from all monkey isolates were identical. These regions were based on the short amino acid motif KLKQP and EWTPCSVTCGNGVRLRRK, respectively (Table 4.16). The central repeat regions of *P. fieldi* transformants were variable with size ranging from 396 bp to 561 bp and composed of the following amino acid motif sequences: PGANQ(E/G)G(G/A)(A/K)(A/P)A and (A/G)(N/G)(D/G)AGQNQP (Table 4.17).

The complete *csp* gene sequences of transformants were identical between SG/EHI/WM01/Y07-23 and SG/EHI/WM05/Y09-68, and between SG/EHI/WM15/Y09-163 and SG/EHI/WM44/Y10-64 (Appendix E).

Table 4.16: Comparison of amino acid sequences in the region I and region II-plus of the *P. fieldi* from CDC, and isolates from the macaque samples. Dots represent identical amino acid.

Strain/clone	**Region I**	**Region II-plus**
P. fieldi (CDC)	KLKQP	EWTPCSVTCGKGVRVRRK
SG/EHI/WM01/Y07- 23	. . . . .	. . . . .
SG/EHI/WM05/Y09 - 68	. . . . .	. . . . .
SG/EHI/WM15 Y09- 163	. . . . .	. . . . .
SG/EHI/WM18 Y09 - 92	. . . . .	. . . . .
SG/EHI/WM44 Y10-64	. . . . .	. . . . .

Table 4.17: Comparison of amino acid motifs and the sequence size of the tandem repeat region and full *csp* gene for *P. fieldi* (CDC), and isolates from macaque samples

Isolate/ clone/ strain	**Amino acid motifs of tandem repeat region (single alphabet code for each motif)**	**No. repeats**	**Sequence of tandem repeats**	**Tandem repeat region size (bp)**	**Full length of *csp* gene (bp)**
P. fieldi (CDC), SG/EHI/WM15/Y09-163, SG/EHI/WM44/Y10-64	PGANQEGGAAA (A) PGANQGGGAAA (B) PGANQGGAKPA (C)	13 3 1	AAAAAAAAAAA ABBBAC	561	1149
SG/EHI/WM1/Y07-23*, SG/EHI/WM5/Y09-68*	ANDAGQNQP (D) GGGAGQNQP (E)	13 6	DDDDDDDDDDD DDEEEEEE	513	1131
SG/EHI/WM18/Y09-92	PGANQEGGAAA (A) PGANQGGGAAA (B) PGANQGGAKPA (C)	9 2 1	AAAAAAAABBA C	396	984

* Identical *csp* non-repeat sequences (within same group)

4.3.4.4 *P. inui* transformants

The region I of *P. inui* is based on the amino acid motif NLKQP instead of the usual KLKQP found in *P. knowlesi, P. cynomolgi and P. fieldi*. The region II-plus of transformant SG/EHI/WM91/Y11-61 was identical to the CDC reference strain, while SG/EHI/WM91/Y11-73's region II-plus was identical to that of *P. inui* isolated from East Malaysia (Table 4.18). The size of the central repeat region and the amino acid motif sequences of both transformants varied from each other. SG/EHI/WM91/Y11-61 had a central repeat region of 420bp, with amino acid motif A(G/Q)(D/G/N/K)P(A/G)(P/G)(P/Q), while SG/EHI/WM91/Y11-73's central repeat region was 540 bp long, with the amino acid motif AG(E/Q)(A/P)G(G/A)AGQ(P/A)GA (Table 4.19).

Table 4.18: Comparison of amino acid sequences in the region I and region II-plus of the *P. fieldi* from CDC and East Malaysia, and isolates from the macaque samples. Dots represent identical amino acid and those underlined indicate substituted residue.

Strain/clone	**Region I**	**Region II-plus**
P. inui (CDC)	NLKQP	EWSVCSVSCGQGVRVRRK
P. inui (GU002523)	. . . . .	. . . A . . .T . . T . . . . .
WM91 - 61	. . . . .	. . . . .
WM91 - 73	. . . . .	. . . A . . .T . . T . . . . .

Table 4.19: Comparison of amino acid motifs and the sequence size of the tandem repeat region and full *csp* gene for *P. inui* (CDC), and isolates from macaque samples

Strain/ clone	**Amino acid motifs of tandem reoeat region (single alphabet code for each motif)**	**No. repeats**	**Sequence of tandem repeats**	**Tandem repeat region size(bp)**	**Full length of *csp* gene (bp)**
P. inui (CDC)	AQDPGAP (A)	1	ABBBBBCBBBCBCBDDE	375	1056
	GQDPGAP (B)	10			
	GQAPGAP (C)	3			
	GQDPAAP (D)	2			
	ARDPAAP (E)	1			
SG/EHI/WM91/Y11-61	AGDPAPP (F)	14	FFFFFFFFFFFFFFGHHHII	420	1014
	AGGPAGQ (G)	1			
	AQNPGGP (H)	3			
	AQKPGGP (I)	2			
SG/EHI/WM91/Y11-73	AGEAGGAGQPGA (J)	12	JJJJJJJJJJJKLK	540	1191
	AGQPGAAGQPGA (K)	2			
	AGQAGAAGQAGA (L)	1			

4.4 Discussion

The *csp* gene is a single copy gene which encodes for the highly antigenic surface protein of the sporozoite [96, 105]. Its polymorphism can be observed through cloning and sequencing of the gene isolate. Gene polymorphism can be presented as sequence variation in the conserved non-repeat N- and C- terminal of the *csp* gene, or a difference in the pattern and length of the tandem amino acid motif repeats.

Phylogenetic inferences based on the non-repeat region of *csp* gene from the *Plasmodium* species isolates of 15 infected wild long-tailed macaques confirmed the presence of four species of simian malaria parasites (*P. knowlesi*, *P. cynomolgi*, *P. fieldi* and *P. inui*). These were in accordance to the respective macaques' nested PCR screening results, with the exception of macaque SG/EHIWM59/Y10 and SG/EHI/WM91/Y11. Nested PCR assay detected *P. coatneyi* in SG/EHI/WM59/Y10 and *P. knowlesi* in SG/EHI/WM91/Y11. However, transformants containing *csp* gene of these two simian *Plasmodium* species were not isolated from the respective samples. This could be due to the low parasite load of these parasite species in these macaques. As such, their detection was only revealed by the highly sensitive nested PCR assay.

Characterization of the *csp* gene not only confirmed the nested PCR assay's results, but also revealed the high genotypic diversity of the simian malaria parasites, especially *P. knowlesi* (Table 4.20). Four distinct subclades were observed within the *P. knowlesi* clade constructed with isolates from the macaques (Figure 4.4). There

Table 4.20: Comparison of the species of malaria parasites from the 15 wild macaques, identified by nested PCR assay and *csp* gene characterization

Monkey isolate	**Age**	**Malaria parasites identified by:**	
		Nested PCR	*Csp* gene sequencing (no. genotypes)
SG/EHI/WM01/Y07	Adult	Pk, Pfi	Pk (1), Pfi (1)
SG/EHI/WM02/Y07	Adult	Pk, Pcy	Pk (2), Pcy (1)
SG/EHI/WM04/Y09	Adult	Pk	Pk (6)
SG/EHI/WM05/Y09	Adult	Pk, Pcy, Pfi	Pk (2), Pcy (1), Pfi (1)
SG/EHI/WM11/Y09	Juvenile	Pk	Pk (1)
SG/EHI/WM15/Y09	Juvenile	Pk, Pfi, Pct	Pk (1), Pfi (1)
SG/EHI/WM16/Y09	Adult	Pk, Pcy	Pk (1), Pcy (2)
SG/EHI/WM17/Y09	Juvenile	Pk	Pk (3)
SG/EHI/WM18/Y09	Juvenile	Pcy, Pfi	Pcy (1), Pfi (1)
SG/EHI/WM26/Y09	Adult	Pk	Pk (6)
SG/EHI/WM33/Y09	Adult	Pk, Pcy	Pk (1), Pcy (1)
SG/EHI/WM35/Y09	Juvenile	Pk	Pk (2)
SG/EHI/WM42/Y10	Adult	Pcy	Pcy (1)
SG/EHI/WM44/Y10	Adult	Pcy, Pfi	Pcy (2), Pfi (1)
SG/EHI/WM91/Y11	Adult	Pk, Pin	Pin (2)

Pct, Pcy, Pfi, Pin and Pk denodes *Plasmodium coatneyi*, *P. cynomolgi*, *P. fieldi*, *P. inui* and *P. knowlesi*, respectively.

were six macaques infected with multiple genotypes of *P. knowlesi*. Of these, two had six genotypes; one had three genotypes, while the rest had two genotypes each. The high prevalence of *P. knowlesi*, together with the multiple genotypes and species infections in both adult and juvenile macaques, reflects a high malaria transmission intensity among these macaques in the forest. In a study conducted by Arez and co-workers in Antula, Republic of Guinea Bissau; an area with intense falciparium transmission, the mosquito vector *An. gambiae* was found to acquire one to two genotypes of *P. falciparium*, while multiple-genotype infections were common in the human population. The asynchronous gametocyte production of different genotypes of malaria parasites resulted in the low number of genotypes in the vector hosts, even though the vertebrate hosts were infected with more than one genotypes of the parasite. Hence, for a vertebrate host to acquire multiple genotypes of the parasites, the host will have to receive multiple infective bites from several individual mosquitoes harbouring different genotypes. Therefore, the presence of multiple species and genotypes infection in the vertebrate hosts indicates the presence of high malaria transmission intensity [124].

As discussed in previous chapter, the high malaria infection rate (80%) observed among juvenile macaques might be a result of an intense malaria transmission in the area, or the occurrence of vertical transmission of simian malaria parasites. Cases of congenital malaria had been reported for *P. vivax* and *P. falciparium* [125], although it is not yet known if the five simian malaria parasites can cross the placental barrier. Nonetheless, seven out of 20 of the malaria-positive juvenile macaques were infected with more than one species of *Plasmodium* parasites (Table 3.1). In addition, two of the knowlesi-infected juvenile macaques which were randomly selected for the

Plasmodium csp gene characterization were found to harbour multiple *P. knowlesi* genotypes (Table 4.20). These observations too suggest the presence of a high malaria transmission intensity in the restricted-access forest.

In contrast to the high genotypic diversity of parasites detected in the wild macaques, all the locally-acquired human knowlesi cases were each infected with single knowlesi genotype. This might be due to a lower exposure to mosquito bites for humans as compared to the simian hosts, indicating that the vector responsible for the simian malaria transmission in Singapore may be highly simiophagic (attracted to monkey hosts for blood meal). In an entomological surveillance conducted by the Singapore military in the affected forest, at least 6 species of anopheline mosquitoes were caught through human landing catch. These include *An. barbirostris* sp. group, *An. sinensis*, *An. tesselatus*, *An. sundaicus*, *An. lesteri*, and *An. kochi* [20]. Majority of these mosquitoes caught were known to be anthropohilic [126] . Of the three thousand mosquitoes screened by PCR, none was tested positive for malaria parasites (EHI unpublished data), suggesting that these anopheline mosquitoes caught may not be involved in the sylvatic transmission cycle of simian malaria parasites. Humans are their accidental hosts and the risk of acquiring knowlesi infection is probably only highest when they enter the forest. This could explain the low number of human knowlesi cases reported in Singapore despite thousands of military personnel accessing the affected forest area.

The genotypes found in the four local human cases were shared with five of the macaques, some of which were caught at different time points. The identical *csp* sequences between macaques and human cases suggested a strong molecular

epidemiological linkage between the two. SG/EHI/H01/Y07 and SG/EHI/H02/Y07, the first two human knowlesi cases reported in 2007, had identical *csp* gene sequences with *P. knowlesi* isolates from SG/EHI/WM02/Y07. Similarly, the third human knowlesi case in 2007 (SG/EHI/H07/Y07) had shared *csp* gene sequences with the two macaques caught in the same year (SG/EHI/WM01/Y07 and SG/EHI/WM02/Y07). Both macaques were trapped in the restricted access forest upon the notification of the human knowlesi cases. These suggested that the wild long-tailed macaques were reservoir hosts of *P. knowlesi* and the human cases might have acquired the infection in the same vicinity where the infected macaques were found. Interestingly, identical *P. knowlesi csp* gene sequences in two of the human cases (SG/EHI/H01/Y07 and SG/EHI/H02/Y07) were also detected in macaques sampled two years later (SG/EHI/WM11/Y09, SG/EHI/WM33/Y09, SG/EHI/WM35/Y09). The detection of identical *P. knowlesi csp* gene sequences in macaques caught across different years may suggest an ongoing sylvatic transmission of simian malaria parasites among the wild macaques.

Unlike the locally-acquired knowlesi cases, the two imported human knowlesi cases (SG/EHI/H1-im/Y09 and SG/EHI/H2-im/Y09), had strikingly higher number of *P. knowlesi* genotypes. These two patients had contracted the infection after visiting Pahang, a state in peninsular Malaysia with reports of *P. knowlesi* trasmission [24]. Each patient sample had at least five distinct genotypes, indicating that these patients might have acquired their infection from an area with intense *P. knowlesi* transmission. This conclusion is drawn with reference to the observations made by Arez and co-workers; the high genetic diversity of the parasites in humans is likely a result of superinfection in an area with intense malaria transmission [124]. As the *csp*

gene sequences of *P. knowlesi* isolates from these two cases were distinct from the *P. knowlesi* isolates of our local cases and macaques, *csp* gene characterization may be potentially useful in the differentiatiation of imported *P. knowlesi* cases from the locally-acquired infections.

Sequence alignment of the *csp* non-repeat region of the *P. inui* isolates from SG/EHI/WM91/Y11 revealed two sub-variants of *P. inui*, as demonstrated by the high pair-wise divergence rate of 7.6% and amino acid polymorphisms at the conserved region I and region II-plus (Table 4.18 and 4.19). One of the sub-variant was similar to the isolate from CDC while the other was closer to the *P. inui* isolate from Kapit, Sarawak. The high pair-wise divergence rate may suggest the presence of a sub-variant of *P. inui* or possibly a novel species of malaria parasite that is closely related to *P. inui.* However, a second gene needs to be sequenced to lend further support to this finding.

In conclusion, analyses based on the *csp* genes of *Plasmodium* isolates from the human and macaque samples revealed that knowlesi malaria is a zoonosis in Singapore and the wild macaques are also reservoir hosts for a panel of simian malaria parasites. The high diversity of simian malaria parasites found in local macaques strongly suggests an intense and continuous sylvatic malaria transmission among the wild macaques. Unfortunately, the vector responsible for this transmission has yet to be identified.

CHAPTER FIVE

SUMMARY AND INDICATIONS FOR FUTURE WORK

5.1 Summary

Natural infection of simian malaria in man was once considered to be rare and of no public health significance, until the discovery of a large focus of human knowlesi cases in Kapit Division of Sarawak, Malaysian Borneo in 2004 [45]. With the knowledge of this zoonotic transmission and its clinical manifestations, coupled with the advent of molecular assays for its detection, human knowlesi cases were first reported in Singapore in 2007 [46]. This prompted an epidemiological investigation that led to the identification of long-tailed macaques as the reservoir of infection. Furthermore, only macaques found in the restricted forest were infected with *P. knowlesi*. This suggests that transmission of *P. knowlesi* might be occurring in the forest and not in urbanized areas [20]. Aside from *P. knowlesi*, long-tailed macaques are also natural host to other species of malaria parasites, of which *P. cynomolgi* and *P. inui* are potentially infectious to humans [7]. In order to evaluate the risk of zoonotic malaria transmission in Singapore, this study was undertaken to determine the prevalence and genetic diversity of simian malaria parasites in local macaques.

Microscopic observation of the giemsa-stained blood film for the detection and identification of simian malaria parasites can be difficult and inaccurate due to the overlapping morphological characteristics between different simian *Plasmodium* species. Morphological identification is further complicated in macaques hosts with mixed infection and low parasitemia [1, 7]. Therefore, a sensitive screening assay for

Plasmodium parasites and a species-specific nested PCR assay to identify the five different species of simian malaria parasite were developed.

The *Plasmodium*-genus specific PCR assay developed in this study was highly specific and sensitive for malaria parasites. Its single amplification reaction reduces the chance of PCR product cross contamination, a common problem associated with the nested PCR assay. In addition, its short run-time and compatibility for use in both conventional and real-time PCR format makes it useful for the screening of malaria parasites in large number of samples, and allow laboratories with different resources to have comparable results. The simian malaria-species specific assay designed in this study was found to be specific to the five simian malaria species, making it a valuable tool for the prevalence study of simian malaria parasites in macaques. As the published knowlesi-specific PCR primers have shown random cross-reaction with *P. vivax* [90], it has limited use in areas where these two parasites co-exist. Therefore, the simian malaria-species specific assay designed in this study can also aid in the confirmation of zoonotic transmission of *P. knowlesi*, and possibly *P. cynomolgi* and *P. inui* in humans.

A total of 157 local long-tailed macaques were screened using the PCR assays developed in this study. Of these, 92 macaques were caught in the restricted-access forests, while 65 were caught from areas near to human habitations. Among the macaques caught in the forest, 71.7% were found to harbour malaria parasites while none of the peri-domestic macaques were infected. The difference in the infection rate among macaques caught from different geographical locations could be due to the lack of competent vectors for simian malaria transmission in the urban areas. Using

the simian malaria species-specific nested PCR assay, all five simian *Plasmodium* species (namely *P. knowlesi, P. coatneyi, P. cynomolgi, P. fieldi* and *P. inui*), which long-tailed macaques are known natural host, were identified. *Plasmodium knowlesi* (68.2%) was the most predominant malaria parasites found, followed by *P. cynomolgi* (60.6%), *P. fieldi* (16.7%), *P. coatneyi* (3.0%) and *P. inui* (1.5%). All of these malaria parasites, except *P. inui*, were found in wild macaques caught from mainland Singapore, while *P. knowlesi* and *P. inui* were detected in the sole malaria-positive macaque collected from the forest of an offshore military island.

The high infection rate of macaques, particularly among the juvenile macaques, might be a result of an intense malaria transmission in the restricted forest. Although we cannot negate the possibility of a vertical transmission of simian malaria parasites from the infected mothers to the juveniles, the co-infection of multiple *Plasmodium* species and genotypes among the infected macaques is also a reflection of an intense malaria transmission in the forest. In addition, the sylvatic transmission of simian malaria parasites among local macaques must have been ongoing in Singapore as *P. knowlesi* was first discovered in a Singapore macaque in 1931 [14, 87]. The detection of identical *P. knowlesi csp* gene sequences from macaques trapped across different years (2007-2009) further accentuates this observation.

Molecular characterization and phylogentic analyses of the *Plasmodium csp* gene obtained from Singapore isolates revealed that *P. knowlesi* found in local human cases were identical to those found in some macaques. This clearly illustrates that four of the human knowlesi cases were acquired locally and the transmission was zoonotic. In contrast to the majority of the macaques which were infected with multiple *P.*

knowlesi genotypes, only single genotype was found in local human cases. This may reflect a high intensity of transmission in the forest and the human cases had acquired the infection when they entered these forests. On the contrary, the two human cases which contracted *P. knowlesi* infection after visiting Pahang, peninsular Malaysia, harboured multiple knowlesi genotypes. These genotypes were distinct when compared to the local isolates. These findings, together with other epidemiological data, strongly confirmed that these *P. knowlesi* infections were not acquired locally.

5.2 Indications for future research

Despite the high intensity and ongoing transmission of malaria parasites among the wild macaques, the vectors involved in the sylvatic transmission of zoonotic knowlesi in Singapore have yet to be identified. To date, only mosquitoes from the *Anopheles leucosphyrus* group have been incriminated as vectors of *P. knowlesi* and other simian malaria parasites. These include *An. hackeri* [127] and *An. cracens* [24] in peninsular Malaysia, *An. latens* in Sarawak, East Malaysia [128, 129] and *An. dirus* in Vietnam [130, 131]. Although Singapore lies within the distribution limit of the *An. leucosphyrus* group [12], there has been no records of the presence of these species group of mosquitoes [132]. Entomological surveillance conducted by the Singapore military in the affected forest also did not reveal the presence of the *An. leucosphyrus* group of mosquitoes. Instead, at least six species of human-biting anopheline species were found, but none of these mosquitoes collected was positive for malaria parasites.

The vector for simian malaria transmission in Singapore is postulated to be highly attracted to monkey hosts instead of human, as inferred from the high intensity of transmission among the macaques and a contrasting low sporadic human knowlesi

incidence. Hence, future study that aims to elucidate the vector may require the use of monkey-baited traps [24, 128].

The identity of the vector is important in ascertaining and mitigating the risk of zoonotic knowlesi transmission so that targeted vector control operations can be implemented. The efficiency of any vector in transmitting malaria in any given geographical area is largely due to their bionomics, which are major elements in determining the appropriateness of control measures to be initiated [133]. Since the notification of the first human knowlesi case, the Singapore Armed Forces has put into operation general mosquito control measures such as environmental management, insecticide-treated uniforms and the use of *Bacilus thuringiensis* var. Israelensis [20]. These measures were implemented from 2007 to current, and have resulted in a reduction of total mosquito population from 64.1 mosquitoes per sampling site in August 2007 to 4.3 per site by June 2011 (Patrick Lam, personal communication, 14th November 2011). Despite the drop in mosquito population, the infection rate of wild macaques, particularly the juvenile macaques, remains high. Majority of these juvenile macaques (age three years and below) were sampled from 2009 to 2011. The high infection rate, along with the multiple species and genotypes infection among these juveniles, suggests that the current mosquito control measures may not be effective against the simiophagic vector responsible for simian malaria transmission among the macaques. The knowledge of the vector identity and its bionomics will allow the design of a targeted vector control strategy. Along with the available epidemiological and molecular data obtained from the human cases and macaque hosts, the vector identity and its associated entomological data will also provide a full epidemiological picture of *P. knowlesi* transmission in Singapore.

Apart from determining the vector responsible for the transmission, it may also be important to monitor the movement and home range of the wild macaques through global positioning system (GPS) tracking. This will determine the spatial distribution of these macaques, and consequently aid in the identification of risk areas for *P. knowlesi* transmission. Although peri-domesticated macaques in this study were found to be negative for malaria parasite, it is possible that wild macaques from the restricted-access forest may migrate to areas near human habitations due to deforestation and habitat destruction. On the other hand, there is also a possibility of simiophagic mosquito vector, migrating out with the monkeys to areas close to human habitations. This future changes in land use and exploitation of the forest may result in greater contact between monkeys, mosquitoes and humans, increasing the risk of zonnotic transmission of simian malaria parasites to the general population. Therefore, the elucidation of the vectors, together with the distribution and the dispersal of the wild macaques in the restricted forest area should be investigated for the identification of risk areas for *P. knowlesi* transmission.

5.3 Conclusion

This study has provided evidence to illustrate the presence of an intense and ongoing sylvatic transmission of malaria parasites among local macaques. However, this should not be a cause for alarm as the risk of the general population acquiring zoonotic malaria should be low, due to the absence of malaria parasite in peri-domestic macaques. Moreover, all reported local cases thus far were associated with occupation or travel history. However, the risk of acquiring simian malaria in Singapore can be better demonstrated with information on the spatial distribution of

macaques and the identification of vectors involved in the transmission among macaques and between macaques and humans.

References

1. Garnham, P.C.C., *Malaria parasites and other haemosporidia*. 1966: Blackwell Scientific Publications, Oxford, United Kingdom.

2. WHO, *World Malaria Report*. 2010.

3. WHO. *Malaria: Disease burden in SEA Region*. 2011 27 September 2011; Available from: http://www.searo.who.int/en/Section10/Section21/Section340_4018.htm.

4. Warrell, D.A. and H.M. Gilles, *Essential Malariology*. Fourth edition ed. 2002: Arnold

5. WHO. *Malaria, countries or areas at risk of transmission, 2010*. 2011 04/10/2011]; Available from: http://gamapserver.who.int/mapLibrary/Files/Maps/Global_Malaria_2010.png.

6. WHO. *Malaria: What is Malaria?* 2007 09 January 2007; Available from: http://www.searo.who.int/EN/Section10/Section21/Section334.htm.

7. Coatney, G.R., Collins, W.E., Warren, M. & Contacos, P.G. , *The Primate Malarias.* 1971: U.S. Government Printing Office: Washington D.C.

8. Collins, W.E. and P.G. Contacos, *Observations on the Relapse Activity of Plasmodium fieldi in the Rhesus Monkey.* Journal of Parasitology, 1971. **57**(1): p. 29-32.

9. Cogswell, F.B. *Malaria and piroplasm of non-human primates. In: Companion and Exotic Animal Parasitology (Bowman, D.D. ed.)* 2000; Available from: http://www.ivis.org/advances/Parasit_Bowman/cogswell_primate/ivis.pdf.

10. Cogswell, F.B., *The hypnozoite and relapse in primate malaria.* Clin Microbiol Rev, 1992. **5**(1): p. 26-35.

11. Wirth, D.F., *Biological revelations.* Nature, 2002. **419**(6906): p. 495-6.

12. Fooden, J., *Malaria in macaques.* International Journal of Primatology, 1994. **15** (4): p. 573-596.

13. Blacklock, B. and S. Adler, *A parasite resembling Plasmodium falciparum in a chimpanzee.* Annals of Tropical Medicine and Parasitology, 1922. **16**: p. 99-106.

14. Knowles, R. and B.M.D. Gupta, *A study of monkey-malaria and its experimental transmission to man.* Indian Medical Gazette, 1932. **67**: p. 301-320.

15. C. E.van Rooyen and G.R. Pile, *Observations on infection by Plasmodium knowlesi (ape malaria) in the treatment of general paralysis of the insane* The British Medical Journal, 1935. **2**(3901): p. 662-666.

16. Das-Gupta, B.M., *Transmission of P. inui to man.* Proceedings of the National Institute of Science, India, 1938. **4**: p. 241-244.

17. Gysin, J., *Animal models: Primates*, in *Malaria: Parasite Biology, Pathogenesis and Protection*, I.W. Sherman, Editor. 1998, American Society for Microbiology Washington D.C. p. 419-441.

18. Kuntz, R.E. and B.J. Myers, *A Checklist of Parasites and Commensals Reported for the Taiwan Macaque.* PRIMATES, 1969. **10**: p. 71-80.

19. Huang, C.C., D.D. Ji, Y.C. Chiang, H.J. Teng, H.J. Liu, C.D. Chang, and Y.H. Wu, *Prevalence and molecular characterization of Plasmodium inui among Formosan macaques (Macaca cyclopis) in Taiwan.* J Parasitol, 2009. **96**(1): p. 8-15.

20. Jeslyn, W.P., T.C. Huat, L. Vernon, L.M. Irene, L.K. Sung, L.P. Jarrod, B. Singh, and N.L. Ching, *Molecular Epidemiological Investigation of Plasmodium knowlesi in Humans and Macaques in Singapore.* Vector Borne Zoonotic Dis, 2010.

21. Eyles, D.E., F.L. Dunn, M. Warren, and E. Guinn, *Plasmodium coatneyi from the Philippines.* J Parasitol, 1963. **49**: p. 1038.

22. Eyles, D.E., *The species of simian malaria: Taxonomy, morphology, life cycle and geographical distribution of the monkey species.* Journal of Parasitolology, 1963. **49**: p. 866-87.

23. Tsukamoto, M., A. Miyata, and I. Miyagi, *Surveys on simian malaria parasites and their vector in Palawan Island, the Philippines.* Tropical Medicine, 1978. **20**(1): p. 39-50.

24. Vythilingam, I., Y.M. Noorazian, T.C. Huat, A.I. Jiram, Y.M. Yusri, A.H. Azahari, I. Norparina, A. Noorrain, and S. Lokmanhakim, *Plasmodium knowlesi in humans, macaques and mosquitoes in peninsular Malaysia.* Parasit Vectors, 2008. **1**(1): p. 26.

25. Lee, K.S., P.C. Divis, S.K. Zakaria, A. Matusop, R.A. Julin, D.J. Conway, J. Cox-Singh, and B. Singh, *Plasmodium knowlesi: Reservoir Hosts and Tracking the Emergence in Humans and Macaques.* PLoS Pathog, 2011. **7**(4): p. e1002015.

26. Eyles, D.E., A.B. Laing, and Y.L. Fong, *Plasmodium fieldi sp. nov., a new species of malaria parasite from the pigtailed macaque in Malaya.* Ann Trop Med Parasitol, 1962. **56**: p. 242-7.

27. Sandosham, A.A., D.E. Eyles, and L.F. Yap, *Plasmodium cynomolgi bastianellii from a Malayan monkey and its similarity to aberrant P. vivax in man. The Medical Journal of Malaya*, 1962. **17**: p. 78-79.

28. Seethamchai, S., C. Putaporntip, S. Malaivijitnond, L. Cui, and S. Jongwutiwes, *Malaria and Hepatocystis species in wild macaques, southern Thailand.* Am J Trop Med Hyg, 2008. **78**(4): p. 646-53.

29. Sallum, M.A.M., E.L. Peyton, B.A. Harrison, and R.C. Wilkerson, *Revision of the Leucosphyrus group of Anopheles (Cellia) (Diptera, Culicidae).* Revista Brasileira de Entomologia, 2005. **49**: p. 01-152.

30. Clark, H.C. and L.H. Dunn, *Experimental efforts to transfer monkey malaria to man.* Am J Trop Med Hyg, 1931. **11**: p. 1-7.

31. Sinton, J.A., E.L. Hutton, and P.G. Shute, *Failure to transmit an infection of P. cynomolgi to man by blood injection and by mosquito bites.* Journal of Tropical Medicine and Hygiene, 1938. **41**: p. 245-246.

32. *The Eighth World Health Assembly.* Public Health Rep. , 1955. **70**(11): p. 1057-1060.

33. Eyles, D., G.R. Coatney, and M.E. Getz, *Vivax-type malaria parasite of macaques transmissible to man.* Science, 1960. **13**: p. 1812-1813.

34. Schmidt, L.H., R. Greenland, and C.S. Genther, *The transmission of Plasmodium cynomolgi to man.* Am J Trop Med Hyg, 1961. **10**: p. 679-88.

35. Chin, W., P.G. Contacos, W.E. Collins, M.H. Jeter, and E. Alpert, *Experimental mosquito-transmission of Plasmodium knowlesi to man and monkey.* Am J Trop Med Hyg, 1968. **17**(3): p. 355-8.

36. Coatney, G.R., W. Chin, P.G. Contacos, and H.K. King, *Plasmodium inui, a quartan-type malaria parasite of old world monkeys transmissible to man.* Journal of Parasitology, 1966. **52**: p. 660-663.

37. Coatney, G.R., H.A. Elder, P.G. Contacos, M.E. Getz, R. Greenland, R.N. Rossan, and L.H. Schmidt, *Transmission of the M strain of Plasmodium cynomolgi to man.* Am J Trop Med Hyg, 1961. **10**: p. 673-8.

38. Contacos, P.G., G.R. Coatney, T.C. Orihel, W.E. Collins, W. Chin, and M.H. Jeter, *Transmission of Plasmodium schwetzi from the chimpanzee to man by mosquito bite.* Am J Trop Med Hyg, 1970. **19**(2): p. 190-5.

39. Contacos, P.G., H.A. Elder, and G.B. Coatney, *Man to man transfer of two strains of Plasmodium cynomolgi by mosquito bite.* Am J Trop Med Hyg, 1962. **11**: p. 186-193.

40. Contacos, P.G., J.S. Lunn, G.R. Coatney, J.W. Kilpatrick, and F.E. Jones, *Quartan-type malaria parasite of the New World monkeys transmissible to man.* Science, 1963. **142**: p. 676.

41. Deane, L.M., M.P. Deane, and J.F. Neto, *Studies on transmission of simian malaria and on a natural infection of man with Plasmodium simium in Brazil.* Bulletin of the World Health Organization, 1966. **35**: p. 805-808.

42. Warren, M., G.F. Bennett, and S. A.A, *A new malaria parasite from the white-handed gibbon, Hylobates lar lar in Malaya.* Singapore Medical Journal, 1965. **6**: p. 50.

43. Chin, W., P.G. Contacos, G.R. Coatney, and H.R. Kimball, *A Naturally Acquired Quotidian-Type Malaria in Man Transferable to Monkeys.* Science, 1965. **149**(3686): p. 865-.

44. Fong, Y.L., F.C. Cadigan, and G.R. Coatney, *A presumptive case of naturally occurring Plasmodium knowlesi malaria in man in Malaysia.* Trans R Soc Trop Med Hyg, 1971. **65**(6): p. 839-40.

45. Singh, B., L. Kim Sung, A. Matusop, A. Radhakrishnan, S.S. Shamsul, J. Cox-Singh, A. Thomas, and D.J. Conway, *A large focus of naturally acquired Plasmodium knowlesi infections in human beings.* Lancet, 2004. **363**(9414): p. 1017-24.

46. Ng, O.T., E.E. Ooi, C.C. Lee, P.J. Lee, L.C. Ng, S.W. Pei, T.M. Tu, J.P. Loh, and Y.S. Leo, *Naturally acquired human Plasmodium knowlesi infection, Singapore.* Emerg Infect Dis, 2008. **14**(5): p. 814-6.

47. Figtree, M., R. Lee, L. Bain, T. Kennedy, S. Mackertich, M. Urban, Q. Cheng, and B.J. Hudson, *Plasmodium knowlesi in human, Indonesian Borneo.* Emerg Infect Dis, 2010. **16**(4): p. 672-4.

48. Sulistyaningsih, E., L.E. Fitri, T. Loscher, and N. Berens-Riha, *Diagnostic difficulties with Plasmodium knowlesi infection in humans.* Emerg Infect Dis, 2010. **16**(6): p. 1033-4.

49. Joveen-Neoh, W.F., K.L. Chong, C.M. Wong, and T.Y. Lau, *Incidence of malaria in the interior division of sabah, malaysian borneo, based on nested PCR.* J Parasitol Res, 2011. **2011**: p. 104284.

50. Luchavez, J., F. Espino, P. Curameng, R. Espina, D. Bell, P. Chiodini, D. Nolder, C. Sutherland, K.S. Lee, and B. Singh, *Human Infections with Plasmodium knowlesi, the Philippines.* Emerg Infect Dis, 2008. **14**(5): p. 811-3.

51. Jongwutiwes, S., C. Putaporntip, T. Iwasaki, T. Sata, and H. Kanbara, *Naturally acquired Plasmodium knowlesi malaria in human, Thailand.* Emerg Infect Dis, 2004. **10**(12): p. 2211-3.

52. Putaporntip, C., T. Hongsrimuang, S. Seethamchai, T. Kobasa, K. Limkittikul, L. Cui, and S. Jongwutiwes, *Differential prevalence of Plasmodium infections and cryptic Plasmodium knowlesi malaria in humans in Thailand.* J Infect Dis, 2009. **199**(8): p. 1143-50.

53. Jongwutiwes, S., P. Buppan, R. Kosuvin, S. Seethamchai, U. Pattanawong, J. Sirichaisinthop, and C. Putaporntip, *Plasmodium knowlesi Malaria in Humans and Macaques, Thailand.* Emerg Infect Dis, 2011. **17**(10): p. 1799-806.

54. Jiang, N., Q. Chang, X. Sun, H. Lu, J. Yin, Z. Zhang, M. Wahlgren, and Q. Chen, *Co-infections with Plasmodium knowlesi and other malaria parasites, Myanmar.* Emerg Infect Dis, 2010. **16**(9): p. 1476-8.

55. Zhu, H.M., J. Li, and H. Zheng, *[Human natural infection of Plasmodium knowlesi].* Zhongguo Ji Sheng Chong Xue Yu Ji Sheng Chong Bing Za Zhi, 2006. **24**(1): p. 70-1.

56. Van den Eede, P., H.N. Van, C. Van Overmeir, I. Vythilingam, T.N. Duc, X. Hung le, H.N. Manh, J. Anne, U. D'Alessandro, and A. Erhart, *Human Plasmodium knowlesi infections in young children in central Vietnam.* Malar J, 2009. **8**: p. 249.

57. Van den Eede, P., I. Vythilingam, D.T. Ngo, V.H. Nguyen, X.H. Le, U. D'Alessandro, and A. Erhart, *Plasmodium knowlesi malaria in Vietnam: some clarifications.* Malar J, 2010. **9**: p. 20.

58. Khim, N., S. Siv, S. Kim, T. Mueller, E. Fleischmann, B. Singh, P.C. Divis, N. Steenkeste, L. Duval, C. Bouchier, S. Duong, F. Ariey, and D. Menard, *Plasmodium knowlesi Infection in Humans, Cambodia, 2007-2010.* Emerg Infect Dis, 2011. **17**(10): p. 1900-2.

59. Bronner, U., P.C. Divis, A. Farnert, and B. Singh, *Swedish traveller with Plasmodium knowlesi malaria after visiting Malaysian Borneo.* Malar J, 2009. **8**: p. 15.

60. Kantele, A., H. Marti, I. Felger, D. Muller, and T.S. Jokiranta, *Monkey malaria in a European traveler returning from Malaysia.* Emerg Infect Dis, 2008. **14**(9): p. 1434-6.

61. Ta, T.T., A. Salas, M. Ali-Tammam, C. Martinez Mdel, M. Lanza, E. Arroyo, and J.M. Rubio, *First case of detection of Plasmodium knowlesi in Spain by Real Time PCR in a traveller from Southeast Asia.* Malar J, 2010. **9**: p. 219.

62. Hoosen, A. and M.T. Shaw, *Plasmodium knowlesi in a traveller returning to New Zealand.* Travel Med Infect Dis, 2011.

63. *Simian Malaria in a U.S. Traveller - New York, 2008.* JAMA: The Journal of the American Medical Association, 2009. **301**(18): p. 1871-1872.

64. Cox-Singh, J., J. Hiu, S.B. Lucas, P.C. Divis, M. Zulkarnaen, P. Chandran, K.T. Wong, P. Adem, S.R. Zaki, B. Singh, and S. Krishna, *Severe malaria - a case of fatal Plasmodium knowlesi infection with post-mortem findings: a case report.* Malar J, 2010. **9**: p. 10.

65. Galinski, M.R. and J.W. Barnwell, *Monkey malaria kills four humans.* Trends Parasitol, 2009. **25**(5): p. 200-4.

66. Collins, W.E. and J.W. Barnwell, *Plasmodium knowlesi: finally being recognized.* J Infect Dis, 2009. **199**(8): p. 1107-8.

67. Cox-Singh, J. and B. Singh, *Knowlesi malaria: newly emergent and of public health importance?* Trends Parasitol, 2008. **24**(9): p. 406-10.

68. Sabbatani, S., S. Fiorino, and R. Manfredi, *The emerging of the fifth malaria parasite (Plasmodium knowlesi): a public health concern?* Braz J Infect Dis, 2010. **14**(3): p. 299-309.

69. White, N.J., *Plasmodium knowlesi: the fifth human malaria parasite.* Clin Infect Dis, 2008. **46**(2): p. 172-3.

70. Yegneswaran, B., D. Alcid, and J. Mohan, *Plasmodium knowlesi: an important yet overlooked human malaria parasite.* Mayo Clin Proc, 2009. **84**(7): p. 664; author reply 664.

71. Ameri, M., *Laboratory diagnosis of malaria in nonhuman primates.* Vet Clin Pathol, 2010. **39**(1): p. 5-19.

72. Rain, A.N., J.W. Mak, and R. Zamri, *Simian malaria infection in wild caught Macaca fascicularis and Presbytis spp in Malaysia.* Southeast Asian J Trop Med Public Health, 1993. **24**(2): p. 386-7.

73. Snounou G., S.V., X.P. Zhu, W. Jarra, L. Pinheiro, V.E. do Rosario, *High sensitivity of detection of human malaria parasites by the use of nested polymerase chain reaction.* Molecular and Biochemical Parasitology, 1993b. **61**(315-320).

74. Singh, B., A. Bobogare, J. Cox-Singh, G. Snounou, M.S. Abdullah, and H.A. Rahman, *A genus- and species-specific nested polymerase chain reaction malaria detection assay for epidemiologic studies.* Am J Trop Med Hyg, 1999. **60**(4): p. 687-92.

75. Arez, A.P., D. Lopes, J. Pinto, A.S. Franco, G. Snounou, and V.E. do Rosario, *Plasmodium sp.: optimal protocols for PCR detection of low parasite numbers from mosquito (Anopheles sp.) samples.* Exp Parasitol, 2000. **94**(4): p. 269-72.

76. Pinto, J., A.P. Arez, S. Franco, V.E. do Rosario, K. Palsson, T.G. Jaenson, and G. Snounou, *Simplified methodology for PCR investigation of midguts from mosquitoes of the Anopheles gambiae complex, in which the vector and Plasmodium species can both be identified.* Ann Trop Med Parasitol, 1997. **91**(2): p. 217-9.

77. Vythilingam, I., K. Nitiavathy, P. Yi, B. Bakotee, B. Hugo, B. Singh, R.A. Wirtz, and K. Palmer, *A highly sensitive, nested polymerase chain reaction based method using simple dna extraction to detect malaria sporozoites in mosquitos.* Southeast Asian J Trop Med Public Health, 1999. **30**(4): p. 631-5.

78. Cox-Singh, J., S. Mahayet, M.S. Abdullah, and B. Singh, *Increased sensitivity of malaria detection by nested polymerase chain reaction using simple sampling and DNA extraction.* Int J Parasitol, 1997. **27**(12): p. 1575-7.

79. Bottius, E., A. Guanzirolli, J.F. Trape, C. Rogier, L. Konate, and P. Druilhe, *Malaria: even more chronic in nature than previously thought; evidence for subpatent parasitaemia detectable by the polymerase chain reaction.* Trans R Soc Trop Med Hyg, 1996. **90**(1): p. 15-9.

80. Mixson-Hayden, T., N.W. Lucchi, and V. Udhayakumar, *Evaluation of three PCR-based diagnostic assays for detecting mixed Plasmodium infection.* BMC Res Notes. **3**: p. 88.

81. Sandosham, A.A. and M.N.b. Keling, *Animal parasites of animals which affect man in malaysia.* The Medical Journal of Malaya, 1967. **22**(1): p. 16-25.

82. Sandosham, A.A., *History of Malaria in Malaya*, in *Malariology with special reference to Malaya*. 1959, University of Malaya Press.

83. Goh, K.T., *Eradication of malaria from Singapore.* Singapore Med J, 1983. **24**(5): p. 255-68.

84. Lee, Y.C., C.S. Tang, L.W. Ang, H.K. Han, L. James, and K.T. Goh, *Epidemiological characteristics of imported and locally-acquired malaria in Singapore.* Ann Acad Med Singapore, 2009. **38**(10): p. 840-9.

85. Ng, L.C., K.S. Lee, C.H. Tan, P.L. Ooi, S.G. Lam-Phua, R. Lin, S.C. Pang, Y.L. Lai, S. Solhan, P.P. Chan, K.Y. Wong, S.T. Ho, and I. Vythilingam, *Entomologic and molecular investigation into Plasmodium vivax transmission in Singapore, 2009.* Malar J. **9**: p. 305.

86. Quek, D.L., Y. Zhang, M. La, P. Fernandez, and J.-M. Chavattee, *Laboratory test results on Plasmodium knowlesi in clinical malaria positive cases registered in 2009 in Singapore.* Epidemiological News Bulletin, 2010. **36**: p. 83-94.

87. Sinton, J.A. and H.W. Mulligan, *A critical review of the literature relating to the identification of the malarial parasites recorded from monkeys of the families circopithecidæ and colobidæ*. 1932, Calcutta: Thacker's Press & Directories.

88. Payne, D., *Use and limitations of light microscopy for diagnosing malaria at the primary health care level.* Bull World Health Organ, 1988. **66**(5): p. 621-6.

89. Coleman, R.E., J. Sattabongkot, S. Promstaporm, N. Maneechai, B. Tippayachai, A. Kengluecha, N. Rachapaew, G. Zollner, R.S. Miller, J.A. Vaughan, K. Thimasarn, and B. Khuntirat, *Comparison of PCR and microscopy for the detection of asymptomatic malaria in a Plasmodium falciparum/vivax endemic area in Thailand.* Malar J, 2006. **5**: p. 121.

90. Imwong, M., N. Tanomsing, S. Pukrittayakamee, N.P. Day, N.J. White, and G. Snounou, *Spurious amplification of a Plasmodium vivax small-subunit RNA gene by use of primers currently used to detect P. knowlesi.* J Clin Microbiol, 2009. **47**(12): p. 4173-5.

91. Cross, J.H., M.Y. Hsu-Kuo, and J.C. Lien, *Accidental human infection with Plasmodium cynomolgi bastianellii.* Southeast Asian J Trop Med Public Health, 1973. **4**(4): p. 481-3.

92. Bennett, G.F. and M. Warren, *Transmission of a new strain of Plasmodium cynomolgi to man.* J Parasitol, 1965. **51**: p. 79-80.

93. CDC. *Laboratory identification of Parasites of Public Health Concern*. 17/09/2010 23/03/2011]; Available from: http://www.dpd.cdc.gov/dpdx/hTML/Frames/DiagnosticProcedures/body_dp_blood_DNAextraction.htm.

94. Brown, T.A., *Gene Cloning and DNA analysis*. 4th ed. 2001: Blackwell Science Ltd: Oxford.

95. WHO, *Basic Malaria Microscopy. Part 1. Learner's Guide*. 2nd ed. 2010, Geneva: WHO.

96. McCutchan, T.F., J.C. Kissinger, M.G. Touray, M.J. Rogers, J. Li, M. Sullivan, E.M. Braga, A.U. Krettli, and L.H. Miller, *Comparison of circumsporozoite proteins from avian and mammalian malarias: biological and phylogenetic implications.* Proc Natl Acad Sci U S A, 1996. **93**(21): p. 11889-94.

97. Dorak, M.T. *Real-time PCR*. 2011 03/09/2011; Available from: http://dorakmt.tripod.com/genetics/realtime.html.

98. Roche *LightCycler (R) 480 Instrument Operature's Manual Software Version 1.5*.

99. McCutchan, T.F., J. Li, G.A. McConkey, M.J. Rogers, and A.P. Waters, *The cytoplasmic ribosomal RNAs of Plasmodium spp.* Parasitology Today, 1995. **11**(4): p. 134-138.

100. Rosner, A., L. Maslenin, and S. Spiegel, *The use of short and long PCR products for improved detection of prunus necrotic ringspot virus in woody plants.* J Virol Methods, 1997. **67**(2): p. 135-41.

101. Fuentes, A., S. Kalchik, L. Gettler, A. Kwiatt, M. Konecki, and L. Jones-Engel, *Characterizing human-macaque interactions in Singapore.* Am J Primatol, 2008. **70**(9): p. 879-83.

102. Sha, J.C., M.D. Gumert, B.P. Lee, L. Jones-Engel, S. Chan, and A. Fuentes, *Macaque-human interactions and the societal perceptions of macaques in Singapore.* Am J Primatol, 2009. **71**(10): p. 825-39.

103. Swindler, D.R., *Primate dentition: an introduction to the teeth of non-human primates*. 2002, New York: Cambridge University Press.

104. di Giovanni, L., A.H. Cochrane, and V. Enea, *On the evolutionary history of the circumsporozoite protein in plasmodia.* Exp Parasitol, 1990. **70**(4): p. 373-81.

105. Hughes, A.L., *Circumsporozoite protein genes of malaria parasites (Plasmodium spp.): evidence for positive selection on immunogenic regions.* Genetics, 1991. **127**(2): p. 345-53.

106. Tewari, R., R. Spaccapelo, F. Bistoni, A.A. Holder, and A. Crisanti, *Function of region I and II adhesive motifs of Plasmodium falciparum circumsporozoite protein in sporozoite motility and infectivity.* J Biol Chem, 2002. **277**(49): p. 47613-8.

107. Rathore, D., S.C. Hrstka, J.B. Sacci, Jr., P. De la Vega, R.J. Linhardt, S. Kumar, and T.F. McCutchan, *Molecular mechanism of host specificity in Plasmodium falciparum infection: role of circumsporozoite protein.* J Biol Chem, 2003. **278**(42): p. 40905-10.

108. Sinnis, P., P. Clavijo, D. Fenyo, B.T. Chait, C. Cerami, and V. Nussenzweig, *Structural and functional properties of region II-plus of the malaria circumsporozoite protein.* J Exp Med, 1994. **180**(1): p. 297-306.

109. Wang, Q., H. Fujioka, and V. Nussenzweig, *Exit of Plasmodium sporozoites from oocysts is an active process that involves the circumsporozoite protein.* PLoS Pathog, 2005. **1**(1): p. e9.

110. Vargas-Serrato, E., V. Corredor, and M.R. Galinski, *Phylogenetic analysis of CSP and MSP-9 gene sequences demonstrates the close relationship of Plasmodium coatneyi to Plasmodium knowlesi.* Infect Genet Evol, 2003. **3**(1): p. 67-73.

111. Jongwutiwes, S., K. Tanabe, M.K. Hughes, H. Kanbara, and A.L. Hughes, *Allelic variation in the circumsporozoite protein of Plasmodium falciparum from Thai field isolates.* Am J Trop Med Hyg, 1994. **51**(5): p. 659-68.

112. Arnot, D.E., M.J. Stewart, and J.W. Barnwell, *Antigenic diversity in Thai Plasmodium vivax circumsporozoite proteins.* Mol Biochem Parasitol, 1990. **43**(1): p. 147-9.

113. Qari, S.H., W.E. Collins, H.O. Lobel, F. Taylor, and A.A. Lal, *A study of polymorphism in the circumsporozoite protein of human malaria parasites.* Am J Trop Med Hyg, 1994. **50**(1): p. 45-51.

114. Ozaki, L.S., P. Svec, R.S. Nussenzweig, V. Nussenzweig, and G.N. Godson, *Structure of the plasmodium knowlesi gene coding for the circumsporozoite protein.* Cell, 1983. **34**(3): p. 815-22.

115. Sharma, S., P. Svec, G.H. Mitchell, and G.N. Godson, *Diversity of circumsporozoite antigen genes from two strains of the malarial parasite Plasmodium knowlesi.* Science, 1985. **229**(4715): p. 779-82.

116. Galinski, M.R., D.E. Arnot, A.H. Cochrane, J.W. Barnwell, R.S. Nussenzweig, and V. Enea, *The circumsporozoite gene of the Plasmodium cynomolgi complex.* Cell, 1987. **48**(2): p. 311-9.

117. Huang, C.C., Y.C. Chiang, D.D. Ji, H.J. Teng, M.H. Liao, C.D. Chang, and Y.H. Wu, *Molecular cloning and characterization of the circumsporozoite protein gene of Plasmodium inui isolated from Formosan macaques (Macaca cyclopis) in Taiwan.* J Parasitol. **96**(6): p. 1145-51.

118. Goldman, I.F., S.H. Qari, P.G. Millet, W.E. Collins, and A.A. Lal, *Circumsporozoite protein gene of Plasmodium simium, a Plasmodium vivax-like monkey malaria parasite.* Mol Biochem Parasitol, 1993. **57**(1): p. 177-80.

119. Qari, S.H., Y.P. Shi, M.M. Povoa, M.P. Alpers, P. Deloron, G.S. Murphy, S. Harjosuwarno, and A.A. Lal, *Global occurrence of Plasmodium vivax-like human malaria parasite.* J Infect Dis, 1993. **168**(6): p. 1485-9.

120. Lockyer, M.J., C.S. Davies, A. Suhrbier, and R.E. Sinden, *Nucleotide sequence of the Plasmodium berghei circumsporozoite protein gene from the ANKA clone 2.34L.* Nucleic Acids Res, 1990. **18**(2): p. 376.

121. Saitou, N. and M. Nei, *The neighbor-joining method: a new method for reconstructing phylogenetic trees.* Mol Biol Evol, 1987. **4**(4): p. 406-25.

122. Felsenstein, J., *Evolutionary trees from DNA sequences: a maximum likelihood approach.* J Mol Evol, 1981. **17**(6): p. 368-76.

123. Kumar, S., K. Tamura, and M. Nei, *MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment.* Brief Bioinform, 2004. **5**(2): p. 150-63.

124. Arez, A.P., J. Pinto, K. Palsson, G. Snounou, T.G. Jaenson, and V.E. do Rosario, *Transmission of mixed Plasmodium species and Plasmodium falciparum genotypes.* Am J Trop Med Hyg, 2003. **68**(2): p. 161-8.

125. Poespoprodjo, J.R., W. Fobia, E. Kenangalem, A. Hasanuddin, P. Sugiarto, E. Tjitra, N.M. Anstey, and R.N. Price, *Highly effective therapy for maternal malaria associated with a lower risk of vertical transmission.* J Infect Dis. **204**(10): p. 1613-9.

126. Reid, J.A., *Anopheline mosquitoes of Malaya and Borneo*. 1968: Govt. of Malaysia.

127. Wharton, R.H. and D.E. Eyles., *Anopheles hackeri, a vector of Plasmodium knowlesi in Malaya.* Science, 1961. **134**(279-280).

128. Tan, C.H., I. Vythilingam, A. Matusop, S.T. Chan, and B. Singh, *Bionomics of Anopheles latens in Kapit, Sarawak, Malaysian Borneo in relation to the transmission of zoonotic simian malaria parasite Plasmodium knowlesi.* Malar J, 2008. **7**: p. 52.

129. Vythilingam, I., C.H. Tan, M. Asmad, S.T. Chan, K.S. Lee, and B. Singh, *Natural transmission of Plasmodium knowlesi to humans by Anopheles latens in Sarawak, Malaysia.* Trans R Soc Trop Med Hyg, 2006. **100**(11): p. 1087-8.

130. Nakazawa, S., R.P. Marchand, N.T. Quang, R. Culleton, N.D. Manh, and Y. Maeno, *Anopheles dirus co-infection with human and monkey malaria parasites in Vietnam.* Int J Parasitol, 2009. **39**(14): p. 1533-7.

131. Marchand, R.P., R. Culleton, Y. Maeno, N.T. Quang, and S. Nakazawa, *Co-infections of Plasmodium knowlesi, P. falciparum, and P. vivax among Humans and Anopheles dirus Mosquitoes, Southern Vietnam.* Emerg Infect Dis, 2011. **17**(7): p. 1232-9.

132. Apiwathnasorn, C., S.-T.N.C.o.T. Museum, and R. Centre, *A list of mosquito species in Southeast Asia*. 1986: Museum and Reference Centre, SEAMED-TROPMED National Centre of Thailand, Faculty of Tropical Medicine, Mahidol University.

133. Trung, H.D., W.V. Bortel, T. Sochantha, K. Keokenchanh, O.J. Briet, and M. Coosemans, *Behavioural heterogeneity of Anopheles species in ecologically different localities in Southeast Asia: a challenge for vector control.* Trop Med Int Health, 2005. **10**(3): p. 251-62.

Appendix A: List of simian *Plasmodium* species controls and their source

Species	Source	Date of bloodspot
Plasmodium coatneyi	Isolated from *Anopheles hackeri*	22/02/2010
Plasmodium cynomolgi B	Propagated in rhesus monkey	12/04/2001
Plasmodium fieldi	Isolated from *Anopheles hackeri* followed by Rhesus monkey	24/04/2006
Plasmodium inui	Isolated from leaf monkey	23/03/2004

Appendix B: Binding sites of primers for simian *Plasmodium* species-specific PCR

Appendix B continued

Appendix C: Details of peri-domestic long-tailed macaques

Code	Location	Gender	Age	Date collected
SG/EHI/PM01/Y08		Male	Adult	January 2008
SG/EHI/PM02/Y08		Male	Adult	January 2008
SG/EHI/PM03/Y08		Male	Adult	January 2008
SG/EHI/PM04/Y08		Male	Adult	January 2008
SG/EHI/PM05/Y08	Central Nature Reserve	Male	Adult	January 2008
SG/EHI/PM06/Y08		Male	Adult	January 2008
SG/EHI/PM07/Y08		Male	Adult	January 2008
SG/EHI/PM08/Y08		Male	Adult	January 2008
SG/EHI/PM09/Y08		Male	Adult	January 2008
SG/EHI/PM10/Y08		Male	Adult	January 2008
SG/EHI/PM11/Y10	Mandai Lake Road	Male	Adult	23/02/2010
SG/EHI/PM12/Y10	Mandai Lake Road	Male	Adult	23/02/2010
SG/EHI/PM13/Y10	Mayfair Park	Male	Adult	23/02/2010
SG/EHI/PM14/Y10	Singapore Island Country Club	Male	Adult	23/02/2010
SG/EHI/PM15/Y10	Bukit Manis Road	Male	Adult	23/02/2010
SG/EHI/PM16/Y10	Windsor Park	Male	Adult	03/03/2010
SG/EHI/PM17/Y10	Windsor Park	Male	Adult	03/03/2010
SG/EHI/PM18/Y10	Venus Drive	Male	Juvenile	03/03/2010
SG/EHI/PM19/Y10	Meng Suan Road	Male	Juvenile	03/03/2010
SG/EHI/PM20/Y10	Meng Suan Road	Male	Juvenile	03/03/2010
SG/EHI/PM21/Y10	Meng Suan Road	Female	Juvenile	09/03/2010
SG/EHI/PM22/Y10	Jalan Asas	Male	Juvenile	09/03/2010
SG/EHI/PM23/Y10	Windsor Park	Male	Juvenile	10/03/2010
SG/EHI/PM24/Y10	Windsor Park	Male	Adult	10/03/2010
SG/EHI/PM25/Y10	Windsor Park	Male	Adult	10/03/2010
SG/EHI/PM26/Y10	Mayfair Park	Male	Adult	15/03/2010
SG/EHI/PM27/Y10	Rifle Range Road	Male	Adult	17/03/2010
SG/EHI/PM28/Y10	Bukit Gombak Rise	Male	Adult	05/04/2010
SG/EHI/PM29/Y10	Andrews Crescent	Male	Adult	09/04/2010
SG/EHI/PM30/Y10	Changi Coast Road, NSRCC	Male	Juvenile	12/04/2010
SG/EHI/PM31/Y10	Pulau Ubin	Male	Adult	16/04/2010
SG/EHI/PM32/Y10	Pulau Ubin	Male	Juvenile	16/04/2010
SG/EHI/PM33/Y10	Chestnut Road	Male	Juvenile	07/05/2010
SG/EHI/PM34/Y10	Chestnut Road	Male	Juvenile	21/05/2010
SG/EHI/PM35/Y10	Pulau Ubin	Male	Adult	24/05/2010
SG/EHI/PM36/Y10	Pulau Ubin	Male	Adult	24/05/2010
SG/EHI/PM37/Y10	Windsor Park	Male	Adult	24/08/2010
SG/EHI/PM38/Y10	Linden Drive, Bukit Timah	Male	Adult	22/10/2010
SG/EHI/PM39/Y10	Sim Road	Male	Adult	26/10/2010
SG/EHI/PM40/Y10	Sim Road	Male	Adult	26/10/2010
SG/EHI/PM41/Y10	Sime Road	Male	Adult	12/11/2010
SG/EHI/PM42/Y10	Sime Road	Male	Adult	12/11/2010
SG/EHI/PM43/Y10	Sime Road	Female	Adult	12/11/2010
SG/EHI/PM44/Y10	Pulau Ubin	Female	Adult	19/11/2010
SG/EHI/PM45/Y11	Rifle Range Road	Male	Adult	09/02/2011
SG/EHI/PM46/Y11	Mount Pleasant Drive	Male	Juvenile	11/02/2011
SG/EHI/PM47/Y11	Turf Club Road	Female	Adult	24/02/2011

Appendix C continued

Code	Location	Gender	Age	Date collected
SG/EHI/PM48/Y11	Chestnut Avenue	Female	Juvenile	16/03/2011
SG/EHI/PM49/Y11	Chestnut Avenue	Female	Adult	01/04/2011
SG/EHI/PM50/Y11	Singapore Island Country Club	Male	Adult	08/04/2011
SG/EHI/PM51/Y11	West Lake Avenue	Female	Adult	08/04/2011
SG/EHI/PM52/Y11	Rifle Range Road	Male	Adult	11/04/2011
SG/EHI/PM53/Y11	Island Club Road	Male	Adult	11/04/2011
SG/EHI/PM54/Y11	Island Club Road	Male	Adult	15/04/2011
SG/EHI/PM55/Y11	Island Club Road	Male	Juvenile	15/04/2011
SG/EHI/PM56/Y11	Island Club Road	Male	Adult	15/04/2011
SG/EHI/PM57/Y11	Chestnut Avenue	Female	Adult	20/04/2011
SG/EHI/PM58/Y11	Tanjong Pagar Community Club	Male	Adult	20/04/2011
SG/EHI/PM59/Y11	Rifle Range Road	Female	Juvenile	20/04/2011
SG/EHI/PM60/Y11	Chestnut Avenue	Female	Adult	26/04/2011
SG/EHI/PM61/Y11	Chestnut Avenue	Female	Juvenile	26/04/2011
SG/EHI/PM62/Y11	Chestnut Avenue	Male	Adult	26/04/2011
SG/EHI/PM63/Y11	Chestnut Avenue	Male	Adult	26/04/2011
SG/EHI/PM64/Y11	Chestnut Avenue	Male	Adult	26/04/2011
SG/EHI/PM65/Y11	Keppel Club	Male	Adult	29/04/2011

Appendix D: Details of <u>wild</u> long-tailed macaques and results of the species-specific nested PCR assay. Macaques highlighted were tested negative for malaria parasites. Macques SG/EHI/WM1/Y07 to SG/EHI/WM86/Y11 were sampled from the restricted-access forest located in the northwestern Singapore, while SG/EHI/WM87/Y11 to SG/EHI/WM93/Y11 were from a military offshore island.

Code	Gender	Age	Date collected	Malaria screening	Species
SG/EHI/WM01/Y07	Male	Adult	01/11/2007	+	Pk, Pfi
SG/EHI/WM02/Y07	Male	Adult	01/11/2007	+	Pk, Pcy
SG/EHI/WM03/Y09	Male	Juvenile	31/03/2009	+	Pk
SG/EHI/WM04/Y09	Male	Adult	01/04/2009	+	Pk
SG/EHI/WM05/Y09	Male	Adult	06/04/2009	+	Pk, Pcy, Pfi
SG/EHI/WM06/Y09	Male	Juvenile	08/04/2009	+	Pcy
SG/EHI/WM07/Y09	Male	Juvenile	08/04/2009	+	Pcy
SG/EHI/WM08/Y09	Female	Juvenile	08/04/2009	-	-
SG/EHI/WM09/Y09	Male	Juvenile	08/04/2009	+	Pcy
SG/EHI/WM10/Y09	Male	Juvenile	08/04/2009	-	-
SG/EHI/WM11/Y09	Male	Juvenile	08/04/2009	+	Pk
SG/EHI/WM12/Y09	Male	Adult	14/04/2009	+	Pk, Pcy
SG/EHI/WM13/Y09	Male	Juvenile	27/04/2009	+	Pcy
SG/EHI/WM14/Y09	Female	Juvenile	27/04/2009	+	Pk, Pcy
SG/EHI/WM15/Y09	Male	Juvenile	27/04/2009	+	Pk, Pfi, Pct
SG/EHI/WM16/Y09	Female	Adult	07/05/2009	+	Pk, Pcy
SG/EHI/WM17/Y09	Female	Juvenile	25/06/2009	+	Pk
SG/EHI/WM18/Y09	Female	Juvenile	09/07/2009	+	Pcy, Pfi
SG/EHI/WM19/Y09	Female	Juvenile	09/07/2009	+	Pk, Pcy, Pfi
SG/EHI/WM20/Y09	Female	Adult	07/08/2009	+	Pk
SG/EHI/WM21/Y09	Female	Juvenile	14/08/2009	+	Pcy
SG/EHI/WM22/Y09	Male	Adult	18/08/2009	-	-
SG/EHI/WM23/Y09	Female	Adult	01/09/2009	+	Pcy
SG/EHI/WM24/Y09	Female	Juvenile	09/09/2009	+	Pk
SG/EHI/WM25/Y09	Male	Adult	15/09/2009	+	Pcy
SG/EHI/WM26/Y09	Male	Adult	25/09/2009	+	Pk
SG/EHI/WM27/Y09	Male	Adult	02/10/2009	-	-
SG/EHI/WM28/Y09	Female	Adult	06/10/2009	-	-
SG/EHI/WM29/Y09	Male	Adult	13/10/2009	+	Pk, Pcy
SG/EHI/WM30/Y09	Male	Adult	16/10/2009	+	Pcy
SG/EHI/WM31/Y09	Female	Adult	20/10/2009	+	Pk, Pcy
SG/EHI/WM32/Y09	Male	Adult	05/11/2009	+	Pk, Pcy, Pfi
SG/EHI/WM33/Y09	Male	Adult	05/11/2009	+	Pk, Pcy
SG/EHI/WM34/Y09	Male	Juvenile	05/11/2009	+	Pcy
SG/EHI/WM35/Y09	Female	Juvenile	12/11/2009	+	Pk
SG/EHI/WM36/Y09	Male	Adult	25/11/2009	+	Pcy
SG/EHI/WM37/Y09	Male	Adult	04/12/2009	+	Pcy, Pfi
SG/EHI/WM38/Y09	Female	Adult	22/12/2009	-	-
SG/EHI/WM39/Y09	Male	Adult	22/12/2009	+	Pk, Pcy, Pfi
SG/EHI/WM40/Y10	Female	Adult	06/01/2010	-	-
SG/EHI/WM41/Y10	Male	Adult	06/01/2010	+	Pcy
SG/EHI/WM42/Y10	Male	Adult	22/01/2010	+	Pcy
SG/EHI/WM43/Y10	Male	Adult	27/01/2010	-	-
SG/EHI/WM44/Y10	Female	Adult	11/02/2010	+	Pcy, Pfi

Appendix D continued

Code	Gender	Age	Date collected	Malaria screening	Species
SG/EHI/WM45/Y10	Male	Juvenile	12/02/2010	+	Pk, Pcy
SG/EHI/WM46/Y10	Male	Adult	12/02/2010	+	Pcy
SG/EHI/WM47/Y10	Male	Adult	23/02/2010	+	Pcy
SG/EHI/WM48/Y10	Male	Juvenile	23/02/2010	+	Pk
SG/EHI/WM49/Y10	Female	Adult	23/02/2010	+	Pk
SG/EHI/WM50/Y10	Female	Adult	16/11/2010	+	Pk
SG/EHI/WM51/Y10	Male	Adult	19/11/2010	+	Pk, Pcy
SG/EHI/WM52/Y10	Male	Adult	23/11/2010	+	Pk
SG/EHI/WM53/Y10	Male	Juvenile	23/11/2010	-	-
SG/EHI/WM54/Y10	Female	Adult	23/11/2010	+	Pk
SG/EHI/WM55/Y10	Male	Juvenile	25/11/2010	+	Pcy
SG/EHI/WM56/Y10	Female	Adult	29/11/2010	+	Pk, Pfi
SG/EHI/WM57/Y10	Female	Adult	29/11/2010	+	Pk, Pcy
SG/EHI/WM58/Y09	Male	Adult	03/12/2010	-	-
SG/EHI/WM59/Y10	Male	Adult	08/12/2010	+	Pk, Pct
SG/EHI/WM60/Y10	Female	Adult	08/12/2010	-	-
SG/EHI/WM61/Y10	Female	Adult	10/12/2010	-	-
SG/EHI/WM62/Y10	Male	Adult	14/12/2010	-	-
SG/EHI/WM63/Y10	Male	Adult	14/12/2010	+	Pk
SG/EHI/WM64/Y10	Too small for blood collection				
SG/EHI/WM65/Y10	Female	Adult	16/12/2010	+	Pk, Pfi
SG/EHI/WM66/Y10	Female	Adult	16/12/2010	+	Pk, Pcy
SG/EHI/WM67/Y10	Female	Adult	20/12/2010	-	-
SG/EHI/WM68/Y10	Male	Adult	22/12/2010	+	Pk
SG/EHI/WM69/Y10	Male	Juvenile	22/12/2010	+	Pk
SG/EHI/WM70/Y10	Female	Adult	22/12/2010	+	Pk, Pcy
SG/EHI/WM71/Y10	Female	Adult	27/12/2010	+	Pcy
SG/EHI/WM72/Y10	Female	Adult	30/12/2010	-	-
SG/EHI/WM73/Y10	Female	Adult	30/12/2010	+	Pk, Pcy
SG/EHI/WM74/Y11	Female	Juvenile	04/01/2011	+	Pk
SG/EHI/WM75/Y11	Male	Juvenile	04/01/2011	+	Pcy
SG/EHI/WM76/Y11	Male	Adult	07/01/2011	+	Pk
SG/EHI/WM77/Y11	Female	Juvenile	18/01/2011	-	-
SG/EHI/WM78/Y11	Female	Adult	28/01/2011	-	-
SG/EHI/WM79/Y11	Female	Adult	03/03/2011	+	Pcy
SG/EHI/WM80/Y11	Male	Adult	10/03/2011	+	Pk
SG/EHI/WM81/Y11	Female	Juvenile	17/03/2011	+	Pk, Pcy
SG/EHI/WM82/Y11	Male	Adult	25/03/2011	-	-
SG/EHI/WM83/Y11	Male	Juvenile	29/03/2011	+	Pk
SG/EHI/WM84/Y11	Female	Juvenile	01/04/2011	+	Pk, Pcy
SG/EHI/WM85/Y11	Male	Juvenile	15/04/2011	-	-
SG/EHI/WM86/Y11	Male	Juvenile	26/04/2011	-	-
SG/EHI/WM87/Y11	Male	Adult	20/05/2011	-	-
SG/EHI/WM88/Y11	Male	Adult	20/05/2011	-	-
SG/EHI/WM89/Y11	Male	Adult	20/05/2011	-	-
SG/EHI/WM90/Y11	Male	Adult	20/05/2011	-	-
SG/EHI/WM91/Y11	Male	Adult	20/05/2011	+	Pk, Pin
SG/EHI/WM92/Y11	Male	Adult	20/05/2011	-	-
SG/EHI/WM93/Y111	Female	Adult	20/05/2011	-	-

Plasmodium coatneyi, *P. cynomolgi*, *P. fieldi*, *P. inui* and *P. knowlesi* were denoted by Pct, Pcy, Pfi, Pin and Pk, respectively.

Appendix E: DNA sequences of the *csp* genes (with GenBank accession number)

P. coatneyi, CDC (JQ219880)

```
ATGAAGAACTTCATTCTCTTGGCCGTTTCTTCCATCCTGTTGGTGGACTTGTTCCCCACGCACTTCGGACATAATGTAGATCTCTCCAGGGCCATAAATT  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  F  P  T  H  F  G  H  N  V  D  L  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAATGTAGACACCAGTTTACTTGGCGCAGCACAGGTAAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAAACCAAA  200
L  N  G  V  S  F  N  N  V  D  T  S  L  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  K  P  K

AAAAAAGGCGGAAAAAAAAGAAGAAGAACCAAAAAAGCCAAATGAAAATAAGCTGAAGCAACCAGTAGATGGAGCACGAGATGGGCCAGCACCAGCAGCA  300
  K  K  A  E  K  K  E  E  E  P  K  K  P  N  E  N  K  L  K  Q  P  V  D  G  A  R  D  G  P  A  P  A  A

GATGGAGCAAGAGATGGACCAGCACCAGCAGCAGATGGAGCACGAGATGGACCAGCACCAGCAGCAGATGGAGCAAGAGATGGACCAGCACCAGCAGCAG  400
 D  G  A  R  D  G  P  A  P  A  A  D  G  A  R  D  G  P  A  P  A  A  D  G  A  R  D  G  P  A  P  A  A

ATGGAGCAAGAGATGGACCAGCACCAGCAGCAGATGGAGCAAGAGATGGACCAGCACCAGCAGCAGATGGAGCACGAGATGGACCAGCACCAGCAGCAGA  500
D  G  A  R  D  G  P  A  P  A  A  D  G  A  R  D  G  P  A  P  A  A  D  G  A  R  D  G  P  A  P  A  A  D

TGGAGCAAGAGATGGGCCAGCACCACCAGCCGATGGAGCAAGAGATGGGCCAGCACCACCAGCAGCAGATGGAGCACGAGATGGACCAGCACCACCAGCA  600
  G  A  R  D  G  P  A  P  P  A  D  G  A  R  D  G  P  A  P  P  A  A  D  G  A  R  D  G  P  A  P  P  A

GCAGATGGAGCACGAGATGGGCCAGCACCACCAGCAGGACAAGGAGGAGGAGGAAATGCAGCAGGCCAAGCACAAGGAGGAGGAGGAAATGCCGGAAACAAAAAAG  700
 A  D  G  A  R  D  G  P  A  P  P  A  G  Q  G  G  G  N  A  A  G  Q  A  Q  G  G  G  N  A  G  N  K  K

CAGGAGACGCAGCTGGAAACGCAGGAGCAGCAAAAGGACAGGGACAAAATAATGAAGGTGCGAATGTCCCAAATGAGAAAGTTGTGAATGATTACCTACA  800
A  G  D  A  A  G  N  A  G  A  A  K  G  Q  G  Q  N  N  E  G  A  N  V  P  N  E  K  V  V  N  D  Y  L  Q

GAAAATTAGATCTACCGTTACCACCGAATGGACTCCATGCAGTGTAACCTGTGGAAATGGTGTAAGACTTAGAAGAAAAGCTCATGCAGAAAAGAAAAAA  900
  K  I  R  S  T  V  T  T  E  W  T  P  C  S  V  T  C  G  N  G  V  R  L  R  R  K  A  H  A  E  K  K  K

CCAGAGGACCTTACCATGGATGACCTTGACGTGGAAGTTTGTGCAATGGATAAGTGCGCTGGCATATTTAACTTTGTGAGTAATTCATTAGGGCTAGTCA  1000
 P  E  D  L  T  M  D  D  L  D  V  E  V  C  A  M  D  K  C  A  G  I  F  N  F  V  S  N  S  L  G  L  V

TATTGTTAGTCCTAGCATTATTCAATTAA
I  L  L  V  L  A  L  F  N  .
```

P. fieldi, CDC (JQ219881)

```
ATGAAGAACTTCATTCTCTTGGCCGTTTCTTCCATCCTGTTGGTGGACTTGTTCCCCACACACTGCGGGCATAATGTAGATCTCTCCAGGGCCATAAATT  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  F  P  T  H  C  G  H  N  V  D  L  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAATGTAGACGCCAGTTCACTTGGCGCAGCACAGGTAAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAACCCAAA  200
L  N  G  V  S  F  N  N  V  D  A  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  N  P  K

AGACGATGAAAAAGCTGATAAACCAAAAAAAAAGGACGAAAAAAAAGTAGAACCAAAAAAGCCACATGAAAATAAGCTGAAACAACCAGTCCCAGGAGCA  300
  D  D  E  K  A  D  K  P  K  K  K  D  E  K  K  V  E  P  K  K  P  H  E  N  K  L  K  Q  P  V  P  G  A

AATCAGGAAGGCGGAGCAGCAGCCCCAGGAGCAAATCAGGAAGGTGGAGCAGCAGCACCAGGAGCAAACCAGGGAGGTGGAGCAGCAGCACCAGGAGCAA  400
 N  Q  E  G  G  A  A  A  P  G  A  N  Q  E  G  G  A  A  A  P  G  A  N  Q  G  G  G  A  A  A  P  G  A

ACCAGGAAGGTGGAGCAGCAGCACCAGGAGCAAACCAGGGAGGTGCAAAGCCAGCAGGAGTACAGGGACAAAATAATGAAGGTGCGAATAAGCCAGATGA  500
N  Q  E  G  G  A  A  A  P  G  A  N  Q  G  G  A  K  P  A  G  V  Q  G  Q  N  N  E  G  A  N  K  P  D  E

AAAGCATGTGAAAGAATATCTAGAGAAAATTAGATCTACCGTTGGCACCGAATGGACTCCATGCAGTGTAACCTGTGGAAAGGGTGTAAGAGTTAGAAGA  600
  K  H  V  K  E  Y  L  E  K  I  R  S  T  V  G  T  E  W  T  P  C  S  V  T  C  G  K  G  V  R  V  R  R

AAACTTAATGCAGGTGACAAAAAACCAGATAAGCTTACTCTGAATGACCTTGAGGCAGAAGTTTGTACAATGGATAAGTGCGCTGGCATATTTAACGTTG  700
 K  L  N  A  G  D  K  K  P  D  K  L  T  L  N  D  L  E  A  E  V  C  T  M  D  K  C  A  G  I  F  N  V

TGAGTAATTCATTAGGGCTAGTCATATTATTAGTCCTAGCATTATTCAATTAA
V  S  N  S  L  G  L  V  I  L  L  V  L  A  L  F  N  .
```

P. inui, CDC (JQ219882)

```
ATGAAGAACTTCATTCTCTTGGCCGTTTCGTCCATCCTGTTGGTGGACTTATTCCCCACACACTGCGGGCATGATGTAGATCTCTCCAGAGCCATAAATT  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  F  P  T  H  C  G  H  D  V  D  L  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAATGTAGACGCCAGTTCACTTGGCGCCGCACAGGTAAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAAGACCCAAA  200
L  N  G  V  S  F  N  N  V  D  A  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  D  P  K

AGACCAGGAAGGAGGTCCTAAGGGAAAAAAGAAGGGAAAAAAAGGAGAACCAAAAAACCCACCTGAAAAGAACCTGAAACAGCCAGCACAAGACCCCGGA  300
  D  Q  E  G  G  P  K  G  K  K  K  G  K  K  G  E  P  K  N  P  P  E  K  N  L  K  Q  P  A  Q  D  P  G

GCACCAGGACAGGATCCAGGAGCACCAGGACAGGATCCAGGAGCACCAGGACAGGATCCAGGAGCACCAGGACAGGACCCAGGAGCACCAGGACAGGACC  400
 A  P  G  Q  D  P  G  A  P  G  Q  D  P  G  A  P  G  Q  D  P  G  A  P  G  Q  D  P  G  A  P  G  Q  D

CAGGAGCACCAGGACAGGCCCCAGGAGCACCAGGACAGGACCCAGGAGCACCAGGACAGGACCCAGGAGCACCAGGACAGGACCCAGGAGCACCAGGACA  500
P  G  A  P  G  Q  A  P  G  A  P  G  Q  D  P  G  A  P  G  Q  D  P  G  A  P  G  Q  D  P  G  A  P  G  Q

GGCCCCAGGAGCACCAGGACAGGACCCAGGAGCACCAGGACAGGCCCCAGGAGCACCAGGACAGGACCCAGGAGCACCAGGACAGGATCCAGCAGCACCA  600
  A  P  G  A  P  G  Q  D  P  G  A  P  G  Q  A  P  G  A  P  G  Q  D  P  G  A  P  G  Q  D  P  A  A  P

GGACAGGATCCAGCAGCACCAGCCCGAGACCCAGCAGCACCAGGACAGGATCCAGGAGCACCAGCACAAAACCCAGGTGGTCCAGCACAAAACCCAGGAG  700
 G  Q  D  P  A  A  P  A  R  D  P  A  A  P  G  Q  D  P  G  A  P  A  Q  N  P  G  G  P  A  Q  N  P  G

GACCAGCACAAAAGCCAGCAGGACCAGCACAGAAGCCAGCAGGACCAGCACGAAACGCAGGAGAAGGAGCAGGAAATGCGGGAGATGATAATGCCCCAGA  800
G  P  A  Q  K  P  A  G  P  A  Q  K  P  A  G  P  A  R  N  A  G  E  G  A  G  N  A  G  D  D  N  A  P  D

TGAAAAGGTTGTGAAAGATTACCTAGAGAAAGTGAAATCTAACCTTACCACCGAATGGAGTGTATGCAGTGTAAGCTGTGGACAGGGTGTAAGAGTTAGA  900
  E  K  V  V  K  D  Y  L  E  K  V  K  S  N  L  T  T  E  W  S  V  C  S  V  S  C  G  Q  G  V  R  V  R

AGAAAAGTTAGTGCATCTAACAAGAAACCAGAGGAACTTACTCTGGATGACCTTGAGGTAGAAATTTGTAAAATGGAGAAGTGCGCTAGCATATTTAACG  1000
 R  K  V  S  A  S  N  K  K  P  E  E  L  T  L  D  D  L  E  V  E  I  C  K  M  E  K  C  A  S  I  F  N

TTGTGAGTAATTCATTAGGGGTAGTAATATTGCTAGTCCTAGCATTATTCAATTAA
V  V  S  N  S  L  G  V  V  I  L  L  V  L  A  L  F  N  .
```

SG/EHI/H1/Y07-15 (*P. knowlesi*; JQ219893)

```
ATGAAGAACTTCATTCTCTTGGCCGTCTCCTCCATCCTGCTGGTGGACTTGCTCCCCACACACTTCGAACATAATGTAGATCTCTCCAGGGCCATAAATG  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  L  P  T  H  F  E  H  N  V  D  L  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAATGTAGACACCAGTTCACTTGGCGCAGCACAGGTGAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAAGCCAAA  200
V  N  G  V  S  F  N  N  V  D  T  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  K  P  K

AGAAGGAGCTGATAAAGAAAAGAAAAAAGAAAAAGAAAAAGAAAAAGAAGAACCAAAGAAGCCAAATGAAAATAAGCTGAAACAACCGGAACAACCAGCA  300
  E  G  A  D  K  E  K  K  K  E  K  E  K  E  K  E  E  P  K  K  P  N  E  N  K  L  K  Q  P  E  Q  P  A

GCAGGAGCAGGGGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAAC  400
 A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E

AACCAGCGGCAGGAGCAGGGGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCAGCAGGAGCAGG  500
Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G

AGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCAGCAGGAGCAGGGGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCACCA  600
  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  P

GCACCAAGGAGGGAACAACCAGCAGCAGGAGCAGGGGGCGAACAACCAGCACCAGCACCAAGGAGGGAACAACCAGCACCAGGAGCAGGTGCGGGAGATG  700
 A  P  R  R  E  Q  P  A  A  G  A  G  G  E  Q  P  A  P  A  P  R  R  E  Q  P  A  P  G  A  G  A  G  D

GAGCACGAGGAGGAAACGCAGGGGCAGGTAAAGGACAGGGACAAAACAATCAGGGTGCGAATGTCCCAAATGAAAAAGTTGTGAATGATTACCTACACAA  800
G  A  R  G  G  N  A  G  A  G  K  G  Q  G  Q  N  N  Q  G  A  N  V  P  N  E  K  V  V  N  D  Y  L  H  K

AATTAGATCTAGCGTTACCACCGAGTGGACTCCATGCAGTGTAACCTGTGGAAATGGTGTAAGAATTAGAAGAAGACAGAATGCTGGTAATAAAAAGGCA  900
  I  R  S  S  V  T  T  E  W  T  P  C  S  V  T  C  G  N  G  V  R  I  R  R  R  Q  N  A  G  N  K  K  A

GAGGACCTTACTATGGATGACCTTGAGGTGGAAGCTTGTGTAATGGATAAGTGCGCTGGCATATTTAACGTTGTGAGTAATTCATTAGGCTTAGTCATAT  1000
 E  D  L  T  M  D  D  L  E  V  E  A  C  V  M  D  K  C  A  G  I  F  N  V  V  S  N  S  L  G  L  V  I

TGTTAGTCTTAGCATTATTCAATTAA
L  L  V  L  A  L  F  N  .
```

SG/EHI/H2/Y07-12 (*P. knowlesi*; JQ219894)

```
ATGAAGAACTTCATTCTCTTGGCCGTCTCCTCCATCCTGCTGGTGGACTTGCTCCCCACACACTTCGAACATAATGTAGATCTCTCCAGGGCCATAAATG  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  L  P  T  H  F  E  H  N  V  D  L  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAATGTAGACACCAGTTCACTTGGCGCAGCACAGGTGAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAAGCCAAA  200
V  N  G  V  S  F  N  N  V  D  T  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  K  P  K

AGAAGGAGCTGATAAAGAAAAGAAAAAAGAAAAAGAAAAAGAAAAAGAAGAACCAAAGAAGCCAAATGAAAATAAGCTGAAACAACCGGAACAACCAGCA  300
  E  G  A  D  K  E  K  K  K  E  K  E  K  E  K  E  E  P  K  K  P  N  E  N  K  L  K  Q  P  E  Q  P  A

GCAGGAGCAGGGGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAAC  400
 A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E

AACCAGCGGCAGGAGCAGGGGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCAGCAGGAGCAGG  500
Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G

AGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCAGCAGGAGCAGGGGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCACCA  600
  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  P

GCACCAAGGAGGGAACAACCAGCAGCAGGAGCAGGGGGCGAACAACCAGCACCAGCACCAAGGAGGGAACAACCAGCACCAGGAGCAGGTGCGGGAGATG  700
 A  P  R  R  E  Q  P  A  A  G  A  G  G  E  Q  P  A  P  A  P  R  R  E  Q  P  A  P  G  A  G  A  G  D

GAGCACGAGGAGGAAACGCAGGGGCAGGTAAAGGACAGGGACAAAACAATCAGGGTGCGAATGTCCCAAATGAAAAAGTTGTGAATGATTACCTACACAA  800
G  A  R  G  G  N  A  G  A  G  K  G  Q  G  Q  N  N  Q  G  A  N  V  P  N  E  K  V  V  N  D  Y  L  H  K

AATTAGATCTAGCGTTACCACCGAGTGGACTCCATGCAGTGTAACCTGTGGAAATGGTGTAAGAATTAGAAGAAGACAGAATGCTGGTAATAAAAAGGCA  900
  I  R  S  S  V  T  T  E  W  T  P  C  S  V  T  C  G  N  G  V  R  I  R  R  R  Q  N  A  G  N  K  K  A

GAGGACCTTACTATGGATGACCTTGAGGTGGAAGCTTGTGTAATGGATAAGTGCGCTGGCATATTTAACGTTGTGAGTAATTCATTAGGCTTAGTCATAT 1000
 E  D  L  T  M  D  D  L  E  V  E  A  C  V  M  D  K  C  A  G  I  F  N  V  V  S  N  S  L  G  L  V  I

TGTTAGTCTTAGCATTATTCAATTAA
L  L  V  L  A  L  F  N  .
```

SG/EHI/H7/Y07-01 (*P. knowlesi*; JQ219895)

```
ATGAAGAACTTCATTCTCTTGGCCGTCTCCTCCATCCTGCTGGTGGACTTGCTCCCCACACACTTCGAACATAATGTAGATCTCTCCAGGGCCATAAATG  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  L  P  T  H  F  E  H  N  V  D  L  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAATGTAGACACCAGTTCACTTGGCGCAGCACAGGTGAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAAGCCAAA  200
V  N  G  V  S  F  N  N  V  D  T  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  K  P  K

AGAAGGAGCTGATAAAGAAAAGAAAAAAGAAAAAGGAAAAGAAAAAGAAGAAGAACCAAAGAAGCCAAATGAAAATAAGCTGAAACAACCGAATGAAGGA  300
  E  G  A  D  K  E  K  K  K  E  K  G  K  E  K  E  E  E  P  K  K  P  N  E  N  K  L  K  Q  P  N  E  G

CAACCACAAGCACAGGGTGATGGAGCAAATGCAGGACAACCACAAGCACAAGGAGATGGAGCAAATGCAGGACAACCACAAGCACAGGGTGATGGAGCAA  400
Q  P  Q  A  Q  G  D  G  A  N  A  G  Q  P  Q  A  Q  G  D  G  A  N  A  G  Q  P  Q  A  Q  G  D  G  A

ATGCAGGACAACCACAAGCACAGGGTGATGGAGCAAATGCAGGACAACCACAAGCACAAGGAGATGGAGCAAATGCAGGACAACCACAAGCACAGGGTGA  500
N  A  G  Q  P  Q  A  Q  G  D  G  A  N  A  G  Q  P  Q  A  Q  G  D  G  A  N  A  G  Q  P  Q  A  Q  G  D

TGGAGCAAATGCAGGGCAACCACAAGCACAGGGTGATGGAGCAAATGCAGGACAACCACAAGCACAAGGAGATGGAGCAAATGCAGGACAACCACAAGCA  600
  G  A  N  A  G  Q  P  Q  A  Q  G  D  G  A  N  A  G  Q  P  Q  A  Q  G  D  G  A  N  A  G  Q  P  Q  A

CAAGGAGATGGAGCAAATGCAGGACAACCACAAGCACAGGGTGATGGAGCAAATGCAGGACAACCACAAGCACAGGGTGATAGGGCGAATGCAGGACAAC  700
Q  G  D  G  A  N  A  G  Q  P  Q  A  Q  G  D  G  A  N  A  G  Q  P  Q  A  Q  G  D  R  A  N  A  G  Q

CACAAGCACAAGGAGATGGGGCAAATGTACCACGACAAGGAAGAAACGGGGGAGGTGCACCAGCAGGAGGAAATGAGGGGAATAAACAAGCAGGAAAAGG  800
P  Q  A  Q  G  D  G  A  N  V  P  R  Q  G  R  N  G  G  G  A  P  A  G  G  N  E  G  N  K  Q  A  G  K  G

ACAGGGACAAAACAATCAGGGTGCGAATGCCCCAAATGAAAAAGTTGTGAATGATTACCTACACAAAATTAGATCTAGCGTTACCACCGAGTGGACTCCA  900
  Q  G  Q  N  N  Q  G  A  N  A  P  N  E  K  V  V  N  D  Y  L  H  K  I  R  S  S  V  T  T  E  W  T  P

TGCAGTGTAACCTGTGGAAATGGTGTAAGAATTAGAAGAAAAGCTCATGCAGGTAATAAAAAGGCAGAGGACCTTACTATGGATGACCTTGAGGTGGAAG 1000
 C  S  V  T  C  G  N  G  V  R  I  R  R  K  A  H  A  G  N  K  K  A  E  D  L  T  M  D  D  L  E  V  E

CTTGTGTAATGGATAAGTGCGCTGGCATATTTAACGTTGTGAGTAATTCATTAGGCTTAGTCATATTGTTAGTCCTAGCATTATTCAATTAA
A  C  V  M  D  K  C  A  G  I  F  N  V  V  S  N  S  L  G  L  V  I  L  L  V  L  A  L  F  N  .
```

SG/EHI/H24/Y08-10 (*P. knowlesi*; JQ219896)

```
ATGAAGAACTTCATTCTCTTGGCCGTCTCCTCCATCCTGCTGGTGGACTTGCTCCCCACACACTTCGAACATAATGTAGATCTCTCCAGGGCCATAAATG  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  L  P  T  H  F  E  H  N  V  D  L  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAATGTAGACACCAGTTCACTTGGCGCAGCACAGGTGAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAAGCCAAA  200
V  N  G  V  S  F  N  N  V  D  T  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  K  P  K

AGAAGGAGCTGATAAAGAAAAGAAAAAAGAAAAAGAAGAAGAACCAAAGAAGCCAAATGAAAATAAGCTGAAACAACCGGATGCAGTACCAGGGGGCGAA  300
  E  G  A  D  K  E  K  K  K  E  K  E  E  E  P  K  K  P  N  E  N  K  L  K  Q  P  D  A  V  P  G  G  E

GAACCAGCACCAGGAAGGGAACAGCCAGCACCAGGAAGGGAGGAACCAGCACCAGGAAGGGAACAGCCAGCACCAGGAAGGGAAGAACCAGCACCAGGAA  400
 E  P  A  P  G  R  E  Q  P  A  P  G  R  E  E  P  A  P  G  R  E  Q  P  A  P  G  R  E  E  P  A  P  G

GGGAACAGCCAGCACCAGGAAGGGAAGAACCAGCGCCAGGAAGGGAACAGCCAGCACCAGGAAGGGAACAGCCAGCACCAGGAAGGGAGGAACCAGCACC  500
R  E  Q  P  A  P  G  R  E  E  P  A  P  G  R  E  Q  P  A  P  G  R  E  Q  P  A  P  G  R  E  E  P  A  P

AGGAAGGGAACAACCAGCACCAGGAAGGGAGGAACCAGCACCAGGAAGGGAACAACCAGCACCAGGAAGGGAACAGCCAGCACCAGGAAGGGAACAGCCA  600
  G  R  E  Q  P  A  P  G  R  E  E  P  A  P  G  R  E  Q  P  A  P  G  R  E  Q  P  A  P  G  R  E  Q  P

GCACCGGGGGGTGAACAACCAGCACCAGGAAGGGAACAGCCAGCACCGGGTGGTGAACAACCAGCACCAGCACCAGGAGCAGGTGCGGGAGATGGAGCAC  700
 A  P  G  G  E  Q  P  A  P  G  R  E  Q  P  A  P  G  G  E  Q  P  A  P  A  P  G  A  G  A  G  D  G  A

GAGGAGGAAACGCAGGGGCAGGGAAAGGACAGGGACAAAACAATCAGGGTGCAAATGTCCCAAATGAAAAAGTTGTGAATGATTACCTACACAAAATTAG  800
R  G  G  N  A  G  A  G  K  G  Q  G  Q  N  N  Q  G  A  N  V  P  N  E  K  V  V  N  D  Y  L  H  K  I  R

ATCTAGCGTTACCACCGAGTGGACTCCATGCAGTGTAACCTGTGGAAATGGTGTAAGAATTAGAAGAAAAGGTCATGCAGGTAATAAAAAGGCAGAGGAC  900
  S  S  V  T  T  E  W  T  P  C  S  V  T  C  G  N  G  V  R  I  R  R  K  G  H  A  G  N  K  K  A  E  D

CTTACTATGGATGACCTTGAGGTGGAAGCTTGTGTAATGGATAAGTGCGCTGGCATATTTAACGTTGTGAGTAATTCATTAGGCTTAGTCATATTGTTAG  1000
 L  T  M  D  D  L  E  V  E  A  C  V  M  D  K  C  A  G  I  F  N  V  V  S  N  S  L  G  L  V  I  L  L

TCCTAGCATTATTCAATTAA
V  L  A  L  F  N  .
```

SG/EHI/H1-im/Y09 -17 (*P. knowlesi*; JQ219883)

```
ATGAGGAACTTCATTCTCTTGGCCGTCTCCTCCATCCTGCTGGTGGACTTGCTCCCCACACACTTCGAACATAATGTAGATCTCTCCAGGGCCATAAATG  100
 M  R  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  L  P  T  H  F  E  H  N  V  D  L  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAATGTAGACACCAGTTCACTTGGCGCAGCACAGGTAAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAAGCCAAA  200
V  N  G  V  S  F  N  N  V  D  T  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  K  P  K

AGAAGGAGATGATAAAGAAAAGAAAAAAGAAAAAGAAAAAGAAGAAGAACCAAAGAAGCTAAATGAAAATAATCCGAAAAAACCGAATGAAGAAGGTGAT  300
  E  G  D  D  K  E  K  K  K  E  K  E  K  E  E  E  P  K  K  L  N  E  N  N  P  K  K  P  N  E  E  G  D

GGAGCTAAGCTGAAACAACCGAATGAAGAAGGTGATGGAGCTAAGCTGAAACAACCGAATGCAGAAGGTGGAGCTAAGCTGAAACAACCGAATGCGGAAG  400
 G  A  K  L  K  Q  P  N  E  E  G  D  G  A  K  L  K  Q  P  N  A  E  G  G  A  K  L  K  Q  P  N  A  E

GTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGGTGG  500
G  G  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  G  G

AGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGA  600
  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G

GCAAATGCACGACAACCGAATGCAGAAGGTGGAGCAAATGCACGACAACCTAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAG  700
 A  N  A  R  Q  P  N  A  E  G  G  A  N  A  R  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G

CAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCACGACAACCTAATGCAGAAGGTGGAGC  800
A  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  R  Q  P  N  A  E  G  G  A

AAACGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCACGACAGCCACAGGCAGAAGGTGGTGGA  900
  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  R  Q  P  Q  A  E  G  G  G

GCAAATGCACGACAGCCACAGGCAGAAGGTGGTGGAGCAAATGCACGACAAGGAGGAAATGAGGGGAATAAACAAGCAGGAAAAGGACAGGGACAAAACA  1000
 A  N  A  R  Q  P  Q  A  E  G  G  G  A  N  A  R  Q  G  G  N  E  G  N  K  Q  A  G  K  G  Q  G  Q  N

ATCAGGGTGCGAATGCCCCAAATGAAAAAGTTGTAAATGATTACCTACAGAAAATTAGATCTAGCGTTACCATCGAGTGGACTCCATGCAGTGTAACCTG  1100
N  Q  G  A  N  A  P  N  E  K  V  V  N  D  Y  L  Q  K  I  R  S  S  V  T  I  E  W  T  P  C  S  V  T  C

TGGAAATGGTGTAAGAATTAGAAGAAGAGCTCATGCAGATAAGAAAAAGGCAGAGGACCTTACTATGGATGACCTTGAAGTGGAAGCTTGTGTAATGGAT  1200
  G  N  G  V  R  I  R  R  R  A  H  A  D  K  K  K  A  E  D  L  T  M  D  D  L  E  V  E  A  C  V  M  D

AAGTGTGCTGGCATATTTAACGTTGTGAGTAATTCATTAGGGTTAGTAATATTGTTAGTCCTAGCATTATTCAATTAA
 K  C  A  G  I  F  N  V  V  S  N  S  L  G  L  V  I  L  L  V  L  A  L  F  N  .
```

SG/EHI/H1-im/Y09 -25 (*P. knowlesi*; JQ219884)

```
ATGAAGAACTTCATTCTCTTGGCCGTCTCCTCCATCCTGCTGGTGGACTTGCTCCCCACACACTTCGAACATAATGTAGATCTCTCCAGGGCCATAAATG  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  L  P  T  H  F  E  H  N  V  D  L  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAATGTAGACACCAGTTCACTTGGCGCAGCACAGGTAAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAAGCCAAA  200
V  N  G  V  S  F  N  N  V  D  T  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  K  P  K

AGAAGGAGATGATAAAGAAAAGAAAAAAGAAAAAGAAAAAGAAGAAGAACCAAAGAAGCTAAATGAAAATAATCCGAAAAAACCGAATGAAGAAGGTGAT  300
  E  G  D  D  K  E  K  K  K  E  K  E  K  E  E  E  P  K  K  L  N  E  N  N  P  K  K  P  N  E  E  G  D

GGAGCTAAGCTGAAACAACCGAATGAAGAAGGTGATGGAGCTAAGCTGAAACAACCGAATGCAGAAGGTGGAGCTAAGCTGAAACAACCGAATGCAGAAG  400
 G  A  K  L  K  Q  P  N  E  E  G  D  G  A  K  L  K  Q  P  N  A  E  G  G  A  K  L  K  Q  P  N  A  E

GTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGGTGG  500
G  G  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  G  G

AGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGA  600
  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G

GCAAATGCACGACAACCGAATGCAGAAGGTGGAGCAAATGCACGACAACCTAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAG  700
 A  N  A  R  Q  P  N  A  E  G  G  A  N  A  R  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G

CAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCACGACAACCTAATGCAGAAGGTGGAGC  800
A  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  R  Q  P  N  A  E  G  G  A

AAACGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCACGACAGCCACAGGCAGAAGGTGGTGGA  900
  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  R  Q  P  Q  A  E  G  G  G

GCAAATGCACGACAGCCACAGGCAGAAGGTGGTGGAGCAAATGCACGACAAGGAGGAAATGAGGGGAATAAACAAGCAGGAAAAGGACAGGGACAAAACA  1000
 A  N  A  R  Q  P  Q  A  E  G  G  G  A  N  A  R  Q  G  G  N  E  G  N  K  Q  A  G  K  G  Q  G  Q  N

ATCAGGGTGCGAATGCCCCAAATGAAAAAGTTGTAAATGACTACCTACAGAAAATTAGATCTAGCGTTACCATCGAGTGGACTCCATGCAGTGTAACTTG  1100
N  Q  G  A  N  A  P  N  E  K  V  V  N  D  Y  L  Q  K  I  R  S  S  V  T  I  E  W  T  P  C  S  V  T  C

TGGAAATGGTGTAAGAATTAGAAGAAGAGCTCATGCAGATAAGAAAAAGGCAGAGGACCTTACTATGGATGACCTTGAAGTGGAAGCTTGTGTAATGGAT  1200
  G  N  G  V  R  I  R  R  R  A  H  A  D  K  K  K  A  E  D  L  T  M  D  D  L  E  V  E  A  C  V  M  D

AAGTGTGCTGGCATATTTAACGTTGTGAGTAATTCATTAGGGTTAGTAATATTGTTAGTCCTAGCATTATTCAATTAA
 K  C  A  G  I  F  N  V  V  S  N  S  L  G  L  V  I  L  L  V  L  A  L  F  N  .
```

SG/EHI/H1-im/Y09 -80 (*P. knowlesi*; JQ219885)

```
ATGAAGAACTTCATTCTCTTGGCCGTCTCCTCCATCCTGCTGGTGGACTTGCTCCCCACACACTTCGAACATAATGTAGATCTCTCCAGGGCCATAAATG  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  L  P  T  H  F  E  H  N  V  D  L  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAATGTAGACACCAGTTCACTTGGCGCAGCACAGGTAAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAAGCCAAA  200
V  N  G  V  S  F  N  N  V  D  T  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  K  P  K

AGAAGGAGATGATAAAGAAAAGAAAAAAGAAAAAGAAAAAGAAGAAGAACCAAAGAAGCTAAATGAAAATAATCCGAAAAAACCGAATGAAGAAGGTGAT  300
  E  G  D  D  K  E  K  K  K  E  K  E  K  E  E  E  P  K  K  L  N  E  N  N  P  K  K  P  N  E  E  G  D

GGAGCTAAGCTGAAACAACCGAATGAAGAAGGTGATGGAGCTAAGCTGAAACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAG  400
 G  A  K  L  K  Q  P  N  E  E  G  D  G  A  K  L  K  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E

GTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCACGACAACCTAATGCAGAAGGTGGAGCAAACGCAGGACAACCGAATGCAGAAGG  500
G  G  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  R  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G

TGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCACGACAGCCACAGGCAGAAGGTGGTGGAGCAAATGCACGACAGCCACAGGCAGAA  600
  G  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  R  Q  P  Q  A  E  G  G  G  A  N  A  R  Q  P  Q  A  E

GGTGGTGGAGCAAATGCACGACAAGGAGGAAATGAGGGGAATAAACAAGCAGGAAAAGGACAGGGACAAAACAATCAGGGTGCGAATGCCCCAAATGAAA  700
 G  G  G  A  N  A  R  Q  G  G  N  E  G  N  K  Q  A  G  K  G  Q  G  Q  N  N  Q  G  A  N  A  P  N  E

AAGTTGTAAATGATTACCTACAGAAAATTAGATCTAGCGTTACCATCGAGTGGACTCCATGCAGTGTAACCTGTGGAAATGGTGTAAGAATTAGAAGAAG  800
K  V  V  N  D  Y  L  Q  K  I  R  S  S  V  T  I  E  W  T  P  C  S  V  T  C  G  N  G  V  R  I  R  R  R

AGCTCATGCAGATAAGAAAAAGGCAGAGGACCTTACTATGGATGACCTTGAAGTGGAAGCTTGTGTAATGGATAAGTGTGCTGGCATATTTAACGTTGTG  900
  A  H  A  D  K  K  K  A  E  D  L  T  M  D  D  L  E  V  E  A  C  V  M  D  K  C  A  G  I  F  N  V  V

AGTAATTCATTAGGGTTAGTAATATTGTTAGTCCTAGCATTATTCAATTAA
 S  N  S  L  G  L  V  I  L  L  V  L  A  L  F  N  .
```

SG/EHI/H1-im/Y09 -95 (*P. knowlesi*; JQ219886)

```
ATGAAGAACTTCATTCTCTTGGCCGTCTCCTCCATCCTGCTGGTGGACTTGCTCCCCACACACTTCGAACATAATGTAGATCTCTCCAGGGCCATAAATG  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  L  P  T  H  F  E  H  N  V  D  L  S  R  A  I  N

TAAATGGAGTAAGCCTCAATAATGTAGACACCAGTTCACTTGGCGCAGCACAGGTAAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAAGCCAAA  200
V  N  G  V  S  L  N  N  V  D  T  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  K  P  K

AGAAGGAGATGATAAAGAAAAGAAAAAAGAAAAAGAAAAAGAAGAAGAACCAAAGAAGCTAAATGAAAATAATCCGAAAAAACCGAATGAAGAAGGTGAT  300
  E  G  D  D  K  E  K  K  K  E  K  E  K  E  E  E  P  K  K  L  N  E  N  N  P  K  K  P  N  E  E  G  D

GGAGCTAAGCTGAAACAACCGAATGCAGAAGGTGGAGCTAAGCTGAAACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTG  400
 G  A  K  L  K  Q  P  N  A  E  G  G  A  K  L  K  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G

GAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGC  500
G  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  G  G  A  N  A  G  Q  P  N  A  E  G  G  A

AAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGTAGGACAACCGAATGCAGAAGGTGGAGCAAATGCACGACAACCGAATGCAGAAGGTGGAGCA  600
  N  A  G  Q  P  N  A  E  G  G  A  N  V  G  Q  P  N  A  E  G  G  A  N  A  R  Q  P  N  A  E  G  G  A

AATGCACGACAACCTAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAA  700
 N  A  R  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G  A

ATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCACGACAACCTAATGCAGAAGGTGGAGCAAACGCAGGACAACCGAATGCAGAAGGTGGAGCAAA  800
N  A  G  Q  P  N  A  E  G  G  A  N  A  R  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G  A  N

TGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCACGACAGCCACAGGCAGAAGGTGGTGGAGCAAATGCACGACAGCCACAGGCAGAAGGTGGTGGA  900
  A  G  Q  P  N  A  E  G  G  A  N  A  R  Q  P  Q  A  E  G  G  G  A  N  A  R  Q  P  Q  A  E  G  G  G

GCAAATGCACGACAAGGAGGAAATGAGGGGAATAAACGAGCAGGAAAAGGACAGGGACAAAACAATCAGGGTGCGAATGCCCCAAATGAAAAAGTTGTAA  1000
 A  N  A  R  Q  G  G  N  E  G  N  K  R  A  G  K  G  Q  G  Q  N  N  Q  G  A  N  A  P  N  E  K  V  V

ATGATTACCTACAGAAAATTAGATCTAGCGTTACCATCGAGTGGACTCCATGCAGTGTAACCTGTGGAAATGGTGTAAGAATTAGAAGAAGAGCTCATGC  1100
N  D  Y  L  Q  K  I  R  S  S  V  T  I  E  W  T  P  C  S  V  T  C  G  N  G  V  R  I  R  R  R  A  H  A

AGATAAGAAAAAGGCAGAGGACCTTACTATGGATGACCTTGAAGTGGAAGCTTGTGTAATGGATAAGTGTGCTGGCATATTTAACGTTGTGAGTAATTCA  1200
  D  K  K  K  A  E  D  L  T  M  D  D  L  E  V  E  A  C  V  M  D  K  C  A  G  I  F  N  V  V  S  N  S

TTAGGGTTAGTAATATTGTTAGTCCTAGCATTATTCAATTAA
 L  G  L  V  I  L  L  V  L  A  L  F  N  .
```

SG/EHI/H1-im/Y09 -102 (*P. knowlesi*; JQ219887)

```
ATGAAGAACTTCATTCTCTTGGCCGTCTCCTCCATCCTGCTGGTGGACTTGCTCCCCACACACTTCGAACATAATGTAGATCTCTCCAGGGCCATAAATG  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  L  P  T  H  F  E  H  N  V  D  L  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAATGTAGACACCAGTTCACTTGGCGCAGCACAGGTAAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAAGCCAAA  200
V  N  G  V  S  F  N  N  V  D  T  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  K  P  K

AGAAGGAGATGATAAAGAAAAGAAAAAAGAAAAAGAAAAAGAAGAAGAACCAAAGAAGCTAAATGAAAATAATCCGAAAAAACCGAATGAAGAAGGTGAT  300
  E  G  D  D  K  E  K  K  K  E  K  E  K  E  E  E  P  K  K  L  N  E  N  N  P  K  K  P  N  E  E  G  D

GGAGCTAAGCTGAAACAACCGAATGAAGAAGGTGATGGAGCTAAGCTGAAACAACCGAATGCAGAAGGTGGAGCTAAGCTGAAACAACCGAATGCAGAAG  400
 G  A  K  L  K  Q  P  N  E  E  G  D  G  A  K  L  K  Q  P  N  A  E  G  G  A  K  L  K  Q  P  N  A  E

GTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGGTGG  500
G  G  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  G  G

AGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGA  600
  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G

GCAAATGCACGACAACCGAATGCAGAAGGTGGAGCAAATGCACGACAACCTAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAG  700
 A  N  A  R  Q  P  N  A  E  G  G  A  N  A  R  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G

CAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCACGACAACCTAATGCAGAAGGTGGAGC  800
A  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  R  Q  P  N  A  E  G  G  A

AAACGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCACGACAGCCACAGGCAGAAGGTGGTGGA  900
  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  R  Q  P  Q  A  E  G  G  G

GCAAATGCACGACAGCCACAGGCAGAAGGTGGTGGAGCAAATGCACGACAAGGAGGAAATGAGGGGAATAAACAAGCAGGAAAAGGACAGGGACAAAACA  1000
 A  N  A  R  Q  P  Q  A  E  G  G  G  A  N  A  R  Q  G  G  N  E  G  N  K  Q  A  G  K  G  Q  G  Q  N

ATCAGGGTGCGAATGCCCCAAATGAAAAAGTTGTAAATGATTACCTACAGAAAATTAGATCTAGCGTTACCATCGAGTGGACTCCATGCAGTGTAACCTG  1100
N  Q  G  A  N  A  P  N  E  K  V  V  N  D  Y  L  Q  K  I  R  S  S  V  T  I  E  W  T  P  C  S  V  T  C

TGGAAATGGTGTAAGAATTAGAAGAAGAGCTCATGCAGATAAGAAAAAGGCAGAGGACCTTACTATGGATGACCTTGAAGTGGAAGCTTGTGTAATGGAT  1200
  G  N  G  V  R  I  R  R  R  A  H  A  D  K  K  K  A  E  D  L  T  M  D  D  L  E  V  E  A  C  V  M  D

AGGTGTGCTGGCATATTTAACGTTGTGAGTAATTCATTAGGGTTAGTAATATTGTTGGTCCTAGCATTATTCAATTAA
 R  C  A  G  I  F  N  V  V  S  N  S  L  G  L  V  I  L  L  V  L  A  L  F  N  .
```

SG/EHI/H2-im/Y09 -3 (*P. knowlesi*; JQ219888)

```
ATGAAGAACTTCATTCTCTTGGCCGTCTCCTCCATCCTGCTGGTGGACTTGCTCCCCACACACTTCGAACATAATGTAGATCTCTCCAGGGCCATAAATG  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  L  P  T  H  F  E  H  N  V  D  L  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAATGTAGACACCAGTTCACTTGGCGCAGCACAGGTAAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAAGCCAAA  200
V  N  G  V  S  F  N  N  V  D  T  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  K  P  K

AGAAGGAGATGATAAAGAAAAGAAAAAAGAAAAAGAAAAAGAAGAAGAACCAAAGAAGCTAAATGAAAATAATCCGAAAAAACCGAATGAAGAAGGTGAT  300
  E  G  D  D  K  E  K  K  K  E  K  E  K  E  E  E  P  K  K  L  N  E  N  N  P  K  K  P  N  E  E  G  D

GGAGCTAAGCTGAAACAACCGAATGAAGAAGGTGATGGAGCTAAGCTGAAACAACCGAATGCAGAAGGTGGAGCTAAGCTGAAACAACCGAATGCAGAAG  400
 G  A  K  L  K  Q  P  N  E  E  G  D  G  A  K  L  K  Q  P  N  A  E  G  G  A  K  L  K  Q  P  N  A  E

GTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGGTGG  500
G  G  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  G  G

AGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGA  600
  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G

GCAAATGCACGACAACCGAATGCAGAAGGTGGAGCAAATGCACGACAACCTAATGCAGGAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAG  700
 A  N  A  R  Q  P  N  A  E  G  G  A  N  A  R  Q  P  N  A  G  G  G  A  N  A  G  Q  P  N  A  E  G  G

CAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCACGACAGCCTAATGCAGAAGGTGGAGC  800
A  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  R  Q  P  N  A  E  G  G  A

AAACGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCACGACAGCCACAGGCAGAAGGTGGTGGA  900
  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  R  Q  P  Q  A  E  G  G  G

GCAAATGCACGACAGCCACAGGCAGAAGGTGGTGGAGCAAATGCACGACAAGGAGGAAATGAGGGGAATAAACAAGCAGGAAAAGGACAGGGACAAAACA  1000
 A  N  A  R  Q  P  Q  A  E  G  G  G  A  N  A  R  Q  G  G  N  E  G  N  K  Q  A  G  K  G  Q  G  Q  N

ATCAGGGTGCGAATGCCCCAAATGAAAAAGTTGTAAATGATTACCTACAGAAAATTAGATCTAGCGTTACCATCGAGTGGACTCCATGCAGTGTAACCTG  1100
N  Q  G  A  N  A  P  N  E  K  V  V  N  D  Y  L  Q  K  I  R  S  S  V  T  I  E  W  T  P  C  S  V  T  C

TGGAAATGGTGTAAGAATTAGAAGAAGAGCTCATGCAGATAAGAAAAAGGCAGAGGACCTTACTGTGGATGACCTTGAAGTGGAAGCTTGTGTAATGGAT  1200
  G  N  G  V  R  I  R  R  R  A  H  A  D  K  K  K  A  E  D  L  T  V  D  D  L  E  V  E  A  C  V  M  D

AAGTGTGCTGGCATATTTAACGTTGTGAGTAATTCATTAGGGTTAGTAATATTGTTAGTCCTAGCATTATTCAATTAA
 K  C  A  G  I  F  N  V  V  S  N  S  L  G  L  V  I  L  L  V  L  A  L  F  N  .
```

SG/EHI/H2-im/Y09 -7 (*P. knowlesi*; JQ219889)

```
ATGAGGAACTTCATTCTCTTGGCCGTCTCCTCCATCCTGCTGGTGGACTTGCTCCCCACACACTTCGAACATAATGTAGATCTCTCCAGGGCCATAAATG  100
 M  R  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  L  P  T  H  F  E  H  N  V  D  L  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAATGTAGACACCAGTTCACTTGGCGCAGCACAGGTAAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAAGCCAAA  200
V  N  G  V  S  F  N  N  V  D  T  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  K  P  K

AGAAGGAGATGATAAAGAAAAGAAAAAAGAAAAAGAAAAAGAAGAAGAACCAAAGAAGCTAAATGAAAATAATCCGAAAAAACCGAATGAAGAAGGTGAT  300
  E  G  D  D  K  E  K  K  K  E  K  E  K  E  E  E  P  K  K  L  N  E  N  N  P  K  K  P  N  E  E  G  D

GGAGCTAAGCTGAAACAACCGAATGAAGAAGGTGATGGAGCTAAGCTGAAACAACCGAATGCAGAAGGTGGAGCTAAGCTGAAACAACCGAATGCAGAAG  400
 G  A  K  L  K  Q  P  N  E  E  G  D  G  A  K  L  K  Q  P  N  A  E  G  G  A  K  L  K  Q  P  N  A  E

GTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGGTGG  500
G  G  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  G  G

AGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGA  600
  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G

GCAAATGCACGACAACCGAATGCAGAAGGTGGAGCAAATGCACGACAACCTAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAG  700
 A  N  A  R  Q  P  N  A  E  G  G  A  N  A  R  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G

CAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCACGACAACCTAATGCAGAAGGTGGAGC  800
A  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  R  Q  P  N  A  E  G  G  A

AAACGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCACGACAGCCACAGGCAGAAGGTGGTGGA  900
  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  R  Q  P  Q  A  E  G  G  G

GCAAATGCACGACAGCCACAGGCAGAAGGTGGTGGAGCAAATGCACGACAAGGAGGAAATGAGGGGAATAAACAAGCAGGAAAAGGACAGGGACAAAACA  1000
 A  N  A  R  Q  P  Q  A  E  G  G  G  A  N  A  R  Q  G  G  N  E  G  N  K  Q  A  G  K  G  Q  G  Q  N

ATCAGGGTGCGAATGCCCCAAATGAAAAAGTTGTAAATGATTACCTACAGAAAATTAGATCTAGCGTTACCATCGAGTGGACTCCATGCAGTGTAACCTG  1100
N  Q  G  A  N  A  P  N  E  K  V  V  N  D  Y  L  Q  K  I  R  S  S  V  T  I  E  W  T  P  C  S  V  T  C

TGGAAATGGTGTAAGAATTAGAAGAAGAGCTCATGCAGATAAGAAAAAGGCAGAGGACCTTACTATGGATGACCTTGAAGTGGAAGCTTGTGTAATGGAT  1200
  G  N  G  V  R  I  R  R  R  A  H  A  D  K  K  K  A  E  D  L  T  M  D  D  L  E  V  E  A  C  V  M  D

AAGTGTGCTGGCATATTTAACGTTGTGAGTAATACATTAGGGTTAGTAATATTGTTAGTCCTAGCATTATTCAATTAA
 K  C  A  G  I  F  N  V  V  S  N  T  L  G  L  V  I  L  L  V  L  A  L  F  N  .
```

SG/EHI/H2-im/Y09 -11 (*P. knowlesi*; JQ219890)

```
ATGAAGAACTTCATTCTCTTGGCCGTCTCCTCCATCCTGCTGGTGGACTTGCTCCCCACACACTTCGAACATAATGTAGATCTCTCCAGGGCCATAAATG  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  L  P  T  H  F  E  H  N  V  D  L  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAATGTGGACACCAGTTCACTTGGCGCAGCACAGGTAAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAAGCCAAA  200
V  N  G  V  S  F  N  N  V  D  T  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  K  P  K

AGAAGGAGATGATAAAGAAAAGAAAAAAGAAAAAGAAAAAGAAGAAGAACCAAAGAAGCTAAATGAAAATAATCCGAAAAAACCGAATGAAGAAGGTGAT  300
  E  G  D  D  K  E  K  K  K  E  K  E  K  E  E  E  P  K  K  L  N  E  N  N  P  K  K  P  N  E  E  G  D

GGAGCTAAGCTGAAACAACCGAATGAAGAAGGTGATGGAGCTAAGCTGAAACAACCGAATGCAGAAGGTGGAGCTAAGCTGAAACAACCGAATGCAGAAG  400
 G  A  K  L  K  Q  P  N  E  E  G  D  G  A  K  L  K  Q  P  N  A  E  G  G  A  K  L  K  Q  P  N  A  E

GTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGGTGG  500
G  G  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  G  G

GGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGA  600
  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G

GCAAATGCACGACAACCGAATGCAGAAGGTGGAGCAAATGCACGACAACCTAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAG  700
 A  N  A  R  Q  P  N  A  E  G  G  A  N  A  R  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G

CAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCACGACAACCTAATGCAGAAGGTGGAGC  800
A  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  R  Q  P  N  A  E  G  G  A

AAACGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCACGACAGCCACAGGCAGAAGGTGGTGGA  900
  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  R  Q  P  Q  A  E  G  G  G

GCAAATGCACGACAGCCACAGGCAGAAGGTGGTGGAGCAAATGCACGACAAGGAGGAAATGAGGGGAATAAACAAGCAGGAAAAGGACAGGGACAGAACA  1000
 A  N  A  R  Q  P  Q  A  E  G  G  G  A  N  A  R  Q  G  G  N  E  G  N  K  Q  A  G  K  G  Q  G  Q  N

ATCAGGGTGCGAATGCCCCAAATGAAAAAGTTGTAAATGATTACCTACAGAAAATTAGATCTAGCGTTACCATCGAGTGGACTCCATGCAGTGTAACCTG  1100
N  Q  G  A  N  A  P  N  E  K  V  V  N  D  Y  L  Q  K  I  R  S  S  V  T  I  E  W  T  P  C  S  V  T  C

TGGAAATGGTGTAAGAATTAGAAGAAGAGCTCATGCAGATAAGAAAAGGGCAGAGGACCTTACTATGGATGACCTTGAGGTGGAAGCTTGTGTAATGGAT  1200
  G  N  G  V  R  I  R  R  R  A  H  A  D  K  K  R  A  E  D  L  T  M  D  D  L  E  V  E  A  C  V  M  D

AAGTGTGCTGGCATATTTAACGTTGTGAGTAATTCATTAGGGTTAGTAATATTGTTAGTCCTAGCATTATTCAATTAA
 K  C  A  G  I  F  N  V  V  S  N  S  L  G  L  V  I  L  L  V  L  A  L  F  N  .
```

SG/EHI/H2-im/Y09 -12 (*P. knowlesi*; JQ219891)

```
ATGAAGAACTTCATTCTCTTGGCCGTCTCCTCCATCCTGCTGGTGGACTTGCTCCCCACACACTTCGAACATAATGTAGATCTCTCCAGGGCCATAAATG  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  L  P  T  H  F  E  H  N  V  D  L  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAATGTAGACACCAGTTCACTTGGCGCAGCACAGGTAAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAAGCCAAA  200
V  N  G  V  S  F  N  N  V  D  T  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  K  P  K

AGAAGGAGATGATAAAGAAAAGAAAAAAGAAAAAGAAAAAGAAGAAGAACCAAAGAAGCTAAATGAAAATAATCCGAAAAAACCGAATGAAGAAGGTGAT  300
  E  G  D  D  K  E  K  K  K  E  K  E  K  E  E  E  P  K  K  L  N  E  N  N  P  K  K  P  N  E  E  G  D

GGAGCTAAGCTGAAACAACCGAATGAAGAAGGTGATGGAGCTAAGCTGAAACAACCGAATGCAGAAGGTGGAGCTAAGCTGAAACAACCGAATGCAGAAG  400
 G  A  K  L  K  Q  P  N  E  E  G  D  G  A  K  L  K  Q  P  N  A  E  G  G  A  K  L  K  Q  P  N  A  E

GTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGGTGG  500
G  G  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  G  G

AGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGA  600
  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G

GCAAATGCACGACAACCGAATGCAGAAGGTGGAGCAAATGCACGACAACCTAATGCAGAGGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAG  700
 A  N  A  R  Q  P  N  A  E  G  G  A  N  A  R  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G

CAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCACGACAACCTAATGCAGAAGGTGGAGC  800
A  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  R  Q  P  N  A  E  G  G  A

AAACGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCA  900
  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G  A

AATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCACGACAACCTAATGCAGAAGGTGGAGCAAACGCAGGACAACCGAATGCAGAAGGTGGAGCAA  1000
 N  A  G  Q  P  N  A  E  G  G  A  N  A  R  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G  A

ATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCACGACAGCCACAGGCAGAAGGTGGTGGAGCAAATGCACGACAGCCACAGGCAGAAGGTGGTGG  1100
N  A  G  Q  P  N  A  E  G  G  A  N  A  R  Q  P  Q  A  E  G  G  G  A  N  A  R  Q  P  Q  A  E  G  G  G

AGCAAATGCACGACAAGGAGGAAATGAGGGGAATAAACAAGCAGGAAAAGGACAGGGACAAAACAATCAGGGTGCGAATGCCCCAAATGAAAAAGTTGTA  1200
  A  N  A  R  Q  G  G  N  E  G  N  K  Q  A  G  K  G  Q  G  Q  N  N  Q  G  A  N  A  P  N  E  K  V  V

AATGATTACCTACAGAAAATTAGATCTAGCGTTACCATCGAGTGGACTCCATGCAGTGTAACCTGTGGAAATGGTGTAAGAATTAGAAGAAGAGCTCATG  1300
 N  D  Y  L  Q  K  I  R  S  S  V  T  I  E  W  T  P  C  S  V  T  C  G  N  G  V  R  I  R  R  R  A  H

CAGATAAGAAAAAGGCAGAGGACCTTACTATGGATGACCTTGAAGTGGAAGCTTGTGTAATGGATAAGTGTGCTGGCATATTTAACGTTGTGAGTAATTC  1400
A  D  K  K  K  A  E  D  L  T  M  D  D  L  E  V  E  A  C  V  M  D  K  C  A  G  I  F  N  V  V  S  N  S

ATTAGGGTTAGTAATATTGTTAGTCCTAGCATTATTCAATTAA
  L  G  L  V  I  L  L  V  L  A  L  F  N  .
```

SG/EHI/H2-im/Y09 -18 (*P. knowlesi*; JQ219892)

```
ATGAAGAACTTCATTCTCTTGGCCGTCTCCTCCATCCTGCTGGTGGACTTGCTCCCCACACACTTCGAACATAATGTAGATCTCACCAGGGCCATAAATG  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  L  P  T  H  F  E  H  N  V  D  L  T  R  A  I  N

TAAATGGAGTAAGCTTCAATAATGTAGACACCAGTTCACTTGGCGCAGCACAGGTAAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAAGCCAAA  200
V  N  G  V  S  F  N  N  V  D  T  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  K  P  K

AGAAGGAGATGATAAAGAAAAGAAAAAAGAAAAAGAAAAAGAAGAAGAACCAAAGAAGCTAAATGAAAATAATCCGAAAAAACCGAATGAAGAAGGTGAT  300
  E  G  D  D  K  E  K  K  K  E  K  E  K  E  E  E  P  K  K  L  N  E  N  N  P  K  K  P  N  E  E  G  D

GGAGCTAAGCTGAAACAACCGAATGAAGAAGGTGATGGAGCTAAGCTGAAACAACCGAATGCAGAAGGTGGAGCTAAGCTGAAACAACCGAATGCAGAAG  400
 G  A  K  L  K  Q  P  N  E  E  G  D  G  A  K  L  K  Q  P  N  A  E  G  G  A  K  L  K  Q  P  N  A  E

GTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGGTGG  500
G  G  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  G  G

AGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGA  600
  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G

GCAAATGCACGACAACCGAATGCAGAAGGTGGAGCAAATGCACGACAACCTAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAG  700
 A  N  A  R  Q  P  N  A  E  G  G  A  N  A  R  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G

CAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCACGACAACCTAATGCAGAAGGTGGAGC  800
A  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  R  Q  P  N  A  E  G  G  A

AAACGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCAGGACAACCGAATGCAGAAGGTGGAGCAAATGCACGACAGCCACAGGCAGAAGGTGGTGGA  900
  N  A  G  Q  P  N  A  E  G  G  A  N  A  G  Q  P  N  A  E  G  G  A  N  A  R  Q  P  Q  A  E  G  G  G

GCAAATGCACGACAGCCACAGGCAGAAGGTGGTGGAGCAAATGCACGACAAGGAGGAAATGAGGGGAATAAATAAGCAGGAAAAGGACAGGGACAAAACA  1000
 A  N  A  R  Q  P  Q  A  E  G  G  G  A  N  A  R  Q  G  G  N  E  G  N  K  .  A  G  K  G  Q  G  Q  N

ATCAGGGTGCGAATGCCCCAAATGAAAAAGTTGTAAATGATTACCTACAGAAAATTAGATCTAGCGTTACCATCGAGTGGACTCCATGCAGTGTAACCTG  1100
N  Q  G  A  N  A  P  N  E  K  V  V  N  D  Y  L  Q  K  I  R  S  S  V  T  I  E  W  T  P  C  S  V  T  C

TGGAAATGGTGTAAGAATTAGAAGAAGAGCTCATGCAGATAAGAAAAAGGCAGAGGACCTTACTATGGATGACCTTGAAGTGGAAGCTTGTGTAATGGAT  1200
  G  N  G  V  R  I  R  R  R  A  H  A  D  K  K  K  A  E  D  L  T  M  D  D  L  E  V  E  A  C  V  M  D

AAGTGTGCTGGCATATTTAACGTTGTGAGTAATTCATTAGGGTTAGTAATATTGTTAGTCCTAGCATTATTCAATTAA
 K  C  A  G  I  F  N  V  V  S  N  S  L  G  L  V  I  L  L  V  L  A  L  F  N  .
```

SG/EHI/WM01/Y07-6 (*P. knowlesi*; JQ219897)

```
ATGAAGAACTTCATTCTCTTGGCCGTCTCCTCCATCCTGCTGGTGGACTTGCTCCCCACACACTTCGAACATAATGTAGATCTCTCCAGGGCCATAAATG  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  L  P  T  H  F  E  H  N  V  D  L  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAATGTAGACACCAGTTCACTTGGCGCAGCACAGGTGAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAAGCCAAA  200
V  N  G  V  S  F  N  N  V  D  T  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  K  P  K

AGAAGGAGCTGATAAAGAAAAGAAAAAAGAAAAAGGAAAAGAAAAAGAAGAAGAACCAAAGAAGCCAAATGAAAATAAGCTGAAACAACCGAATGAAGGA  300
  E  G  A  D  K  E  K  K  K  E  K  G  K  E  K  E  E  E  P  K  K  P  N  E  N  K  L  K  Q  P  N  E  G

CAACCACAAGCACAGGGTGATGGAGCAAATGCAGGACAACCACAAGCACAAGGAGATGGAGCAAATGCAGGACAACCACAAGCACAGGGTGATGGAGCAA  400
 Q  P  Q  A  Q  G  D  G  A  N  A  G  Q  P  Q  A  Q  G  D  G  A  N  A  G  Q  P  Q  A  Q  G  D  G  A

ATGCAGGACAACCACAAGCACAGGGTGATGGAGCAAATGCAGGACAACCACAAGCACAAGGAGATGGAGCAAATGCAGGACAACCACAAGCACAGGGTGA  500
N  A  G  Q  P  Q  A  Q  G  D  G  A  N  A  G  Q  P  Q  A  Q  G  D  G  A  N  A  G  Q  P  Q  A  Q  G  D

TGGAGCAAATGCAGGGCAACCACAAGCACAGGGTGATGGAGCAAATGCAGGACAACCACAAGCACAAGGAGATGGAGCAAATGCAGGACAACCACAAGCA  600
  G  A  N  A  G  Q  P  Q  A  Q  G  D  G  A  N  A  G  Q  P  Q  A  Q  G  D  G  A  N  A  G  Q  P  Q  A

CAAGGAGATGGAGCAAATGCAGGACAACCACAAGCACAGGGTGATGGAGCAAATGCAGGACAACCACAAGCACAGGGTGATAGGGCGAATGCAGGACAAC  700
 Q  G  D  G  A  N  A  G  Q  P  Q  A  Q  G  D  G  A  N  A  G  Q  P  Q  A  Q  G  D  R  A  N  A  G  Q

CACAAGCACAAGGAGATGGGGCAAATGTACCACGACAAGGAAGAAACGGGGGAGGTGCACCAGCAGGAGGAAATGAGGGGAATAAACAAGCAGGAAAAGG  800
P  Q  A  Q  G  D  G  A  N  V  P  R  Q  G  R  N  G  G  G  A  P  A  G  G  N  E  G  N  K  Q  A  G  K  G

ACAGGGACAAAACAATCAGGGTGCGAATGCCCCAAATGAAAAAGTTGTGAATGATTACCTACACAAAATTAGATCTAGCGTTACCACCGAGTGGACTCCA  900
  Q  G  Q  N  N  Q  G  A  N  A  P  N  E  K  V  V  N  D  Y  L  H  K  I  R  S  S  V  T  T  E  W  T  P

TGCAGTGTAACCTGTGGAAATGGTGTAAGAATTAGAAGAAAAGCTCATGCAGGTAATAAAAAGGCAGAGGACCTTACTATGGATGACCTTGAGGTGGAAG  1000
 C  S  V  T  C  G  N  G  V  R  I  R  R  K  A  H  A  G  N  K  K  A  E  D  L  T  M  D  D  L  E  V  E

CTTGTGTAATGGATAAGTGCGCTGGCATATTTAACGTTGTGAGTAATTCATTAGGCTTAGTCATATTGTTAGTCCTAGCATTATTCAATTAA
A  C  V  M  D  K  C  A  G  I  F  N  V  V  S  N  S  L  G  L  V  I  L  L  V  L  A  L  F  N  .
```

SG/EHI/WM01/Y07-23 (*P. fieldi*; JQ219931)

```
ATGAAGAACTTCATTCTCTTAGCCGTTTCTTCCATCCTGTTGGTGGACTTGTTCCCCACACACTGCGGGCATAATGTAGATCTCTCCAGGGACATAAATT  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  F  P  T  H  C  G  H  N  V  D  L  S  R  D  I  N

TAAATGGAGCAAGCTTCAATAATGTAGACGCCAGTTCACTTGGCGCAGCACAGGTAAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAACCCAAA  200
L  N  G  A  S  F  N  N  V  D  A  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  N  P  K

AGATGAAGGAGCTGCTAAAGAAAAAAAGGACGAAAAAAAAGTAGAACCAAAAAAGGCACGTGAAAATAAGCTGAAACAACCAGCTAATGATGCAGGACAA  300
  D  E  G  A  A  K  E  K  K  D  E  K  K  V  E  P  K  K  A  R  E  N  K  L  K  Q  P  A  N  D  A  G  Q

AATCAGCCAGCTAATGATGCAGGACAAAATCAGCCAGCTAATGACGCAGGACAAAATCAGCCAGCTAATGATGCAGGACAAAATCAGCCAGCTAATGATG  400
 N  Q  P  A  N  D  A  G  Q  N  Q  P  A  N  D  A  G  Q  N  Q  P  A  N  D  A  G  Q  N  Q  P  A  N  D

CAGGACAAAATCAGCCAGCTAATGATGCAGGACAAAATCAGCCAGCTAATGATGCAGGACAAAATCAGCCAGCTAATGATGCAGGACAAAATCAGCCAGC  500
A  G  Q  N  Q  P  A  N  D  A  G  Q  N  Q  P  A  N  D  A  G  Q  N  Q  P  A  N  D  A  G  Q  N  Q  P  A

TAATGATGCAGGACAAAATCAACCAGCTAATGATGCAGGACAAAATCAGCCAGCTAATGATGCAGGACAAAATCAACCAGCTAATGATGCAGGACAAAAT  600
  N  D  A  G  Q  N  Q  P  A  N  D  A  G  Q  N  Q  P  A  N  D  A  G  Q  N  Q  P  A  N  D  A  G  Q  N

CAGCCAGCTAATGATGCAGGACAAAATCAGCCAGGTGGTGGAGCAGGACAAAATCAGCCAGGTGGTGGAGCAGGACAAAATCAACCAGGTGGTGGAGCAG  700
 Q  P  A  N  D  A  G  Q  N  Q  P  G  G  G  A  G  Q  N  Q  P  G  G  G  A  G  Q  N  Q  P  G  G  G  A

GACAAAATCAACCAGGTGGTGGAGCAGGACAAAATCAACCAGGTGGTGGAGCAGGACAAAATCAGCCAGGAAATGGAGCAGGACAAAATCAGCCAGGAGA  800
G  Q  N  Q  P  G  G  G  A  G  Q  N  Q  P  G  G  G  A  G  Q  N  Q  P  G  N  G  A  G  Q  N  Q  P  G  D

TGGAGCAGGACGAAATGGTGGAAACGCAGGAGCAGGAGGACAGGGACAAAATAATGAAGGTGCGAATAAGCCAGATGAAAAGCATGTGAAAGAATACCTA  900
  G  A  G  R  N  G  G  N  A  G  A  G  G  Q  G  Q  N  N  E  G  A  N  K  P  D  E  K  H  V  K  E  Y  L

GAGAAAATTAGATCTACCGTTGGCACCGAATGGACTCCATGCAGTGTAACCTGTGGAAAGGGTGTAAGAGTTAGAAGAAAACTTAGTGCAGGTGACAAAA  1000
 E  K  I  R  S  T  V  G  T  E  W  T  P  C  S  V  T  C  G  K  G  V  R  V  R  R  K  L  S  A  G  D  K

AACCAGATAAGCTTACTCTGAATGACCTTGAGGCAGAAGTTTGTACAATGGATAAGTGCGCTGGCATATTTAACGTTGTGAGTAATTCATTAGGGCTAGT  1100
K  P  D  K  L  T  L  N  D  L  E  A  E  V  C  T  M  D  K  C  A  G  I  F  N  V  V  S  N  S  L  G  L  V

CATATTGTTAGTCCTAGCATTATTCAATTAA
  I  L  L  V  L  A  L  F  N  .
```

SG/EHI/WM02/Y07-1 (*P. knowlesi*; JQ219898)

```
ATGAAGAACTTCATTCTCTTGGCCGTCTCCTCCATCCTGCTGGTGGACTTGCTCCCCACACACTTCGAACATAATGTAGATCTCTCCAGGGCCATAAATG  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  L  P  T  H  F  E  H  N  V  D  L  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAATGTAGACACCAGTTCACTTGGCGCAGCACAGGTGAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAAGCCAAA  200
V  N  G  V  S  F  N  N  V  D  T  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  K  P  K

AGAAGGAGCTGATAAAGAAAAGAAAAAAGAAAAAGAAAAAGAAAAAGAAGAACCAAAGAAGCCAAATGAAAATAAGCTGAAACAACCGGAACAACCAGCA  300
  E  G  A  D  K  E  K  K  K  E  K  E  K  E  K  E  E  P  K  K  P  N  E  N  K  L  K  Q  P  E  Q  P  A

GCAGGAGCAGGGGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAAC  400
 A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E

AACCAGCGGCAGGAGCAGGGGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCAGCAGGAGCAGG  500
Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G

AGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCAGCAGGAGCAGGGGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCACCA  600
  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  P

GCACCAAGGAGGGAACAACCAGCAGCAGGAGCAGGGGGCGAACAACCAGCACCAGCACCAAGGAGGGAACAACCAGCACCAGGAGCAGGTGCGGGAGATG  700
 A  P  R  R  E  Q  P  A  A  G  A  G  G  E  Q  P  A  P  A  P  R  R  E  Q  P  A  P  G  A  G  A  G  D

GAGCACGAGGAGGAAACGCAGGGGCAGGTAAAGGACAGGGACAAAACAATCAGGGTGCGAATGTCCCAAATGAAAAAGTTGTGAATGATTACCTACACAA  800
G  A  R  G  G  N  A  G  A  G  K  G  Q  G  Q  N  N  Q  G  A  N  V  P  N  E  K  V  V  N  D  Y  L  H  K

AATTAGATCTAGCGTTACCACCGAGTGGACTCCATGCAGTGTAACCTGTGGAAATGGTGTAAGAATTAGAAGAAGACAGAATGCTGGTAATAAAAAGGCA  900
  I  R  S  S  V  T  T  E  W  T  P  C  S  V  T  C  G  N  G  V  R  I  R  R  R  Q  N  A  G  N  K  K  A

GAGGACCTTACTATGGATGACCTTGAGGTGGAAGCTTGTGTAATGGATAAGTGCGCTGGCATATTTAACGTTGTGAGTAATTCATTAGGCTTAGTCATAT  1000
 E  D  L  T  M  D  D  L  E  V  E  A  C  V  M  D  K  C  A  G  I  F  N  V  V  S  N  S  L  G  L  V  I

TGTTAGTCTTAGCATTATTCAATTAA
L  L  V  L  A  L  F  N  .
```

SG/EHI/WM02/Y07-39 (*P. knowlesi*; JQ219899)

```
ATGAAGAACTTCATTCTCTTGGCCGTCTCCTCCATCCTGCTGGTGGACTTGCTCCCCACACACTTCGAACATAATGTAGATCTCTCCAGGGCCATAAATG  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  L  P  T  H  F  E  H  N  V  D  L  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAATGTAGACACCAGTTCACTTGGCGCAGCACAGGTGAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAAGCCAAA  200
V  N  G  V  S  F  N  N  V  D  T  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  K  P  K

AGAAGGAGCTGATAAAGAAAAGAAAAAAGAAAAAGGAAAAGAAAAAGAAGAAGAACCAAAGAAGCCAAATGAAAATAAGCTGAAACAACCGAATGAAGGA  300
  E  G  A  D  K  E  K  K  K  E  K  G  K  E  K  E  E  E  P  K  K  P  N  E  N  K  L  K  Q  P  N  E  G

CAACCACAAGCACAGGGTGATGGAGCAAATGCAGGACAACCACAAGCACAAGGAGATGGAGCAAATGCAGGACAACCACAAGCACAGGGTGATGGAGCAA  400
 Q  P  Q  A  Q  G  D  G  A  N  A  G  Q  P  Q  A  Q  G  D  G  A  N  A  G  Q  P  Q  A  Q  G  D  G  A

ATGCAGGACAACCACAAGCACAGGGTGATGGAGCAAATGCAGGACAACCACAAGCACAAGGAGATGGAGCAAATGCAGGACAACCACAAGCACAGGGTGA  500
N  A  G  Q  P  Q  A  Q  G  D  G  A  N  A  G  Q  P  Q  A  Q  G  D  G  A  N  A  G  Q  P  Q  A  Q  G  D

TGGAGCAAATGCAGGGCAACCACAAGCACAGGGTGATGGAGCAAATGCAGGACAACCACAAGCACAAGGAGATGGAGCAAATGCAGGACAACCACAAGCA  600
  G  A  N  A  G  Q  P  Q  A  Q  G  D  G  A  N  A  G  Q  P  Q  A  Q  G  D  G  A  N  A  G  Q  P  Q  A

CAAGGAGATGGAGCAAATGCAGGACAACCACAAGCACAGGGTGATGGAGCAAATGCAGGACAACCACAAGCACAGGGTGATAGGGCGAATGCAGGACAAC  700
 Q  G  D  G  A  N  A  G  Q  P  Q  A  Q  G  D  G  A  N  A  G  Q  P  Q  A  Q  G  D  R  A  N  A  G  Q

CACAAGCACAAGGAGATGGGGCAAATGTACCACGACAAGGAAGAAACGGGGGAGGTGCACCAGCAGGAGGAAATGAGGGGAATAAACAAGCAGGAAAAGG  800
P  Q  A  Q  G  D  G  A  N  V  P  R  Q  G  R  N  G  G  G  A  P  A  G  G  N  E  G  N  K  Q  A  G  K  G

ACAGGGACAAAACAATCAGGGTGCGAATGCCCCAAATGAAAAAGTTGTGAATGATTACCTACACAAAATTAGATCTAGCGTTACCACCGAGTGGACTCCA  900
  Q  G  Q  N  N  Q  G  A  N  A  P  N  E  K  V  V  N  D  Y  L  H  K  I  R  S  S  V  T  T  E  W  T  P

TGCAGTGTAACCTGTGGAAATGGTGTAAGAATTAGAAGAAAAGCTCATGCAGGTAATAAAAAGGCAGAGGACCTTACTATGGATGACCTTGAGGTGGAAG  1000
 C  S  V  T  C  G  N  G  V  R  I  R  R  K  A  H  A  G  N  K  K  A  E  D  L  T  M  D  D  L  E  V  E

CTTGTGTAATGGATAAGTGCGCTGGCATATTTAACGTTGTGAGTAATTCATTAGGCTTAGTCATATTGTTAGTCCTAGCATTATTCAATTAA
A  C  V  M  D  K  C  A  G  I  F  N  V  V  S  N  S  L  G  L  V  I  L  L  V  L  A  L  F  N  .
```

SG/EHI/WM02/Y07-110 (*P. cynomolgi*; JQ219922)

```
ATGAAGAACTTCATTCTCTTAGCCGTTTCTTCCATCCTGTTGGTGGACTTGTTTCCCACAAACTGCGGGCATAATGTAGATTTCTCCAGGGCCATAAATT  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  F  P  T  N  C  G  H  N  V  D  F  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAATGTAGACGCCAGTTCCCTTGGCGCAGCACAGGTAAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAACCCAAA  200
L  N  G  V  S  F  N  N  V  D  A  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  N  P  K

AAACGAGGAAGGAGCTGATAAACAAAAAAAGGACGaAAAAAAAGTAGaACCAAAAAAGCCaCGTGAAAATAAGCTGAAACAACCAGACGGAAATAATGCA  300
  N  E  E  G  A  D  K  Q  K  K  D  E  K  K  V  E  P  K  K  P  R  E  N  K  L  K  Q  P  D  G  N  N  A

GCTGATGGAGGTGTACAACCGCCAGCAGGAGGAGGAAATAATGCAGCTGATGGAGGTGTACAACCACCAGCAGGAGGAGGAAATAATGCAGCTGATGGAG  400
 A  D  G  G  V  Q  P  P  A  G  G  G  N  N  A  A  D  G  G  V  Q  P  P  A  G  G  G  N  N  A  A  D  G

GTGTACAACCACCAGCAGGAGGAGGAAATAATGCAGCTGATGGAGGTGTACAACCACCAGCAGGAGGAGGAAATAATGCAGCTGATGGAGGTGTACAACC  500
G  V  Q  P  P  A  G  G  G  N  N  A  A  D  G  G  V  Q  P  P  A  G  G  G  N  N  A  A  D  G  G  V  Q  P

ACCAGCAGGAGGAGGAAATAATGCAGCTGATGGAGGCGTACAACCACCAGCAGGAGGAGGAAATAATGCAGCTGATGGAGGTGTACAACCGCCAGCAGGA  600
  P  A  G  G  G  N  N  A  A  D  G  G  V  Q  P  P  A  G  G  G  N  N  A  A  D  G  G  V  Q  P  P  A  G

GGAGGGAATAATGCAGCTGATGGAGGTGTACAACCACCAGCAGGAGGAGGAAATAATGCAGCTGATGGAGGTGTACAACCACCAGCAAGAGGAGGAAATA  700
 G  G  N  N  A  A  D  G  G  V  Q  P  P  A  G  G  G  N  N  A  A  D  G  G  V  Q  P  P  A  R  G  G  N

ATGCAGCTGATGGAGGTGCACAACCGCCAGCAGGAGGAGGAAATAATGCAGCTGATGGAGGTGTACAACCACCAGCAGCAGGAGGAAATAGGGCAAATAA  800
N  A  A  D  G  G  A  Q  P  P  A  G  G  G  N  N  A  A  D  G  G  V  Q  P  P  A  A  G  G  N  R  A  N  K

AAAAGCAGGARAAGCAGGKGGAAACSCAGGASCAGGACAGGGACAAAATAATGAARGTGCGAATATGCCAAATGTAAAGCTTGTGCAAGAATACCTAGAC  900
  K  A  G  X  A  G  G  N  X  G  X  G  Q  G  Q  N  N  E  X  A  N  M  P  N  V  K  L  V  Q  E  Y  L  D

AAAATTAGATCTACCATTGGCGTCGAGTGGAGTCCATGCAGTGTAACCTGTGGAAAGGGTGTAAGAATGAGAAGAAAAGTTAGTGCAGCTAACAAAAAAC  1000
 K  I  R  S  T  I  G  V  E  W  S  P  C  S  V  T  C  G  K  G  V  R  M  R  R  K  V  S  A  A  N  K  K

CAGAAGAGCTTGATGCGAATGACCTTGAGACTGAAGTTTGTACAATGGATAAGTGCGCTGGTATATTTAACGTTGTGAGTAATTCATTAGGGCTAGTCAT  1100
P  E  E  L  D  A  N  D  L  E  T  E  V  C  T  M  D  K  C  A  G  I  F  N  V  V  S  N  S  L  G  L  V  I

ATTGTTAGTCCTAGCATTATTCAATTAA
  L  L  V  L  A  L  F  N  .
```

SG/EHI/WM04/Y09-8 (*P. knowlesi*; JQ219900)

```
ATGAAGAACTTCATTCTCTTGGCCGTCTCCTCCATCCTGCTGGTGGACTTGCTCCCCACATACTTCGAACATAATGTAGATCTCTCCAGGGCCATAAATG  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  L  P  T  Y  F  E  H  N  V  D  L  S  R  A  I  N

TAAATGGAGTGAGCTTCAATAGTGTAGACACCAGTTCACTTGGCGCAGCACAGGTAAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAAGCCAAA  200
V  N  G  V  S  F  N  S  V  D  T  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  K  P  K

AGAAGGAGCTGATAAAGAAAAGAAAAAAGAAAAAGAAAAAGAAGAAGAACCAAAGAAGCCAAATGAAAATAAGCTGAAACAACCACAAGGAAATGGGGGG  300
  E  G  A  D  K  E  K  K  K  E  K  E  K  E  E  E  P  K  K  P  N  E  N  K  L  K  Q  P  Q  G  N  G  G

GCAGGTCAAGCACAACCACAAGGAAATGGGGGGGCAGGTCAAGCACAACCGGAAGGAAATAGGGAAGCTCCAGCACAACCACAAGGAAATGGAGGGGCAG  400
 A  G  Q  A  Q  P  Q  G  N  G  G  A  G  Q  A  Q  P  E  G  N  R  E  A  P  A  Q  P  Q  G  N  G  G  A

GTCAAGCACAACCACAAAAAAACGAAGGAGGAAACGCAGGAGCACGGAAAGGACAGGGACAAAACAATCAGGGTGCGAATGCCCCAAATGAAAAAGTTGT  500
G  Q  A  Q  P  Q  K  N  E  G  G  N  A  G  A  R  K  G  Q  G  Q  N  N  Q  G  A  N  A  P  N  E  K  V  V

AAATGATTACCTACAGAAAATTAGATCTAGCGTTACCACCGAGTGGACTCCATGCAGTGTAACCTGTGGAAATGGTGTAAGAATTAGAAGAAGAGCTCAT  600
  N  D  Y  L  Q  K  I  R  S  S  V  T  T  E  W  T  P  C  S  V  T  C  G  N  G  V  R  I  R  R  R  A  H

GCAGATAAGAAAAAGGCAGAGGACCTTACTATGGATGACCTTGAAGTGGAAGCTTGTGTAATGGATAAGTGTGCTGGCATATTTAACGTTGTGAGTAATT  700
 A  D  K  K  K  A  E  D  L  T  M  D  D  L  E  V  E  A  C  V  M  D  K  C  A  G  I  F  N  V  V  S  N

CATTAGGGTTAGTCATATCGTTAGTCCTAGCATTATTCAATTAA
S  L  G  L  V  I  S  L  V  L  A  L  F  N  .
```

SG/EHI/WM04/Y09-9 (*P. knowlesi*; JQ219901)

```
ATGAAGAACTTCATTCTCTTGGCCGTCTCCTCCATCCTGCTGGTGGACCTGCTCCCCACATACTTCGAACATAATGTAGATCTCTCCAGGGCCATAAATG  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  L  P  T  Y  F  E  H  N  V  D  L  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAGTGTAGACACCAGTTCACTTGGCGCAGCACAGGTAAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAAGCCAAA  200
V  N  G  V  S  F  N  S  V  D  T  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  K  P  K

AGAAGGAGCTGATAAAGAAAAGAAAAAAGAAAAAGAAAAAGAAGAAGAACCAAAGAAGCCAAATGAAAATAAGCTGAAACAACCACAAGGAAATGGGGGG  300
  E  G  A  D  K  E  K  K  K  E  K  E  K  E  E  E  P  K  K  P  N  E  N  K  L  K  Q  P  Q  G  N  G  G

GCAGGTCAAGCACAACCAGAAGGAAATGGGGGGGCAGGTCAAGCACAACCGGAAGGAAATGGGGGGGCAGGTCAAGCACAACCAGAAGGAAATGGGGGGG  400
 A  G  Q  A  Q  P  E  G  N  G  G  A  G  Q  A  Q  P  E  G  N  G  G  A  G  Q  A  Q  P  E  G  N  G  G

CAGGTCAAGCACAACCGGAAGGAAATGGGGGGGCAGGTCAAGCACAACCGGAAGGAAATGGGGGGGCAGGTCAAGCACAACCGGAAGGAAATGGGGGGGC  500
A  G  Q  A  Q  P  E  G  N  G  G  A  G  Q  A  Q  P  E  G  N  G  G  A  G  Q  A  Q  P  E  G  N  G  G  A

AGGTCAAGCACAACCGGAAGGAAATGGGGGGGCAGGTCAAGCACAACCGGAAGGAAATGGGGGGGCAGGTCAAGCACAACCGGAAGGAAATGGGGGGGCA  600
  G  Q  A  Q  P  E  G  N  G  G  A  G  Q  A  Q  P  E  G  N  G  G  A  G  Q  A  Q  P  E  G  N  G  G  A

GGTCAaGCACAACCGGAAGGAAATgGGGGGGCAGGTCAAGCACAACCGGAAGGAAATGGGGGGGCAGGTCAAGCACAACCGGAAGGAAATGGGGGGGCAG  700
 G  Q  A  Q  P  E  G  N  G  G  A  G  Q  A  Q  P  E  G  N  G  G  A  G  Q  A  Q  P  E  G  N  G  G  A

GTCAAGCACAACCACAAGGAAATGGGGGGGCAGGTCAAGCACAACCACAAAAAAACGAAGGAGGAAACGCAGGAGCACGGAAAGGACAGGGACAAAACAA  800
G  Q  A  Q  P  Q  G  N  G  G  A  G  Q  A  Q  P  Q  K  N  E  G  G  N  A  G  A  R  K  G  Q  G  Q  N  N

TCAGGGTGCGAATGCCCCAAATGAAAAAGTTGTAAATGATTACCTACAGAAAATTAGATCTAGCGTTACCACCGAGTGGACTCCATGCAGTGTAACCTGT  900
  Q  G  A  N  A  P  N  E  K  V  V  N  D  Y  L  Q  K  I  R  S  S  V  T  T  E  W  T  P  C  S  V  T  C

GGAAATGGTGTAAGAATTAGAAGAAGAGCTCATGCAGATAAGAAAAAGGCAGAGGACCTTACTATGGATGACCTTGAAGTGGAAGCTTGTGTAATGGATA  1000
 G  N  G  V  R  I  R  R  R  A  H  A  D  K  K  K  A  E  D  L  T  M  D  D  L  E  V  E  A  C  V  M  D

AGTGTGCTGGCATATTTAACGTTGTGAGTAATTCATTAGGGTTAGTCATATTGTTAGTCCTAGCATTATTCAATTAA
K  C  A  G  I  F  N  V  V  S  N  S  L  G  L  V  I  L  L  V  L  A  L  F  N  .
```

SG/EHI/WM04/Y09-12 (*P. knowlesi*; JQ219902)

```
ATGAAGAACTTCATTCTCTTGGCCGTCTCCTCCATCCTGCTGGTGGACTTGCTCCCCACATACTTCGAACATAATGTAGATCTCTCCAGGGCCATAAATG  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  L  P  T  Y  F  E  H  N  V  D  L  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAGTGTAGACACCAGTTCACTTGGCGCAGCACAGGTAAGACAaAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAAGCCAAA  200
V  N  G  V  S  F  N  S  V  D  T  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  K  P  K

AGAAGGAGCTGATAAAGAAAAGAAAAAAGAAAAAGAAAAAGAAGAAGAACCAAAGAAGCCAAATGAAAATAAGCTGAAACAACCACAAGGAAATGGGGGG  300
  E  G  A  D  K  E  K  K  K  E  K  E  K  E  E  E  P  K  K  P  N  E  N  K  L  K  Q  P  Q  G  N  G  G

GCAGGTCAAGCACAACCAGAAGGAAATGGGGGGGCAGGTCAAGCACAACCGGAAGGAAATGGGGGGGCAGGTCAAGCACAACCAGAAGGAAATGGGGGGG  400
 A  G  Q  A  Q  P  E  G  N  G  G  A  G  Q  A  Q  P  E  G  N  G  G  A  G  Q  A  Q  P  E  G  N  G  G

CAGGTCAAGCACAACCGGAAGGAAATGGGGGGGCAGGTCAAGCACAACCGGAAGGAAATGGGGGGGCAGGTCAAGCACAACCGGAAGGAAATGGGGGGGC  500
A  G  Q  A  Q  P  E  G  N  G  G  A  G  Q  A  Q  P  E  G  N  G  G  A  G  Q  A  Q  P  E  G  N  G  G  A

AGGTCAAGCACAACCGGAAGGAAATGGGGGGGCAGGTCAAGCACAACCGGAAGGAAATGGGGGGGCAGGTCAAGCACAACCGGAAGGAAATGGGGGGGCA  600
  G  Q  A  Q  P  E  G  N  G  G  A  G  Q  A  Q  P  E  G  N  G  G  A  G  Q  A  Q  P  E  G  N  G  G  A

GGTCAAGCACAACCACAAGGAAATGGGGGGGCAGGTCAAGCACAACCACAAGGAAATGGGGGGGCAGGTCAAGCACAACCACAAGGAAATGGGGGGGCAG  700
 G  Q  A  Q  P  Q  G  N  G  G  A  G  Q  A  Q  P  Q  G  N  G  G  A  G  Q  A  Q  P  Q  G  N  G  G  A

GTCAAGCACAACCGGAAGGAAATAGGGAAGCTCCAGCACAACCACAAGGAAATGGGGGGGCAGGTCAAGCACAACCACAAAAAAACGAAGGAGGAAACGC  800
G  Q  A  Q  P  E  G  N  R  E  A  P  A  Q  P  Q  G  N  G  G  A  G  Q  A  Q  P  Q  K  N  E  G  G  N  A

AGGAGCACGGAAAGGACAGGGACAAAACAATCAGGGTGCGAATGCCCCAAATGAAAAAGTTGTAAATGATTACCTACAGAAAATTAGATCTAGCGTTACC  900
  G  A  R  K  G  Q  G  Q  N  N  Q  G  A  N  A  P  N  E  K  V  V  N  D  Y  L  Q  K  I  R  S  S  V  T

ACCGAGTGGACTCCATGCAGTGTAACCTGTGGAAATGGTGTAAGAATTAGAAGAAGAGCTCATGCAGATAAGAAAAAGGCAGAGGACCTTACTATGGATG  1000
 T  E  W  T  P  C  S  V  T  C  G  N  G  V  R  I  R  R  R  A  H  A  D  K  K  K  A  E  D  L  T  M  D

ACCTTGAAGTGGAAGCTTGTGTAATGGATAAGTGTGCTGGCATATTTAACGTTGTGAGTAATTCATTAGGGTTAGTCATATTGTTAGTCCTAGCATTATT  1100
D  L  E  V  E  A  C  V  M  D  K  C  A  G  I  F  N  V  V  S  N  S  L  G  L  V  I  L  L  V  L  A  L  F

CAATTAA
  N  .
```

SG/EHI/WM04/Y09-13 (*P. knowlesi*; JQ219903)

```
ATGAAGAACTTCATTCTCTTGGCCGTCTCCTCCATCCTGCTGGTGGACTTGCTCCCCACATACTTCGAACATAATGTAGATCTCTCCAGGGCCATAAATG  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  L  P  T  Y  F  E  H  N  V  D  L  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAGTGTAGACACCAGTTCACTTGGCGCAGCACAGGTAAGACAGAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAAGCCAAA  200
V  N  G  V  S  F  N  S  V  D  T  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  K  P  K

AGAAGGAGCTGATAAAGAAAAGAAAAAAGAAAAAGAAAAAGAAGAAGAACCAAAGAAGCCAAATGAAAATAAGCTGAAACAACCACAAGGAAATGGGGGG  300
  E  G  A  D  K  E  K  K  K  E  K  E  K  E  E  E  P  K  K  P  N  E  N  K  L  K  Q  P  Q  G  N  G  G

GCAGGTCAAGCACAACCAGAAGGAAATGGGGGGGCAGGTCAAGCACAACCGGAAGGAAATGGGGGGGCAGGTCAAGCACAACCAGAAGGAAATGGGGGGG  400
 A  G  Q  A  Q  P  E  G  N  G  G  A  G  Q  A  Q  P  E  G  N  G  G  A  G  Q  A  Q  P  E  G  N  G  G

CAGGTCAAGCACAACCGGAAGGAAATGGGGGGGCAGGTCAAGCACAACCGGAAGGAAaTGGGGGGGCAGGTCAAGCACAACCGGAAGGAAATGGGGGGGC  500
A  G  Q  A  Q  P  E  G  N  G  G  A  G  Q  A  Q  P  E  G  N  G  G  A  G  Q  A  Q  P  E  G  N  G  G  A

AGGTCAAGCACAACCGGAAGGAAATGGGGGGGCAGGTCAAGCACAACCGGAAGGAAATGGGGGGGCAGGTCAAGCACAACCGGAAGGAAATGGGGGGGCA  600
  G  Q  A  Q  P  E  G  N  G  G  A  G  Q  A  Q  P  E  G  N  G  G  A  G  Q  A  Q  P  E  G  N  G  G  A

GGTCAAGCACAACCACAAGGAAATGGGGGGGCAGGTCAAGCACAACCACAAGGAAATGGGGGGGCAGGTCAAGCACAACCACAAGGAAATGGGGGGGCAG  700
 G  Q  A  Q  P  Q  G  N  G  G  A  G  Q  A  Q  P  Q  G  N  G  G  A  G  Q  A  Q  P  Q  G  N  G  G  A

GTCAAGCACAACCGGAAGGAAATAGGGAAGCTCCAGCACAACCACAAGGAAATGGGGGGGCAGGTCAAGCACAGCCACAAAAAAACGAAGGAGGAAACGC  800
G  Q  A  Q  P  E  G  N  R  E  A  P  A  Q  P  Q  G  N  G  G  A  G  Q  A  Q  P  Q  K  N  E  G  G  N  A

AGGAGCACGGAAAGGACAGGGACAAAACAATCAGGGTGCGAATGCCCCAAATGAAAAAGTTGTAAATGATTACCTACAGAAAATTAGATCTAGCGTTACC  900
  G  A  R  K  G  Q  G  Q  N  N  Q  G  A  N  A  P  N  E  K  V  V  N  D  Y  L  Q  K  I  R  S  S  V  T

ACCGAGTGGACTCCATGCAGTGTAACCTGTGGAAATGGTGTAAGAATTAGAAGAAGAGCTCATGCAGATAAGAAAAAGGCAGAGGACCTTACTGTGGATG  1000
 T  E  W  T  P  C  S  V  T  C  G  N  G  V  R  I  R  R  R  A  H  A  D  K  K  K  A  E  D  L  T  V  D

ACCTTGAAGTGGAAGCTTGTGTAATGGATAAGTGTGCTGGCATATTTAACGTTGTGAGTAATTCATTAGGGTTAGTCATATTGTTAGTCCTAGCATTATT  1100
D  L  E  V  E  A  C  V  M  D  K  C  A  G  I  F  N  V  V  S  N  S  L  G  L  V  I  L  L  V  L  A  L  F

CAATTAA
  N  .
```

SG/EHI/WM04/Y09-14 (*P. knowlesi*; JQ219904)

```
ATGAAGAACTTCATTCTCTTGGCCGTCTCCTCCATCCTGCTGGTGGACTTGCTCCCCACATACTTCGAACATAATGTAGATCTCTCCAGGGCCATAAATG  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  L  P  T  Y  F  E  H  N  V  D  L  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAGTGTAGACACCAGTTCACTTGGCGCAGCACAGGTAAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAAGCCAAA  200
V  N  G  V  S  F  N  S  V  D  T  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  K  P  K

AGAAGGAGCTGATAAAGAAAAGAAAAAAGAAAAAGAAAAAGAAGAAGAACCAAAGAAGCCAAATGAAAATAAGCTGAAACAACCACAAGgAAATGGGGGG  300
  E  G  A  D  K  E  K  K  K  E  K  E  K  E  E  E  P  K  K  P  N  E  N  K  L  K  Q  P  Q  G  N  G  G

GCAGGTCAAGCACAACCAGAAGGAAATGGGGGGGCAGGTCAAGCACAACCGGAAGGAAATGGGGGGGCAGGTCAAGCACAACCAGAAGGAAATGGGGGGG  400
 A  G  Q  A  Q  P  E  G  N  G  G  A  G  Q  A  Q  P  E  G  N  G  G  A  G  Q  A  Q  P  E  G  N  G  G

CAGGTCAAGCACAACCGGAAGGAAATGGGGGGGCAGGTCAAGCACAACCGGAAGGAAATGGGGGGGCAGGTCAAGCACAACCGGAAGGAAATGGGGGGGC  500
A  G  Q  A  Q  P  E  G  N  G  G  A  G  Q  A  Q  P  E  G  N  G  G  A  G  Q  A  Q  P  E  G  N  G  G  A

AGGTCAAGCACAACCGGAAGGAAATGGGGGGGCAGGTCAAGCACAACCGGAAGGAAATGGGGGGGCAGGTCAAGCACAACCGGAAGGAAATGGGGGGGCA  600
  G  Q  A  Q  P  E  G  N  G  G  A  G  Q  A  Q  P  E  G  N  G  G  A  G  Q  A  Q  P  E  G  N  G  G  A

GGTCAAGCACAACCACAAGGAAATGGGGGGGCAGGTCAAGCACAACCACAAgGAAATGGGGGGGCAGGTCAAGCACAACCACAAGGAAATGGGGGGGCAG  700
 G  Q  A  Q  P  Q  G  N  G  G  A  G  Q  A  Q  P  Q  G  N  G  G  A  G  Q  A  Q  P  Q  G  N  G  G  A

GTCAAGCACAACCGGAAGGAAATAGGGAAGCTCCAGCACAACCACAAGGAAATGGGGGGGCAGGTCAAGCACAACCACAAAAAAACGAAGGAGGAAACGC  800
G  Q  A  Q  P  E  G  N  R  E  A  P  A  Q  P  Q  G  N  G  G  A  G  Q  A  Q  P  Q  K  N  E  G  G  N  A

AGGAGCACGGAAAGGACAGGGACAAAACAATCAGGGTGCGAATGCCCCAAATGAAAAAGTTGTAAATGATTACCTACAGAAAATTAGATCTAGCGTTACC  900
  G  A  R  K  G  Q  G  Q  N  N  Q  G  A  N  A  P  N  E  K  V  V  N  D  Y  L  Q  K  I  R  S  S  V  T

ACCGAGTGGACTCCATGCAGTGTAACCTGTGGAAATGGTGTAAGAATTAGAAGAAGAGCTCATGCGGATAAGAAAAAGGCAGAGGACCTTACTATGGATG  1000
 T  E  W  T  P  C  S  V  T  C  G  N  G  V  R  I  R  R  R  A  H  A  D  K  K  K  A  E  D  L  T  M  D

ACCTTGAAGTGGAAGCTTGTGTAATGGATAAGTGTGCTGGCATATTTAACGTTGTGAGTAATTCATTAGGGTTAGTCATATTGTTAGTCCTAGCATTATT  1100
D  L  E  V  E  A  C  V  M  D  K  C  A  G  I  F  N  V  V  S  N  S  L  G  L  V  I  L  L  V  L  A  L  F

CAATTAA
  N  .
```

SG/EHI/WM04/Y09-15 (*P. knowlesi*; JQ219905)

```
ATGAAGAACTTCATTCTCTTGGCCGTCTCCTCCATCCTGCTGGTGGACTTGCTCCCCACATACTTCGAACATAATGTAGATCTCTCCAGGGCCGTAAATG  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  L  P  T  Y  F  E  H  N  V  D  L  S  R  A  V  N

TAAATGGAGTAAGCTTCAATAGTGTAGACACCAGTTCACTTGGCGCAGCACAGGTAAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAAGCCAAA  200
V  N  G  V  S  F  N  S  V  D  T  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  K  P  K

AGAAGGAGCTGATAAAGAAAAGAAAAAAGAAAAAGAAAAAGAAGAAGAACCAAAGAAGCCAAATGAAAATAAGCTGAAACAACCACAAGGAAATGGGGGG  300
  E  G  A  D  K  E  K  K  K  E  K  E  K  E  E  E  P  K  K  P  N  E  N  K  L  K  Q  P  Q  G  N  G  G

GCAGGTCAAGCACAACCAGAAGGAAATGGGGGGGCAGGTCAAGCACAACCGGAAGGAAATGGGGGGGCAGGTCAAGCACAACCAGAAGGAAATGGGGGGG  400
 A  G  Q  A  Q  P  E  G  N  G  G  A  G  Q  A  Q  P  E  G  N  G  G  A  G  Q  A  Q  P  E  G  N  G  G

CAGGTCAAGCACAACCGGAAGGAAATGGGGGGGCAGGTCAAGCACAACCGGAAGGAAATGGGGGGGCAGGTCAAGCACAACCGGAAGGAAATGGGGGGGC  500
A  G  Q  A  Q  P  E  G  N  G  G  A  G  Q  A  Q  P  E  G  N  G  G  A  G  Q  A  Q  P  E  G  N  G  G  A

AGGTCAAGCACAACCGGAAGGAAATGGGGGGGCAGGTCAAGCACAACCACAAGGAAATGGGGGGGCAGGTCAAGCACAACCACAAGGAAATGGGGGGGCA  600
  G  Q  A  Q  P  E  G  N  G  G  A  G  Q  A  Q  P  Q  G  N  G  G  A  G  Q  A  Q  P  Q  G  N  G  G  A

GGTCAAGCACAACCACAAGGAAATGGGGGGGCAGGTCAAGCACAACCACAAGGAAATGGGGGGGCAGGTCAAGCACAACCGGAAGGAAATGGGGGGGCAG  700
 G  Q  A  Q  P  Q  G  N  G  G  A  G  Q  A  Q  P  Q  G  N  G  G  A  G  Q  A  Q  P  E  G  N  G  G  A

GTCAAGCACAACCACAAGGAAATGGGGGGGCAGGTCAAGCACAACCGGAAGGAAATAGGGAAGCTCCAGCACAACCACAAGGAAATGGGGGGGCAGGTCA  800
G  Q  A  Q  P  Q  G  N  G  G  A  G  Q  A  Q  P  E  G  N  R  E  A  P  A  Q  P  Q  G  N  G  G  A  G  Q

AGCACAACCACAAAAAAACGAAGGAGGAAACGCAGGAGCACGGAAAGGACAGGGACAAAACAGTCAGGGTGCGAATGCCCCAAATGAAAAAGTTGTAAAT  900
  A  Q  P  Q  K  N  E  G  G  N  A  G  A  R  K  G  Q  G  Q  N  S  Q  G  A  N  A  P  N  E  K  V  V  N

GATTACCTACAGAAAATTAGATCTAGCGTTACCACCGAGTGGACTCCATGCAGTGTAACCTGTGGAAATGGTGTAAGAATTAGAAGGAGAGCTCATGCAG  1000
 D  Y  L  Q  K  I  R  S  S  V  T  T  E  W  T  P  C  S  V  T  C  G  N  G  V  R  I  R  R  R  A  H  A

ATAAGAAAAAGGCAGGGGACCTTACTATGGATGACCTTGAAGTGGAAGCTTGTGTAATGGATAAGTGTGCTGGCATATTTAACGCTGTGAGTAATTCATT  1100
D  K  K  K  A  G  D  L  T  M  D  D  L  E  V  E  A  C  V  M  D  K  C  A  G  I  F  N  A  V  S  N  S  L

AGGGTTAGTCATATTGTTAGTCCTAGCATTATTCAATTAA
  G  L  V  I  L  L  V  L  A  L  F  N  .
```

SG/EHI/WM05/Y09-70 (*P. knowlesi*; JQ219906)

```
ATGAAGAACTTCATTCTCTTGGCCGTCTCCTCCATCCTGCTGGTGGACTTGCTCCCCACACACTTCGAACATAATGTAGATCTCTCCAGGGCCATAAATG  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  L  P  T  H  F  E  H  N  V  D  L  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAATGTAGACACCAGTTCACTTGGCGCAGCACAGGTGAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAAGCCAAA  200
V  N  G  V  S  F  N  N  V  D  T  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  K  P  K

AGAAGGAGCTGATAAAGAAAAGAAAAAAGAAAAAGAAGAAGAACCAAAGAAGCCAAATGAAAATAAGCTGAAACAACCGGATGCAGTACCAGGGGGCGAA  300
  E  G  A  D  K  E  K  K  K  E  K  E  E  E  P  K  K  P  N  E  N  K  L  K  Q  P  D  A  V  P  G  G  E

GAACCAGCACCAGGAAGGGAACAGCCAGCACCAGGAAGGGAGGAACCAGCACCAGGAAGGGAACAGCCAGCACCAGGAAGGGAAGAACCAGCACCAGGAA  400
 E  P  A  P  G  R  E  Q  P  A  P  G  R  E  E  P  A  P  G  R  E  Q  P  A  P  G  R  E  E  P  A  P  G

GGGAACAGCCAGCACCAGGAAGGGAAGAACCAGCGCCAGGAAGGGAACAGCCAGCACCAGGAAGGGAACAGCCAGCACCAGGAAGGGAGGAACCAGCACC  500
R  E  Q  P  A  P  G  R  E  E  P  A  P  G  R  E  Q  P  A  P  G  R  E  Q  P  A  P  G  R  E  E  P  A  P

AGGAAGGGAACAACCAGCACCAGGAAGGGAGGAACCAGCACCAGGAAGGGAACAACCAGCACCAGGAAGGGAACAGCCAGCACCAGGAAGGGAACAGCCA  600
  G  R  E  Q  P  A  P  G  R  E  E  P  A  P  G  R  E  Q  P  A  P  G  R  E  Q  P  A  P  G  R  E  Q  P

GCACCGGGGGGTGAACAACCAGCACCAGGAAGGGAACAGCCAGCACCGGGTGGTGAACAACCAGCACCAGCACCAGGAGCAGGTGCGGGAGATGGAGCAC  700
 A  P  G  G  E  Q  P  A  P  G  R  E  Q  P  A  P  G  G  E  Q  P  A  P  A  P  G  A  G  A  G  D  G  A

GAGGAGGAAACGCAGGGGCAGGGAAAGGACAGGGACAAAACAATCAGGGTGCAAATGTCCCAAATGAAAAAGTTGTGAATGATTACCTACACAAAATTAG  800
R  G  G  N  A  G  A  G  K  G  Q  G  Q  N  N  Q  G  A  N  V  P  N  E  K  V  V  N  D  Y  L  H  K  I  R

ATCTAGCGTTACCACCGAGTGGACTCCATGCAGTGTAACCTGTGGAAATGGTGTAAGAATTAGAAGAAAGGGTCATGCAGGTAATAAAAAGGCAGAGGAC  900
  S  S  V  T  T  E  W  T  P  C  S  V  T  C  G  N  G  V  R  I  R  R  K  G  H  A  G  N  K  K  A  E  D

CTTACTATGGATGACCTTGAGGTGGAAGCTTGTGTAATGGATAAGTGCGCTGGCATATTTAACGTTGTGAGTAATTCATTAGGCTTAGTCATATTGTTAG  1000
 L  T  M  D  D  L  E  V  E  A  C  V  M  D  K  C  A  G  I  F  N  V  V  S  N  S  L  G  L  V  I  L  L

TCCTAGCATTATTCAATTAA
V  L  A  L  F  N  .
```

SG/EHI/WM05/Y09-79 (*P. knowlesi*; JQ219907)

```
ATGAAGAACTTCATTCTCTTGGCCGTCTCCTCCATCCTGCTGGTGGACTTGCTCCCCACACACTTCGAACATAATGTAGATCTCTCCAGGGCCATAAATG  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  L  P  T  H  F  E  H  N  V  D  L  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAATGTAGACACCAGTTCACTTGGCGCAGCACAGGTGAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAGGCCAAA  200
V  N  G  V  S  F  N  N  V  D  T  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  R  P  K

AGAAGGAGCTGATAAAGAAAAGAAAAAAGAAAAAGAAAAAGAAAAAGAAGAACCAAAGAAGCCAAATGAAAATAAGCTGAAACAACCGGAACAACCAGCA  300
  E  G  A  D  K  E  K  K  K  E  K  E  K  E  K  E  E  P  K  K  P  N  E  N  K  L  K  Q  P  E  Q  P  A

GCAGGAGCAGGGGGCGAACAACCAGCACCAGCACCAAGGAGGGAACAACCAGCACCAGGAGCAGGTGCGGGAGATGGAGCACGAGGAGGAAACGCAGGGG  400
 A  G  A  G  G  E  Q  P  A  P  A  P  R  R  E  Q  P  A  P  G  A  G  A  G  D  G  A  R  G  G  N  A  G

CAGGTAAAGGACAGGGACAAAACAATCAGGGTGCGAATGTCCCAAATGAAAAAGTTGTGAATGATTACCTACACAAAATTAGATCTAGCGTTACCACCGA  500
A  G  K  G  Q  G  Q  N  N  Q  G  A  N  V  P  N  E  K  V  V  N  D  Y  L  H  K  I  R  S  S  V  T  T  E

GTGGACTCCATGCAGTGTAACCTGTGGAAATGGTGTAAGAATTAGAAGAAGACAGAATGCTGGTAATAAAAAGGCAGAGGACCTTACTATGGATGACCTT  600
  W  T  P  C  S  V  T  C  G  N  G  V  R  I  R  R  R  Q  N  A  G  N  K  K  A  E  D  L  T  M  D  D  L

GAGGTGGAAGCTTGTGTAATGGATAAGTGCGCTGGCATATTTAACGTTGTGAGTAATTCATTAGGCTTAGTCATATTGTTAGTCTTAGCATTATTCAATT  700
 E  V  E  A  C  V  M  D  K  C  A  G  I  F  N  V  V  S  N  S  L  G  L  V  I  L  L  V  L  A  L  F  N

AA
```

SG/EHI/WM05/Y09-65 (*P. cynomolgi*; JQ219923)

```
ATGAAGAACTTCATTCTCTTAGCCGTTTCTTCCATCCTGTTGGTGGACTTGTTTCCCACAAACTGCGGGCATAATGTAGATTTCTCCAGGGCCATAAATT  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  F  P  T  N  C  G  H  N  V  D  F  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAATGTAGACGCCAGTTCACTTGGCGCAGCACAGGTAAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAACCCAAA  200
L  N  G  V  S  F  N  N  V  D  A  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  N  P  K

AAACGAGGAAGGAGCTGCTAAACAAAAAAAGGACGAAAAAAAAGTAGAACCAAAAAAGCCACGTGAAAATAAGCTGAAACAACCACCAGCAGCAGCAGAT  300
  N  E  E  G  A  A  K  Q  K  K  D  E  K  K  V  E  P  K  K  P  R  E  N  K  L  K  Q  P  P  A  A  A  D

GGAGCACCCGCAGCAGCAGATGGAGCACCCGCAGCAGCAGATGGAGCACCCGCAGCAGCAGATGGAGCACCAGCAGCAGCAGATGGAGCACCAGCAGCAG  400
 G  A  P  A  A  A  D  G  A  P  A  A  A  D  G  A  P  A  A  A  D  G  A  P  A  A  A  D  G  A  P  A  A

CAGATGGAGCACCAGCAGCAGCAGATGGAGCACCCGCAGCAGCAGATGGGGCACCAGCAGCAGCAGATGGAGCACCAGCAGCAGCAGATGGAGCACCAGC  500
A  D  G  A  P  A  A  A  D  G  A  P  A  A  A  D  G  A  P  A  A  A  D  G  A  P  A  A  A  D  G  A  P  A

AGCAGGAGGAAATCAGGCAGGTGCACAAGCAGGAGCAGGAGGAAATCAGGCAGGTGCACAAGCAGGAGCAGGAGGAAATCAGGCAGGTGCACAAGCAGGA  600
  A  G  G  N  Q  A  G  A  Q  A  G  A  G  G  N  Q  A  G  A  Q  A  G  A  G  G  N  Q  A  G  A  Q  A  G

GCAGGAGGAAATCAGGCAGGTGGACAGCCAGGAGCAGGAGGAAATCAGGCAGGTGGACAGCCAGGAGCAGGAGGAAATCAGGCAGGTGGACAGCCAGGAG  700
 A  G  G  N  Q  A  G  G  Q  P  G  A  G  G  N  Q  A  G  G  Q  P  G  A  G  G  N  Q  A  G  G  Q  P  G

CAGGAGGAAATCAGGCAGGTGGACAAGCAGGAGCAGGAGGAAATCAGGCAGGTGGACAAGCAGGAGGAGCAAATGCGGGAAATAAAAAAGCAGGAGACGC  800
A  G  G  N  Q  A  G  G  Q  A  G  A  G  G  N  Q  A  G  G  Q  A  G  G  A  N  A  G  N  K  K  A  G  D  A

AGGACAGGGACAAAATAATGGAGGTGCGAATGTGCCAAATGTAAAGCTTGTGCAAGAATACCTAGACAAAATTAGATCTACCATTAGCACCGAGTGGAGT  900
  G  Q  G  Q  N  N  G  G  A  N  V  P  N  V  K  L  V  Q  E  Y  L  D  K  I  R  S  T  I  S  T  E  W  S

CCATGCAGTGTAACCTGTGGAAAGGGTGTAAGAATGAGAAAAAAAGTTAATGCAGCTAACAAAAAACCAGAAGAGCTTGATGTGAATGACCTTGAGGCAG  1000
 P  C  S  V  T  C  G  K  G  V  R  M  R  K  K  V  N  A  A  N  K  K  P  E  E  L  D  V  N  D  L  E  A

AAGTTTGTACAATGGATAAGTGCGCTGGTATATTTAACGTTGTGAGTAATTCATTAGGGCTAGTCATATTGTTAGTCCTAGCATTATTCAATTAA
E  V  C  T  M  D  K  C  A  G  I  F  N  V  V  S  N  S  L  G  L  V  I  L  L  V  L  A  L  F  N  .
```

SG/EHI/WM05/Y09-68 (*P. fieldi*; JQ219932)

```
ATGAAGAACTTCATTCTCTTAGCCGTTTCTTCCATCCTGTTGGTGGACTTGTTCCCCACACACTGCGGGCATAATGTAGATCTCTCCAGGGACATAAATT  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  F  P  T  H  C  G  H  N  V  D  L  S  R  D  I  N

TAAATGGAGCAAGCTTCAATAATGTAGACGCCAGTTCACTTGGCGCAGCACAGGTAAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAACCCAAA  200
L  N  G  A  S  F  N  N  V  D  A  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  N  P  K

AGATGAAGGAGCTGCTAAAGAAAAAAAGGACGAAAAAAAAGTAGAACCAAAAAAGGCACGTGAAAATAAGCTGAAACAACCAGCTAATGATGCAGGACAA  300
  D  E  G  A  A  K  E  K  K  D  E  K  K  V  E  P  K  K  A  R  E  N  K  L  K  Q  P  A  N  D  A  G  Q

AATCAGCCAGCTAATGATGCAGGACAAAATCAGCCAGCTAATGACGCAGGACAAAATCAGCCAGCTAATGATGCAGGACAAAATCAGCCAGCTAATGATG  400
 N  Q  P  A  N  D  A  G  Q  N  Q  P  A  N  D  A  G  Q  N  Q  P  A  N  D  A  G  Q  N  Q  P  A  N  D

CAGGACAAAATCAGCCAGCTAATGATGCAGGACAAAATCAGCCAGCTAATGATGCAGGACAAAATCAGCCAGCTAATGATGCAGGACAAAATCAGCCAGC  500
A  G  Q  N  Q  P  A  N  D  A  G  Q  N  Q  P  A  N  D  A  G  Q  N  Q  P  A  N  D  A  G  Q  N  Q  P  A

TAATGATGCAGGACAAAATCAACCAGCTAATGATGCAGGACAAAATCAGCCAGCTAATGATGCAGGACAAAATCAACCAGCTAATGATGCAGGACAAAAT  600
  N  D  A  G  Q  N  Q  P  A  N  D  A  G  Q  N  Q  P  A  N  D  A  G  Q  N  Q  P  A  N  D  A  G  Q  N

CAGCCAGCTAATGATGCAGGACAAAATCAGCCAGGTGGTGGAGCAGGACAAAATCAGCCAGGTGGTGGAGCAGGACAAAATCAACCAGGTGGTGGAGCAG  700
 Q  P  A  N  D  A  G  Q  N  Q  P  G  G  G  A  G  Q  N  Q  P  G  G  G  A  G  Q  N  Q  P  G  G  G  A

GACAAAATCAACCAGGTGGTGGAGCAGGACAAAATCAACCAGGTGGTGGAGCAGGACAAAATCAGCCAGGAAATGGAGCAGGACAAAATCAGCCAGGAGA  800
G  Q  N  Q  P  G  G  G  A  G  Q  N  Q  P  G  G  G  A  G  Q  N  Q  P  G  N  G  A  G  Q  N  Q  P  G  D

TGGAGCAGGACGAAATGGTGGAAACGCAGGAGCAGGAGGACAGGGACAAAATAATGAAGGTGCGAATAAGCCAGATGAAAAGCATGTGAAAGAATACCTA  900
  G  A  G  R  N  G  G  N  A  G  A  G  G  Q  G  Q  N  N  E  G  A  N  K  P  D  E  K  H  V  K  E  Y  L

GAGAAAATTAGATCTACCGTTGGCACCGAATGGACTCCATGCAGTGTAACCTGTGGAAAGGGTGTAAGAGTTAGAAGAAAACTTAGTGCAGGTGACAAAA  1000
 E  K  I  R  S  T  V  G  T  E  W  T  P  C  S  V  T  C  G  K  G  V  R  V  R  R  K  L  S  A  G  D  K

AACCAGATAAGCTTACTCTGAATGACCTTGAGGCAGAAGTTTGTACAATGGATAAGTGCGCTGGCATATTTAACGTTGTGAGTAATTCATTAGGGCTAGT  1100
K  P  D  K  L  T  L  N  D  L  E  A  E  V  C  T  M  D  K  C  A  G  I  F  N  V  V  S  N  S  L  G  L  V

CATATTGTTAGTCCTAGCATTATTCAATTAA
  I  L  L  V  L  A  L  F  N  .
```

SG/EHI/WM11/Y09-74 (*P. knowlesi*; JQ219908)

```
ATGAAGAACTTCATTCTCTTGGCCGTCTCCTCCATCCTGCTGGTGGACTTGCTCCCCACACACTTCGAACATAATGTAGATCTCTCCAGGGCCATAAATG  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  L  P  T  H  F  E  H  N  V  D  L  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAATGTAGACACCAGTTCACTTGGCGCAGCACAGGTGAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAAGCCAAA  200
V  N  G  V  S  F  N  N  V  D  T  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  K  P  K

AGAAGGAGCTGATAAAGAAAAGAAAAAAGAAAAAGAAAAAGAAAAAGAAGAACCAAAGAAGCCAAATGAAAATAAGCTGAAACAACCGGAACAACCAGCA  300
  E  G  A  D  K  E  K  K  K  E  K  E  K  E  K  E  E  P  K  K  P  N  E  N  K  L  K  Q  P  E  Q  P  A

GCAGGAGCAGGGGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAAC  400
 A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E

AACCAGCGGCAGGAGCAGGGGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCAGCAGGAGCAGG  500
Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G

AGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCAGCAGGAGCAGGGGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCACCA  600
  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  P

GCACCAAGGAGGGAACAACCAGCAGCAGGAGCAGGGGGCGAACAACCAGCACCAGCACCAAGGAGGGAACAACCAGCACCAGGAGCAGGTGCGGGAGATG  700
 A  P  R  R  E  Q  P  A  A  G  A  G  G  E  Q  P  A  P  A  P  R  R  E  Q  P  A  P  G  A  G  A  G  D

GAGCACGAGGAGGAAACGCAGGGGCAGGTAAAGGACAGGGACAAAACAATCAGGGTGCGAATGTCCCAAATGAAAAAGTTGTGAATGATTACCTACACAA  800
G  A  R  G  G  N  A  G  A  G  K  G  Q  G  Q  N  N  Q  G  A  N  V  P  N  E  K  V  V  N  D  Y  L  H  K

AATTAGATCTAGCGTTACCACCGAGTGGACTCCATGCAGTGTAACCTGTGGAAATGGTGTAAGAATTAGAAGAAGACAGAATGCTGGTAATAAAAAGGCA  900
  I  R  S  S  V  T  T  E  W  T  P  C  S  V  T  C  G  N  G  V  R  I  R  R  R  Q  N  A  G  N  K  K  A

GAGGACCTTACTATGGATGACCTTGAGGTGGAAGCTTGTGTAATGGATAAGTGCGCTGGCATATTTAACGTTGTGAGTAATTCATTAGGCTTAGTCATAT 1000
 E  D  L  T  M  D  D  L  E  V  E  A  C  V  M  D  K  C  A  G  I  F  N  V  V  S  N  S  L  G  L  V  I

TGTTAGTCTTAGCATTATTCAATTAA
L  L  V  L  A  L  F  N  .
```

SG/EHI/WM15/Y09-149 (*P. knowlesi*; JQ219909)

```
ATGAAGAACTTCATTCTCTTGGCCGTCTCCTCCATCCTGCTGGTGGACTTGCTCCCCACATACTTCGAACATAATGTAGATCTCTCCAGGGCCATAAATG  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  L  P  T  Y  F  E  H  N  V  D  L  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAGTGTAGACACCAGTTCACTTGGCGCAGCACAGGTAAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAAGCCAAA  200
V  N  G  V  S  F  N  S  V  D  T  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  K  P  K

AGAAGGAGCTGATAAAGAAAAGAAAAAAGAAAAAGAAAAAGAAGAAGAACCAAAGAAGCCAAATGAAAATAAGCTGAAACAACCACAAGGAAATGGGGGG  300
  E  G  A  D  K  E  K  K  K  E  K  E  K  E  E  P  K  K  P  N  E  N  K  L  K  Q  P  Q  G  N  G  G

GCAGGTCAAGCACAACCAGAAGGAAATGGGGGGGCAGGTCAAGCACAACCGGAAGGAAATGGGGGGGCAGGTCAAGCACAACCAGAAGGAAATGGGGGGG  400
 A  G  Q  A  Q  P  E  G  N  G  G  A  G  Q  A  Q  P  E  G  N  G  G  A  G  Q  A  Q  P  E  G  N  G  G

CAGGTCAAGCACAACCGGAAGGAAATGGGGGGGCAGGTCAAGCACAACCGGAAGGAAATGGGGGGGCAGGTCAAGCACAACCGGAAGGAAATGGGGGGGC  500
A  G  Q  A  Q  P  E  G  N  G  G  A  G  Q  A  Q  P  E  G  N  G  G  A  G  Q  A  Q  P  E  G  N  G  G  A

AGGTCAAGCACAACCGGAAGGAAATGGGGGGGCAGGTCAAGCACAACCGGAAGGAAATGGGGGGGCAGGTCAAGCACAACCGGAAGGAAATGGGGGGGCA  600
  G  Q  A  Q  P  E  G  N  G  G  A  G  Q  A  Q  P  E  G  N  G  G  A  G  Q  A  Q  P  E  G  N  G  G  A

GGTCAAGCACAACCACAAGGAAATGGGGGGGCAGGTCAAGCACAACCACAAGGAAATGGGGGGGCAGGTCAAGCACAACCACAAGGAAATGGGGGGGCAG  700
 G  Q  A  Q  P  Q  G  N  G  G  A  G  Q  A  Q  P  Q  G  N  G  G  A  G  Q  A  Q  P  Q  G  N  G  G  A

GTCAAGCACAACCGGAAGGAAATAGGGAAGCTCCAGCACAACCACAAGGAAATGGGGGGGCAGGTCAAGCACAACCACAAAAAAACGAAGGAGGAAACGC  800
G  Q  A  Q  P  E  G  N  R  E  A  P  A  Q  P  Q  G  N  G  G  A  G  Q  A  Q  P  Q  K  N  E  G  G  N  A

AGGAGCACGGAAAGGACAGGGACAAAACAATCAGGGTGCGAATGCCCCAAATGAAAAAGTTGTAAATGATTACCTACAGAAAATTAGATCTAGCGTTACC  900
  G  A  R  K  G  Q  G  Q  N  N  Q  G  A  N  A  P  N  E  K  V  V  N  D  Y  L  Q  K  I  R  S  S  V  T

ACCGAGTGGACTCCATGCAGTGTAACCTGTGGAAATGGTGTAAGAATTAGAAGAAGAGCTCATGCAGATAAGAAAAAGGCAGAGGACCTTACTATGGATG 1000
 T  E  W  T  P  C  S  V  T  C  G  N  G  V  R  I  R  R  R  A  H  A  D  K  K  K  A  E  D  L  T  M  D

ACCTTGAAGTGGAAGCTTGTGTAATGGATAAGTGTGCTGGCATATTTAACGTTGTGAGTAATTCATTAGGGTTAGTCATATTGTTAGTCCTAGCATTATT 1100
D  L  E  V  E  A  C  V  M  D  K  C  A  G  I  F  N  V  V  S  N  S  L  G  L  V  I  L  L  V  L  A  L  F

CAATTAA
  N  .
```

SG/EHI/WM15/Y09-163 (*P. fieldi*; JQ219933)

```
ATGAAGAACTTCATTCTCTTGGCCGTTTCTTCCATCCTGTTGGTGGACTTGTTCCCCACACACTGCGGGCATAATGTAGATCTCTCCAGGGCAATAAATT  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  F  P  T  H  C  G  H  N  V  D  L  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAATGTAGACGCCAGTTCACTTGGCGCAGCACAGGTAAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAACCCAAA  200
L  N  G  V  S  F  N  N  V  D  A  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  N  P  K

AGACGATGAAAAAGCTGATAAACCAAAAAAAAAAGGACGAAAAAAAAAGTAGAACCAAAAAAGCCACATGAAAATAAGCTGAAACAACCAGTCCCAGGAGCA  300
  D  D  E  K  A  D  K  P  K  K  K  D  E  K  K  V  E  P  K  K  P  H  E  N  K  L  K  Q  P  V  P  G  A

AATCAGGAAGGCGGAGCAGCAGCCCCAGGAGCAAATCAGGAAGGCGGAGCAGCAGCCCCAGGAGCAAATCAAGAAGGTGGAGCAGCAGCCCCAGGTGCAA  400
 N  Q  E  G  G  A  A  A  P  G  A  N  Q  E  G  G  A  A  A  P  G  A  N  Q  E  G  G  A  A  A  P  G  A

ACCAGGAAGGTGGAGCAGCAGCCCCAGGAGCAAACCAGGAAGGTGGAGCAGCAGCCCCAGGAGCAAATCAGGAAGGCGGAGCAGCAGCCCCAGGAGCAAA  500
N  Q  E  G  G  A  A  A  P  G  A  N  Q  E  G  G  A  A  A  P  G  A  N  Q  E  G  G  A  A  A  P  G  A  N

TCAGGAAGGTGGAGCAGCAGCCCCAGGTGCAAACCAGGAAGGTGGAGCAGCAGCCCCAGGAGCAAATCAGGAAGGCGGAGCAGCAGCCCCAGGAGCAAAT  600
  Q  E  G  G  A  A  A  P  G  A  N  Q  E  G  G  A  A  A  P  G  A  N  Q  E  G  G  A  A  A  P  G  A  N

CAGGAAGGCGGAGCAGCAGCCCCAGGAGCAAATCAGGAAGGTGGAGCAGCAGCCCCAGGAGCAAACCAGGAAGGTGGAGCAGCAGCACCAGGAGCAAACC  700
 Q  E  G  G  A  A  A  P  G  A  N  Q  E  G  G  A  A  A  P  G  A  N  Q  E  G  G  A  A  A  P  G  A  N

AGGGAGGTGGAGCAGCAGCACCAGGAGCAAACCAGGGAGGTGGAGCAGCAGCACCAGGAGCAAACCAGGGAGGTGGAGCAGCAGCACCAGGAGCAAACCA  800
Q  G  G  G  A  A  A  P  G  A  N  Q  G  G  G  A  A  A  P  G  A  N  Q  G  G  G  A  A  A  P  G  A  N  Q

GGAAGGTGGAGCAGCAGCACCAGGAGCAAACCAGGGAGGTGCAAAGCCAGCAGGAGGACAGGGACAAAATAATGAAGGTGCGAATAAGCCAGATGAAAAG  900
  E  G  G  A  A  A  P  G  A  N  Q  G  G  A  K  P  A  G  G  Q  G  Q  N  N  E  G  A  N  K  P  D  E  K

CATGTGAAAGAATACCTAGAGAAAATTAGATCTACCGTTGGCACCGAATGGACTCCATGCAGTGTAACCTGTGGAAAGGGTGTAAGAGTTAGAAGAAAAC  1000
 H  V  K  E  Y  L  E  K  I  R  S  T  V  G  T  E  W  T  P  C  S  V  T  C  G  K  G  V  R  V  R  R  K

TTAATGCAGGTGACAAAAAACCAGATAAGCTTACTCTGAATGACCTTGAGGCAGAAGTTTGTACAATGGATAAGTGCGCTGGCATATTTAACGTTGTGAG  1100
L  N  A  G  D  K  K  P  D  K  L  T  L  N  D  L  E  A  E  V  C  T  M  D  K  C  A  G  I  F  N  V  V  S

TAATTCATTAGGGCTAGTCATATTGTTAGTCCTAGCATTATTCAATTAA
  N  S  L  G  L  V  I  L  L  V  L  A  L  F  N  .
```

SG/EHI/WM16/Y09-85 (*P. knowlesi*; JQ219910)

```
ATGAAGAACTTCATTCTCTTGGCCGTCTCCTCCATCCTGCTGGTGGACTTGCTCCCCACATACTTCGAACATAATGTAGATCTCTCCAGGGCCATAAATG  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  L  P  T  Y  F  E  H  N  V  D  L  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAGTGTAGACACCAGTTCACTTGGCGCAGCACAGGTAAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAAGCCAAA  200
V  N  G  V  S  F  N  S  V  D  T  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  K  P  K

AGAAGGAGCTGATAAAGaAAAGAAAAAAGAAAAAGAAAAAGAAGAAGAACCAAAGAAGCCAAATGAAAATAAGCTGAAACAACCACAAGGAAATGGGGGG  300
  E  G  A  D  K  E  K  K  K  E  K  E  K  E  E  E  P  K  K  P  N  E  N  K  L  K  Q  P  Q  G  N  G  G

GCAGGTCAAGCACAACCAGAAGGAAATGGGGGGGCAGGTCAAGCACAACCGGAAGGAAATGGGGGGGCAGGTCAAGCACAACCAGAAGGAAATGGGGGGG  400
 A  G  Q  A  Q  P  E  G  N  G  G  A  G  Q  A  Q  P  E  G  N  G  G  A  G  Q  A  Q  P  E  G  N  G  G

CAGGTCAAGCACAACCGGAAGGAAATGGGGGGGCAGGTCAAGCACAACCGGAAGGAAATGGGGGGGCAGGTCAAGCACAACCGGAAGGAAATGGGGGGGC  500
A  G  Q  A  Q  P  E  G  N  G  G  A  G  Q  A  Q  P  E  G  N  G  G  A  G  Q  A  Q  P  E  G  N  G  G  A

AGGTCAAGCACAACCGGAAGGAAATGGGGGGGCAGGTCAAGCACAACCGGAAGGAAATGGGGGGGCAGGTCAAGCACAACCGGAAGGAAATGGGGGGGCA  600
  G  Q  A  Q  P  E  G  N  G  G  A  G  Q  A  Q  P  E  G  N  G  G  A  G  Q  A  Q  P  E  G  N  G  G  A

GGTCAAGCACAACCACAAGGAAATGGGGGGGCAGGTCAAGCACAACCACAAGGAAATGGGGGGGCAGGTCAAGCACAACCACAAGGAAATGGGGGGGCAG  700
 G  Q  A  Q  P  Q  G  N  G  G  A  G  Q  A  Q  P  Q  G  N  G  G  A  G  Q  A  Q  P  Q  G  N  G  G  A

GTCAAGCACAACCGGAAGGAAATAGGGAAGCTCCAGCACAACCACAAGGAAATGGGGGGGCAGGTCAAGCACAACCACAAAAAAACGAAGGAGGAAACGC  800
G  Q  A  Q  P  E  G  N  R  E  A  P  A  Q  P  Q  G  N  G  G  A  G  Q  A  Q  P  Q  K  N  E  G  G  N  A

AGGAGCACGGAAAGGACAGGGACAAAACAATCAGGGTGCGAATGCCCCAAATGAAAAAGTTGTAAATGATTACCTACAGAAAATTAGATCTAGCGTTACC  900
  G  A  R  K  G  Q  G  Q  N  N  Q  G  A  N  A  P  N  E  K  V  V  N  D  Y  L  Q  K  I  R  S  S  V  T

ACCGAGTGGACTCCATGCAGTGTAACCTGTGGAAATGGTGTAAGAATTAGAAGAAGAGCTCATGCAGATAAGAAAAAGGCAGAGGACCTTACTATGGATG  1000
 T  E  W  T  P  C  S  V  T  C  G  N  G  V  R  I  R  R  R  A  H  A  D  K  K  K  A  E  D  L  T  M  D

ACCTTGAAGTGGAAGCTTGTGTAATGGATAAGTGTGCTGGCATATTTAACGTTGTGAGTAATTCATTAGGGTTAGTCATATTGTTAGTCCTAGCATTATT  1100
D  L  E  V  E  A  C  V  M  D  K  C  A  G  I  F  N  V  V  S  N  S  L  G  L  V  I  L  L  V  L  A  L  F

CAATTAA
  N  .
```

SG/EHI/WM16/Y09-2 (*P. cynomolgi*; JQ219924)

```
ATGAAGAACTTCATTCTCTTGGCCGTTTCTTCCATCCTGTTGGTGGACTTGTTCCGCACACACTGGGGACATAATGTAGATTTCTCCAAGGCCATAAATT  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  F  R  T  H  W  G  H  N  V  D  F  S  K  A  I  N

TAAATGGAGTAAGCTTCAGTAATGTAGACGCCAGTTCACTTGGCGCAGCACAGGTAAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAACCCAAA  200
L  N  G  V  S  F  S  N  V  D  A  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  N  P  K

AAACGAGGAAGGAGCTGATAAACCAAAAAAAAAGGACGAAAAACAAGTAGAACCAAAAAAGCCACGTGAAAATAAGCTGAAACAACCACAAGCAGGAGGT  300
  N  E  E  G  A  D  K  P  K  K  K  D  E  K  Q  V  E  P  K  K  P  R  E  N  K  L  K  Q  P  Q  A  G  G

GATGCAGGAAATGCACAAGCCGGAGGAGATGCGGGAAATGCACAAGCAGGAGGTGATGCAGGAAATGCACAAGCAGGAGGTGCACAAGCAGGAGGTGATG  400
 D  A  G  N  A  Q  A  G  G  D  A  G  N  A  Q  A  G  G  D  A  G  N  A  Q  A  G  G  A  Q  A  G  G  D

CAGGAAATGCACAAGCAGGAGGTGCACAAGCAGGAGGTGATGCAGGAAATGCACAAGCAGGAGGTGCGCAAGCAGGAGGTGATGCAGGAAATGCACAAGC  500
A  G  N  A  Q  A  G  G  A  Q  A  G  G  D  A  G  N  A  Q  A  G  G  A  Q  A  G  G  D  A  G  N  A  Q  A

GGGAGGTGCGCAAGCAGGAGGAGATGCAGGAAATGCACAAGCGGGAGGTGCGCAAGCAGGAGGAGATGCGGGAAATGCACAAGCAGGAGGTGCACAAGCA  600
  G  G  A  Q  A  G  G  D  A  G  N  A  Q  A  G  G  A  Q  A  G  G  D  A  G  N  A  Q  A  G  G  A  Q  A

GGAGGTGATGCAGGAAATGCACAAGCAGGAGGTGCACAAGCAGGAGGTGATGCAGGAAATGCACAAGCAGGAGGTGCGCAAGCAGGAGGTGCACAAGCAG  700
 G  G  D  A  G  N  A  Q  A  G  G  A  Q  A  G  G  D  A  G  N  A  Q  A  G  G  A  Q  A  G  G  A  Q  A

GAGGTGCACAAGCAGGAGGTGCACAAGCAGGAGGTGCACAAGCAGGAGGTGCACAAGCAGGAGGTGCGCAAGCAGGAGGTGCACAAGCAGGAGGTGCACA  800
G  G  A  Q  A  G  G  A  Q  A  G  G  A  Q  A  G  G  A  Q  A  G  G  A  Q  A  G  G  A  Q  A  G  G  A  Q

AGCAGGAGGTGCACAAGCAGGAGGAGCAAATGCGGGAAATAAAAAAAGCAGGAGACGCAGGAGCAGGACAGGGACAAAATAATGAAGGTGCGAATATGCCA  900
  A  G  G  A  Q  A  G  G  A  N  A  G  N  K  K  A  G  D  A  G  A  G  Q  G  Q  N  N  E  G  A  N  M  P

AATGTAAAGCTTGTGAAAGAATACCTAGACAAAATTAGATCTACCATTGGCGTCGAGTGGAGTCCATACAGTGTAACCTGTGGAAAGGGTGTAAGAATGA  1000
 N  V  K  L  V  K  E  Y  L  D  K  I  R  S  T  I  G  V  E  W  S  P  Y  S  V  T  C  G  K  G  V  R  M

GAAGAAAAGTTAGTGCAGCTAACAAAAAACCAGAAGAGCTTGATGCGAATGACCTTGAGACTGAAGTTTGTACAATGGATAAGTGCGCTGGTATATTTAA  1100
R  R  K  V  S  A  A  N  K  K  P  E  E  L  D  A  N  D  L  E  T  E  V  C  T  M  D  K  C  A  G  I  F  N

CGTTGTGAGTAATTCATTAGGGCTAGTCATATTGTTAGTCCTAGCATTATTCAATTAA
  V  V  S  N  S  L  G  L  V  I  L  L  V  L  A  L  F  N  .
```

SG/EHI/WM16/Y09-34 (*P. cynomolgi*; JQ219925)

```
ATGAAGAACTTCATTCTCTTGGCCGTTTCTTCCATCCTGTTGGTGGACTTGTTCCGCACACACTGGGGACATAATGTAGATTTCTCCAAGGCCATAAATT  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  F  R  T  H  W  G  H  N  V  D  F  S  K  A  I  N

TAAATGGAGTAAGCTTCAGTAATGTAGACGCCAGTTCACTTGGCGCAGCACAGGTAAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAACCCAAA  200
L  N  G  V  S  F  S  N  V  D  A  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  N  P  K

AAACGAGGAAGGAGCTGATAAACCAAAAAAAAAGGACGAAAAACAAGTAGAACCAAAAAAGCCACGTGAAAATAAGCTGAAACAACCACAAGCAGGAGGT  300
  N  E  E  G  A  D  K  P  K  K  K  D  E  K  Q  V  E  P  K  K  P  R  E  N  K  L  K  Q  P  Q  A  G  G

GATGCAGGAAATGCACAAGCCGGAGGAGATGCGGGAAATGCACAAGCAGGAGGTGATGCAGGAAATGCACAAGCAGGAGGTGCACAAGCAGGAGGTGATG  400
 D  A  G  N  A  Q  A  G  G  D  A  G  N  A  Q  A  G  G  D  A  G  N  A  Q  A  G  G  A  Q  A  G  G  D

CAGGAAATGCACAAGCAGGAGGTGCACAAGCAGGAGGTGATGCAGGAAATGCACAAGCAGGAGGTGCGCAAGCAGGAGGTGATGCAGGAAATGCACAAGC  500
A  G  N  A  Q  A  G  G  A  Q  A  G  G  D  A  G  N  A  Q  A  G  G  A  Q  A  G  G  D  A  G  N  A  Q  A

GGGAGGTGCGCAAGCAGGAGGAGATGCAGGAAATGCACAAGCGGGAGGTGCGCAAGCAGGAGGAGATGCGGGAAATGCACAAGCAGGAGGTGCACAAGCA  600
  G  G  A  Q  A  G  G  D  A  G  N  A  Q  A  G  G  A  Q  A  G  G  D  A  G  N  A  Q  A  G  G  A  Q  A

GGAGGTGATGCAGGAAATGCACAAGCAGGAGGTGCACAAGCAGGAGGTGATGCAGGAAATGCACAAGCAGGAGGTGCGCAAGCAGGAGGTGCACAAGCAG  700
 G  G  D  A  G  N  A  Q  A  G  G  A  Q  A  G  G  D  A  G  N  A  Q  A  G  G  A  Q  A  G  G  A  Q  A

GAGGTGCACAAGCAGGAGGTGCACAAGCAGGAGGTGCACAAGCAGGAGGTGCACAAGCAGGAGGTGCGCAAGCAGGAGGTGCACAAGCAGGAGGTGCACA  800
G  G  A  Q  A  G  G  A  Q  A  G  G  A  Q  A  G  G  A  Q  A  G  G  A  Q  A  G  G  A  Q  A  G  G  A  Q

AGCAGGAGGTGCACAAGCAGGAGGAGCAAATGCGGGAAATAAAAAAAGCAGGAGACGCAGGAGCAGGACAGGGACAAAATAATGAAGGTGCGAATATGCCA  900
  A  G  G  A  Q  A  G  G  A  N  A  G  N  K  K  A  G  D  A  G  A  G  Q  G  Q  N  N  E  G  A  N  M  P

AATGTAAAGCTTGTGAAAGAATACCTAGACAAAATTAGATCTACCATTGGCGTCGAGTGGAGTCCATGCAGTGTAACCTGTGGAAAGGGTGTAAGAATGA  1000
 N  V  K  L  V  K  E  Y  L  D  K  I  R  S  T  I  G  V  E  W  S  P  C  S  V  T  C  G  K  G  V  R  M

GAAGAAAAGTTAGTGCAGCTAACAAAAAACCAGAAGAGCTTGATGCGAATGACCTTGAGACTGAAGTTTGTACAATGGATAAGTGCGCTGGTATATTTAA  1100
R  R  K  V  S  A  A  N  K  K  P  E  E  L  D  A  N  D  L  E  T  E  V  C  T  M  D  K  C  A  G  I  F  N

CGTTGTGAGTAATTCATTAGGGCTAGTCATATTGTTAGTCCTAGCATTATTCAATTAA
  V  V  S  N  S  L  G  L  V  I  L  L  V  L  A  L  F  N  .
```

SG/EHI/WM17/Y09-4 *(P. knowlesi*; JQ219911)

```
ATGAAGAACTTCATTCTCTTGGCCGTCTCCTCCATCCTGCTGGTGGACTTGCTCCCCACACACTTCGAACATAATGTAGATCTCTCCAGGGCCATAAATG  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  L  P  T  H  F  E  H  N  V  D  L  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAATGCAGACACCAGTTCACTTGGCGCAGCACAGGTGAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAAGCCAAA  200
V  N  G  V  S  F  N  N  A  D  T  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  K  P  K

AGAAGGAGCTGATAAAGAAAAGAAAAAAGAAAAAGAAGAAGAACCAAAGAAGCCAAATGAAAATAAGCTGAAACAACCGGATGCAGTACCAGGGGGCGAA  300
  E  G  A  D  K  E  K  K  K  E  K  E  E  E  P  K  K  P  N  E  N  K  L  K  Q  P  D  A  V  P  G  G  E

GAACCAGCACCAGGAAGGGAACAGCCAGCACCAGGAAGGGAGGAACCAGCACCAGGAAGGGAACAGCCAGCACCAGGAAGGGAAGAACCAGCACCAGGAA  400
 E  P  A  P  G  R  E  Q  P  A  P  G  R  E  E  P  A  P  G  R  E  Q  P  A  P  G  R  E  E  P  A  P  G

GGGAACAGCCAGCACCAGGAAGGGAAGAACCAGCGCCAGGAAGGGAACAGCCAGCACCAGGAAGGGAACAGCCAGCACCAGGAAGGGAGGAACCAGCACC  500
R  E  Q  P  A  P  G  R  E  E  P  A  P  G  R  E  Q  P  A  P  G  R  E  Q  P  A  P  G  R  E  E  P  A  P

AGGAAGGGAACAACCAGCACCAGGAAGGGAGGAACCAGCACCAGGAAGGGAACAACCAGCACCAGGAAGGGAGCAGCCAGCACCAGGAAGGGAACAGCCA  600
  G  R  E  Q  P  A  P  G  R  E  E  P  A  P  G  R  E  Q  P  A  P  G  R  E  Q  P  A  P  G  R  E  Q  P

GCACCGGGGGGTGAACAACCAGCACCAGGAAGGGAACAGCCAGCACCGGGTGGTGAACAACCAGCACCAGCACCAGGAGCAGGTGCGGGAGATGGAGCAC  700
 A  P  G  G  E  Q  P  A  P  G  R  E  Q  P  A  P  G  G  E  Q  P  A  P  A  P  G  A  G  A  G  D  G  A

GAGGAGGAAACGCAGGGGCAGGGAAAGGACAGGGACAAAACAATCAGGGTGCAAATGTCCCAAATGAAAAAGTTGTGAATGATTACCTACACAAAATTAG  800
R  G  G  N  A  G  A  G  K  G  Q  G  Q  N  N  Q  G  A  N  V  P  N  E  K  V  V  N  D  Y  L  H  K  I  R

ATCTAGCGTTACCACCGAGTGGACTCCATGCAGTGTAACCTGTGGAAATGGTGTAAGAATTAGAAGAAAAGGTCATGCAGGTAATAAAAAGGCAGAGGAC  900
  S  S  V  T  T  E  W  T  P  C  S  V  T  C  G  N  G  V  R  I  R  R  K  G  H  A  G  N  K  K  A  E  D

CTTACTATGGATGACCTTGAGGTGGAAGCTTGTGTAATGGATAAGTGCGCTGGCATATTTAACGTTGTGAGTAATTCATTAGGCTTAGTCATATTGTTAG  1000
 L  T  M  D  D  L  E  V  E  A  C  V  M  D  K  C  A  G  I  F  N  V  V  S  N  S  L  G  L  V  I  L  L

TCCTAGCATTATTCAATTAA
V  L  A  L  F  N  .
```

SG/EHI/WM17/Y09-30 *(P. knowlesi*; JQ219912)

```
ATGAAGAACTTCATTCTCTTGGCCGTCTCCTCCATCCTGCTGGTGGACTTGCTCCCCACACACTTCGAACATAATGTAGATCTCTCCAGGGCCATAAATG  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  L  P  T  H  F  E  H  N  V  D  L  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAATGTAGACACCAGTTCACTTGGCGCAGCACAGGTGAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAAGCCAAA  200
V  N  G  V  S  F  N  N  V  D  T  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  K  P  K

AGAAGGAGCTGATAAAGAAAAGAAAAAAGAAAAAGAAGAAGAACCAAAGAAGCCAAATGAAAATAAGCTGAAACAACCGGATGCAGTACCAGGGGGCGAA  300
  E  G  A  D  K  E  K  K  K  E  K  E  E  E  P  K  K  P  N  E  N  K  L  K  Q  P  D  A  V  P  G  G  E

GAACCAGCACCAGGAAGGGAACAGCCAGCACCAGGAAGGGAGGAACCAGCACCAGGAAGGGAACAGCCAGCACCAGGAAGGGAAGAACCAGCACCAGGAA  400
 E  P  A  P  G  R  E  Q  P  A  P  G  R  E  E  P  A  P  G  R  E  Q  P  A  P  G  R  E  E  P  A  P  G

GGGAACAGCCAGCACCAGGAAGGGAAGAACCAGCGCCAGGAAGGGAACAGCCAGCACCAGGAAGGGAACAGCCAGCACCAGGAAGGGAGGAACCAGCACC  500
R  E  Q  P  A  P  G  R  E  E  P  A  P  G  R  E  Q  P  A  P  G  R  E  Q  P  A  P  G  R  E  E  P  A  P

AGGAAGGGAACAACCAGCACCAGGAAGGGAGGAACCAGCACCAGGAAGGGAACAACCAGCACCAGGAAGGGAACAGCCAGCACCAGGAAGGGAACAGCCA  600
  G  R  E  Q  P  A  P  G  R  E  E  P  A  P  G  R  E  Q  P  A  P  G  R  E  Q  P  A  P  G  R  E  Q  P

GCACCGGGGGGTGAACAACCAGCACCAGGAAGGGAACAGCCAGCACCGGGTGGTGAACAACCAGCACCAGCACCAGGAGCAGGTGCGGGAGATGGAGCAC  700
 A  P  G  G  E  Q  P  A  P  G  R  E  Q  P  A  P  G  G  E  Q  P  A  P  A  P  G  A  G  A  G  D  G  A

GAGGAGGAAACGCAGGGGCAGGGAAAGGACAGGGACAAAACAATCAGGGTGCAAATGTCCCAAATGAAAAAGTTGTGAATGATTACCTACACAAAATTAG  800
R  G  G  N  A  G  A  G  K  G  Q  G  Q  N  N  Q  G  A  N  V  P  N  E  K  V  V  N  D  Y  L  H  K  I  R

ATCTAGCGTTACCACCGAGTGGACTCCATGCAGTGTAACCTGTGGAAATGGTGTAAGAATTAGAAGAAAAGGTCATGCAGGTAATAAAAAGGCAGAGGAC  900
  S  S  V  T  T  E  W  T  P  C  S  V  T  C  G  N  G  V  R  I  R  R  K  G  H  A  G  N  K  K  A  E  D

CTTACTATGGATGACCTTGAGGTGGAAGCTTGTGTAATGGATAAGTGCGCTGGCATATTTAACGTTGTGAGTAATTCATTAGGCTTAGTCATATTGTTAG  1000
 L  T  M  D  D  L  E  V  E  A  C  V  M  D  K  C  A  G  I  F  N  V  V  S  N  S  L  G  L  V  I  L  L

TCCTAGCATTATTCAATTAA
V  L  A  L  F  N  .
```

SG/EHI/WM18/Y09-24 *(P. cynomolgi*; JQ219926)

```
ATGAAGAACTTCATTCTCTTAGCCGTTTCTTCCATCCTGTTGGTGGACTTGTTTCCCACAAACTGCGGGCATAATGTAGATTTCTCCAGGGCCATAAATT  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  F  P  T  N  C  G  H  N  V  D  F  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAATGTAGACGCCAGTTCCCTTGGCGCAGCACAGGTAAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAACCCAAA  200
L  N  G  V  S  F  N  N  V  D  A  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  N  P  K

AAACGAGGAAGGAGCTGATAAACAAAAAAAGGACGAAAAAAAAGTAGAACCAAAAAAGCCACGTGAAAATAAGCTGAAACAACCAGACGGAAATAATGCA  300
  N  E  E  G  A  D  K  Q  K  K  D  E  K  K  V  E  P  K  K  P  R  E  N  K  L  K  Q  P  D  G  N  N  A

GCTGATGGAGGTGTACAACCGCCAGCAGGAGGAGGAAATAATGCAGCTGATGGAGGTGTACAACCACCAGCAGGAGGAGGAAATAATGCAGCTGATGGAG  400
 A  D  G  G  V  Q  P  P  A  G  G  G  N  N  A  A  D  G  G  V  Q  P  P  A  G  G  G  N  N  A  A  D  G

GTGTACAACCACCAGCAGGAGGAGGAAATAATGCAGCTGATGGAGGTGTACAACCACCAGCAGGAGGAGGAAATAATGCAGCTGATGGAGGTGTACAACC  500
G  V  Q  P  P  A  G  G  G  N  N  A  A  D  G  G  V  Q  P  P  A  G  G  G  N  N  A  A  D  G  G  V  Q  P

ACCAGCAGGAGGAGGAAATAATGCAGCTGATGGAGGCGTACAACCACCAGCAGGAGGAGGAAATAATGCAGCTGATGGAGGTGTACAACCGCCAGCAGGA  600
  P  A  G  G  G  N  N  A  A  D  G  G  V  Q  P  P  A  G  G  G  N  N  A  A  D  G  G  V  Q  P  P  A  G

GGAGGGAATAATGCAGCTGATGGAGGTGTACAACCACCAGCAGGAGGAGGAAATAATGCAGCTGATGGAGGTGTACAACCACCAGCAAGAGGAGGAAATA  700
 G  G  N  N  A  A  D  G  G  V  Q  P  P  A  G  G  G  N  N  A  A  D  G  G  V  Q  P  P  A  R  G  G  N

ATGCAGCTGATGGAGGTGCACAACCGCCAGCAGGAGGAGGAAATAATGCAGCTGATGGAGGTGTACAACCACCAGCAGCAGGAGGAAATAGGGCAAATAA  800
N  A  A  D  G  G  A  Q  P  P  A  G  G  G  N  N  A  A  D  G  G  V  Q  P  P  A  A  G  G  N  R  A  N  K

AAAAGCAGGAGAAGCAGGTGGAAACGCAGGAGCAGGACAGGGACAAAATAATGAAGGTGCGAATATGCCAAATGTAAAGCTTGTGCAAGAATACCTAGAC  900
  K  A  G  E  A  G  G  N  A  G  A  G  Q  G  Q  N  N  E  G  A  N  M  P  N  V  K  L  V  Q  E  Y  L  D

AAAATTAGATCTACCATTGGCGTCGAGTGGAGTCCATGCAGTGTAACCTGTGGAAAGGGTGTAAGAATGAGAAGAAAAGTTAGTGCAGCTAACAAAAAAC  1000
 K  I  R  S  T  I  G  V  E  W  S  P  C  S  V  T  C  G  K  G  V  R  M  R  R  K  V  S  A  A  N  K  K

CAGAAGAGCTTGATGCGAATGACCTTGAGACTGAAGTTTGTACAATGGATAAGTGCGCTGGTATATTTAACGTTGTGAGTAATTCATTAGGGCTAGTCAT  1100
P  E  E  L  D  A  N  D  L  E  T  E  V  C  T  M  D  K  C  A  G  I  F  N  V  V  S  N  S  L  G  L  V  I

ATTGTTAGTCCTAGCATTATTCAATTAA
  L  L  V  L  A  L  F  N  .
```

SG/EHI/WM18/Y09-92 *(P. fieldi*; JQ219934)

```
ATGAAGAACTTCATTCTCTTGGCCGTTTCTTCCATCCTGTTGGTGGACTTGTTCCCCACACACTGCGGGCATAATGTAGATCTCTCCAGGGCAATAAATT  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  F  P  T  H  C  G  H  N  V  D  L  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAATGTAGACGCCAGTTCACTTGGCGCAGCACAGGTAAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAACCCAAA  200
L  N  G  V  S  F  N  N  V  D  A  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  N  P  K

AGACGATGAAAAAGCTGATAAACCAAAAAAAAAGGACGAAAAAAAAGTAGAACCAAAAAAGCCACATGAAAATAAGCTGAAACAACCAGTCCCAGGAGCA  300
  D  D  E  K  A  D  K  P  K  K  K  D  E  K  K  V  E  P  K  K  P  H  E  N  K  L  K  Q  P  V  P  G  A

AATCAGGAAGGCGGAGCAGCAGCCCCAGGAGCAAATCAGGAAGGCGGAGCAGCAGCCCCAGGAGCAAATCAAGAAGGTGGAGCAGCAGCCCCAGGTGCAA  400
 N  Q  E  G  G  A  A  A  P  G  A  N  Q  E  G  G  A  A  A  P  G  A  N  Q  E  G  G  A  A  A  P  G  A

ACCAGGAAGGTGGAGCAGCAGCCCCAGGAGCAAACCAGGAAGGTGGAGCAGCAGCCCCAGGAGCAAATCAGGAAGGCGGAGCAGCAGCCCCAGGAGCAAA  500
N  Q  E  G  G  A  A  A  P  G  A  N  Q  E  G  G  A  A  A  P  G  A  N  Q  E  G  G  A  A  A  P  G  A  N

TCAGGAAGGTGGAGCAGCAGCCCCAGGTGCAAACCAGGAAGGTGGAGCAGCAGCCCCAGGAGCAAATCAGGGAGGTGGAGCAGCAGCACCAGGAGCAAAC  600
  Q  E  G  G  A  A  A  P  G  A  N  Q  E  G  G  A  A  A  P  G  A  N  Q  G  G  G  A  A  A  P  G  A  N

CAGGGAGGTGGAGCAGCAGCACCAGGAGCAAACCAGGAAGGTGGAGCAGCAGCACCAGGAGCAAACCAGGGAGGTGCAAAGCCAGCAGGAGGACAGGGAC  700
 Q  G  G  G  A  A  A  P  G  A  N  Q  E  G  G  A  A  A  P  G  A  N  Q  G  G  A  K  P  A  G  G  Q  G

AAAATAATGAAGGTGCGAATAAGCCAGATGAAAAGCATGTGAAAGAATACCTAGAGAAAATTAGATCTACCGTTGGCACCGAATGGACTCCATGCAGTGT  800
Q  N  N  E  G  A  N  K  P  D  E  K  H  V  K  E  Y  L  E  K  I  R  S  T  V  G  T  E  W  T  P  C  S  V

AACCTGTGGAAAGGGTGTAAGAGTTAGAAGAAAACTTAATGCAGGTGACAAAAAACCAGATAAGCTTACTCTGAATGACCTTGAGGCAGAAGTTTGTACA  900
  T  C  G  K  G  V  R  V  R  R  K  L  N  A  G  D  K  K  P  D  K  L  T  L  N  D  L  E  A  E  V  C  T

ATGGATAAGTGCGCTGGCATATTTAACGTTGTGAGTAATTCATTAGGGCTAGTCATATTGTTAGTCCTAGCATTATTCAATTAA
 M  D  K  C  A  G  I  F  N  V  V  S  N  S  L  G  L  V  I  L  L  V  L  A  L  F  N  .
```

SG/EHI/WM26/Y09-1 *(P. knowlesi*; JQ219913)

```
ATGAAGAACTTCATTCTCTTGGCCGTCTCCTCCATCCTGCTGGTGGACTTGCTCCCCACACACTTCGAACATAATGTAGATCTCTCCAGGGCCATAAATG  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  L  P  T  H  F  E  H  N  V  D  L  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAATGTAGACACCAGTTCACTTGGCGCAGCACAGGTGAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAAGCCAAA  200
V  N  G  V  S  F  N  N  V  D  T  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  K  P  K

AGAAGGAGCTGATAAAGAAAAGAAAAAAGAAAAAGAAGAAGAACCAAAGAAGCCAAATGAAAATAAGCTGAAACAACCGGATGCAGTACCAGGGGGCGAA  300
  E  G  A  D  K  E  K  K  K  E  K  E  E  E  P  K  K  P  N  E  N  K  L  K  Q  P  D  A  V  P  G  G  E

GAACCAGCACCAGGAAGGGAACAGCCAGCACCAGGAAGGGAGGAACCAGCACCAGGAAGGGAACAGCCAGCACCAGGAAGGGAAGAACCAGCACCAGGAA  400
 E  P  A  P  G  R  E  Q  P  A  P  G  R  E  E  P  A  P  G  R  E  Q  P  A  P  G  R  E  E  P  A  P  G

GGGAACAGCCAGCACCAGGAAGGGAAGAACCAGCGCCAGGAAGGGAACAGCCAGCACCAGGAAGGGAACAGCCAGCACCAGGAAGGGAGGAACCAGCACC  500
R  E  Q  P  A  P  G  R  E  E  P  A  P  G  R  E  Q  P  A  P  G  R  E  Q  P  A  P  G  R  E  E  P  A  P

AGGAAGGGAACAACCAGCACCAGGAAGGGAGGAACCAGCACCAGGAAGGGAACAACCAGCACCAGGAAGGGAACAGCCAGCACCAGGAAGGGAACAGCCA  600
  G  R  E  Q  P  A  P  G  R  E  E  P  A  P  G  R  E  Q  P  A  P  G  R  E  Q  P  A  P  G  R  E  Q  P

GCACCGGGGGGTGAACAACCAGCACCAGGAAGGGAACAGCCAGCACCGGGTGGTGAACAACCAGCACCAGCACCAGGAGCAGGTGCGGGAGATGGAGCAC  700
 A  P  G  G  E  Q  P  A  P  G  R  E  Q  P  A  P  G  G  E  Q  P  A  P  A  P  G  A  G  A  G  D  G  A

GAGGAGGAAACGCAGGGGCAGGGAAAGGACAGGGACAAAACAATCAGGGTGCAAATGTCCCAAATGAAAAAGTTGTGAATGATTACCTACACAAAATTAG  800
R  G  G  N  A  G  A  G  K  G  Q  G  Q  N  N  Q  G  A  N  V  P  N  E  K  V  V  N  D  Y  L  H  K  I  R

ATCTAGCGTTACCACCGAGTGGACTCCATGCAGTGTAACCTGTGGAAATGGTGTAAGAATTAGAAGAAAAGGTCATGCAGGTAATAAAAAGGCAGAGGAC  900
  S  S  V  T  T  E  W  T  P  C  S  V  T  C  G  N  G  V  R  I  R  R  K  G  H  A  G  N  K  K  A  E  D

CTTACTATGGATGACCTTGAGGTGGAAGCTTGTGTAATGGATAAGTGCGCTGGCATATTTAACGTTGTGAGTAATTCATTAGGCTTAGTCATATTGTTAG  1000
 L  T  M  D  D  L  E  V  E  A  C  V  M  D  K  C  A  G  I  F  N  V  V  S  N  S  L  G  L  V  I  L  L

TCCTAGCATTATTCAATTAA
V  L  A  L  F  N  .
```

SG/EHI/WM26/Y09-13*(P. knowlesi*; JQ219914)

```
ATGAAGAACTTCATTCTCTTGGCCGTCTCCTCCATCCTGCTGGTGGACTTGCTCCCCACACACTTCGAACATAATGTAGATCTCTCCAGGGCCATAAATG  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  L  P  T  H  F  E  H  N  V  D  L  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAATGTAGACACCAGTTCACTTGGCGCAGCACAGGTGAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAAGCCAAA  200
V  N  G  V  S  F  N  N  V  D  T  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  K  P  K

AGAAGGAGCTGATAAAGAAAAGAAAAAAGAAAAAGAAGAAGAACCAAAGAAGCCAAATGAAAATAAGCTGAAACAACCGGATGCAGTACCAGGGGGCGAA  300
  E  G  A  D  K  E  K  K  K  E  K  E  E  E  P  K  K  P  N  E  N  K  L  K  Q  P  D  A  V  P  G  G  E

GAACCAGCACCAGGAAGGGAACAGCCAGCACCAGGAAGGGAGGAACCAGCACCAGGAAGGGAACAGCCAGCACCAGGAAGGGAAGAACCAGCACCAGGAA  400
 E  P  A  P  G  R  E  Q  P  A  P  G  R  E  E  P  A  P  G  R  E  Q  P  A  P  G  R  E  E  P  A  P  G

GGGAACAGCCAGCACCAGGAAGGGAAGAACCAGCGCCAGGAAGGGAACAGCCAGCACCAGGAAGGGAACAGCCAGCACCAGGAAGGGAGGAGCCAGCACC  500
R  E  Q  P  A  P  G  R  E  E  P  A  P  G  R  E  Q  P  A  P  G  R  E  Q  P  A  P  G  R  E  E  P  A  P

AGGAAGGGAACAACCAGCACCAGGAAGGGAGGAACCAGCACCAGGAAGGGAACAACCAGCACCAGGAAGGGAACAGCCAGCACCAGGAAGGGAACAGCCA  600
  G  R  E  Q  P  A  P  G  R  E  E  P  A  P  G  R  E  Q  P  A  P  G  R  E  Q  P  A  P  G  R  E  Q  P

GCACCGGGGGGTGAACAACCAGCACCAGGAAGGGAACAGCCAGCACCGGGTGGTGAACAACCAGCACCAGCACCAGGAGCAGGTGCGGGAGATGGAGCAC  700
 A  P  G  G  E  Q  P  A  P  G  R  E  Q  P  A  P  G  G  E  Q  P  A  P  A  P  G  A  G  A  G  D  G  A

GAGGAGGAAACGCAGGGGCAGGGAAAGGACAGGGACAAAACAATCAGGGTGCAAATGTCCCAAATGAAAAAGTTGTGAATGATTACCTACACAAAATTAG  800
R  G  G  N  A  G  A  G  K  G  Q  G  Q  N  N  Q  G  A  N  V  P  N  E  K  V  V  N  D  Y  L  H  K  I  R

ATCTAGCGTTACCACCGAGTGGACTCCATGCAGTGTAACCTGTGGAAATGGTGTAAGAATTAGAAGAAAAGGTCATGCAGGTAATAAAAAGGCAGAGGAC  900
  S  S  V  T  T  E  W  T  P  C  S  V  T  C  G  N  G  V  R  I  R  R  K  G  H  A  G  N  K  K  A  E  D

CTTACTATGGATGACCTTGAGGTGGAAGCTTGTGTAATGGATAAGTGCGCTGGCATATTTAACGTTGTGAGTAATTCATTAGGCTTAGTCATATTGTTAG  1000
 L  T  M  D  D  L  E  V  E  A  C  V  M  D  K  C  A  G  I  F  N  V  V  S  N  S  L  G  L  V  I  L  L

TCCTAGCATTATTCAATTAA
V  L  A  L  F  N  .
```

SG/EHI/WM26/Y09-47*(P. knowlesi*; JQ219915)

```
ATGAAGAACTTCATTCTCTTGGCCGTCTCCTCCATCCTGCTGGTGGACTTGCTCCCCACACACTTCGAACATAATGTAGATCTCTCCAGGGCCATAAATG  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  L  P  T  H  F  E  H  N  V  D  L  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAATGTAGACACCAGTTCACTTGGCGCAGCACAGGTGAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAAGCCAAA  200
V  N  G  V  S  F  N  N  V  D  T  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  K  P  K

AGAAGGAGCTGATAAAGAAAAGAAAAAAGAAAAAGAAGAAGAACCAAAGAAGCCAAATGAAAATAAGCTGAAACAACCGGATGCAGTACCAGGGGGCGAA  300
  E  G  A  D  K  E  K  K  K  E  K  E  E  E  P  K  K  P  N  E  N  K  L  K  Q  P  D  A  V  P  G  G  E

GAACCAGCACCAGGAAGGGAACAGCCAGCACCAGGAAGGGAGGAACCAGCACCAGGAAGGGAACAGCCAGCACCAGGAAGGGAAGAACCAGCACCAGGAA  400
 E  P  A  P  G  R  E  Q  P  A  P  G  R  E  E  P  A  P  G  R  E  Q  P  A  P  G  R  E  E  P  A  P  G

GGGAACAGCCAGCACCAGGAAGGGAAGAACCAGCGCCAGGAAGGGAACAGCCAGCACCAGGAAGGGAACAGCCAGCACCAGGAAGGGAGGAACCAGCACC  500
R  E  Q  P  A  P  G  R  E  E  P  A  P  G  R  E  Q  P  A  P  G  R  E  Q  P  A  P  G  R  E  E  P  A  P

AGGAAGGGAACAACCAGCACCAGGAAGGGAGGAACCAGCACCAGGAAGGGAACAACCAGCACCAGGAAGGGAACAGCCAGCACCAGGAAGGGAACAGCCA  600
  G  R  E  Q  P  A  P  G  R  E  E  P  A  P  G  R  E  Q  P  A  P  G  R  E  Q  P  A  P  G  R  E  Q  P

GCACCGGGGGGTGAACAACCAGCACCAGGAAGGGAACAGCCAGCACCGGGTGGTGAACAACCAGCACCAGCACCAGGAGCAGGTGCGGGAGATGGAGCAC  700
 A  P  G  G  E  Q  P  A  P  G  R  E  Q  P  A  P  G  G  E  Q  P  A  P  A  P  G  A  G  A  G  D  G  A

GAGGAGGAAACGCAGGGGCAGGGAAAGGACAGGGACAAAACAATCAGGGTGCAAATGTCCCGAATGAAAAAGTTGTGAATGATTACCTACACAAAATTAG  800
R  G  G  N  A  G  A  G  K  G  Q  G  Q  N  N  Q  G  A  N  V  P  N  E  K  V  V  N  D  Y  L  H  K  I  R

ATCTAGCGTTACCACCGAGTGGACTCCATGCAGTGTAACCTGTGGAAATGGTGTAAGAATTAGAAGAAAAGGTCATGCAGGTAATAAAAAGGCAGAGGAC  900
  S  S  V  T  T  E  W  T  P  C  S  V  T  C  G  N  G  V  R  I  R  R  K  G  H  A  G  N  K  K  A  E  D

CTTACTATGGATGACCTTGAGGTGGAAGCTTGTGTAATGGATAAGTGCGCTGGCATATTTAACGTTGTGAGTAATTCATTAGGCTTAGTCATATTGTTAG  1000
 L  T  M  D  D  L  E  V  E  A  C  V  M  D  K  C  A  G  I  F  N  V  V  S  N  S  L  G  L  V  I  L  L

TCCTAGCATTATTCAATTAA
V  L  A  L  F  N  .
```

SG/EHI/WM26/Y09-60 *(P. knowlesi*; JQ219916)

```
ATGAAGAACTTCATTCTCTTGGCCGTCTCCTCCATCCTGCTGGTGGACTTGCTCCCCACACACTTCGAACATAATGTAGATCTCTCCAGGGCCATAAATG  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  L  P  T  H  F  E  H  N  V  D  L  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAATGTAGACACCAGTTCACTTGGCGCAGCACGGGTGAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAAGCCAAA  200
V  N  G  V  S  F  N  N  V  D  T  S  S  L  G  A  A  R  V  R  Q  S  A  S  R  G  R  G  L  G  E  K  P  K

AGAAGGAGCTGATAAAGAAAAGAAAAAAGAAAAAGAAAAAGAAAAAGAAGAACCAAAGAAGCCAAATGAAAATAAGCTGAAACAACCGGAACAACCAGCA  300
  E  G  A  D  K  E  K  K  K  E  K  E  K  E  K  E  E  P  K  K  P  N  E  N  K  L  K  Q  P  E  Q  P  A

GCAGGAGCAGGGGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAAC  400
 A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E

AACCAGCGGCAGGAGCAGGGGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCAGCAGGAGCAGG  500
Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G

AGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCAGCAGGAGCAGGGGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCACCA  600
  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  P

GCACCAAGGAGGGAACAACCAGCAGCAGGAGCAGGGGGCGAACAACCAGCACCAGCACCAAGGAGGGAACAACCAGCACCAGGAGCAGGTGCGGGAGATG  700
 A  P  R  R  E  Q  P  A  A  G  A  G  G  E  Q  P  A  P  A  P  R  R  E  Q  P  A  P  G  A  G  A  G  D

GAGCACGAGGAGGAAACGCAGGGGCAGGTAAAGGACAGGGACAAAACAATCAGGGTGCGAATGTCCCAAATGAAAAAGTTGTGAATGATTACCTACACAA  800
G  A  R  G  G  N  A  G  A  G  K  G  Q  G  Q  N  N  Q  G  A  N  V  P  N  E  K  V  V  N  D  Y  L  H  K

AATTAGATCTAGCGTTACCACCGAGTGGACTCCATGCAGTGTAACCTGTGGAAATGGTGTAAGAATTAGAAGAAGACAGAATGCTGGTAATAAAAAGGCA  900
  I  R  S  S  V  T  T  E  W  T  P  C  S  V  T  C  G  N  G  V  R  I  R  R  R  Q  N  A  G  N  K  K  A

GAGGACCTTACTATGGACGACCTTGAGGTGGAAGCTTGTGTAATGGATAAGTGCGCTGGCATATTTAACGTTGTGAGTAATTCATTAGGCTTAGTCATAT  1000
 E  D  L  T  M  D  D  L  E  V  E  A  C  V  M  D  K  C  A  G  I  F  N  V  V  S  N  S  L  G  L  V  I

TGTTAGTCTTAGCATTATTCAATTAA
L  L  V  L  A  L  F  N  .
```

SG/EHI/WM26/Y09-98 *(P. knowlesi*; JQ219917)

```
ATGAAGAACTTCATTCTCTTGGCCGTCTCCTCCATCCTGCTGGTGGACTTGCTCCCCACACACTTCGAACATAATGTAGATCTCTCCAGGGCCATAAATG  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  L  P  T  H  F  E  H  N  V  D  L  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAATGTAGACACCAGTTCACTTGGCGCAGCACAGGTGAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAAGCCAAA  200
V  N  G  V  S  F  N  N  V  D  T  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  K  P  K

AGAAGGAGCTGATAAAGAAAAGAAAAAAGAAAAAGAAGAAGAACCAAAGAAGCCAAATGAAAATAAGCTGAAACAACCGGATACAGTACCAGGGGGCGAA  300
  E  G  A  D  K  E  K  K  K  E  K  E  E  E  P  K  K  P  N  E  N  K  L  K  Q  P  D  T  V  P  G  G  E

GAACCAGCACCAGGAAGGGAACAGCCAGCACCAGGAAGGGAGGAACCAGCACCAGGAAGGGAACAGCCAGCACCAGGAAGGGAAGAACCAGCACCAGGAA  400
 E  P  A  P  G  R  E  Q  P  A  P  G  R  E  E  P  A  P  G  R  E  Q  P  A  P  G  R  E  E  P  A  P  G

GGGAACAGCCAGCACCAGGAAGGGAAGAACCAGCGCCAGGAAGGGAACAGCCAGCACCAGGAAGGGAACAGCCAGCACCAGGAAGGGAGGAACCAGCACC  500
R  E  Q  P  A  P  G  R  E  E  P  A  P  G  R  E  Q  P  A  P  G  R  E  Q  P  A  P  G  R  E  E  P  A  P

AGGAAGGGAACAACCAGCACCAGGAAGGGAGGAACCAGCACCAGCACCAGGAGCAGGTGCGGGAGATGGAGCACGAGGAGGAAACGCAGGGGCAGGGAAA  600
  G  R  E  Q  P  A  P  G  R  E  E  P  A  P  A  P  G  A  G  A  G  D  G  A  R  G  G  N  A  G  A  G  K

GGACAGGGACAAAACAATCAGGGTGCAAATGTCCCAAATGAAAAAGTTGTGAATGATTACCTACACAAAATTAGATCTAGCGTTACCACCGAGTGGACTC  700
 G  Q  G  Q  N  N  Q  G  A  N  V  P  N  E  K  V  V  N  D  Y  L  H  K  I  R  S  S  V  T  T  E  W  T

CATGCAGTGTAACCTGTGGAAATGGTGTAAGAATTAGAAGAAAAGGTCATGCAGGTAATAAAAAGGCAGAGGACCTTACTATGGATGACCTTGAGGTGGA  800
P  C  S  V  T  C  G  N  G  V  R  I  R  R  K  G  H  A  G  N  K  K  A  E  D  L  T  M  D  D  L  E  V  E

AGCTTGTGTAATGGATAAGTGCGCTGGCATATTTAACGTTGTGAGTAATTCATTAGGCTTAGTCATATTGTTAGTCCTAGCATTATTCAATTAA
  A  C  V  M  D  K  C  A  G  I  F  N  V  V  S  N  S  L  G  L  V  I  L  L  V  L  A  L  F  N  .
```

SG/EHI/WM26/Y09-123 *(P. knowlesi*; JQ219918)

```
ATGAAGAACTTCATTCTCTTGGCCGTCTCCTCCATCCTGCTGGTGGACTTGCTCCCCACACACTTCGAACACAATGTAGATCTCTCCAGGGCCATAAATG  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  L  P  T  H  F  E  H  N  V  D  L  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAATGTAGACACCAGTTCACTTGGCGCAGCACAGGTGAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAAGCCAAA  200
V  N  G  V  S  F  N  N  V  D  T  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  K  P  K

AGAAGGAGCTGATAAAGAAAAGAAAAAAGAAAAAGAAGAAGAACCAAAGAAGCCAAATGAAAATAAGCTGAAACAACCGGATGCAGTACCAGGGGGCGAA  300
  E  G  A  D  K  E  K  K  K  E  K  E  E  E  P  K  K  P  N  E  N  K  L  K  Q  P  D  A  V  P  G  G  E

GAACCAGCACCAGGAAGGGAACAGCCAGCACCAGGAAGGGAGGAACCAGCACCAGGAAGGGAACAGCCAGCACCAGGAAGGGAAGAACCAGCACCAGGAA  400
 E  P  A  P  G  R  E  Q  P  A  P  G  R  E  E  P  A  P  G  R  E  Q  P  A  P  G  R  E  E  P  A  P  G

GGGAACAGCCAGCACCAGGAAGGGAAGAACCAGCGCCAGGAAGGGAACAGCCAGCACCAGGAAGGGAACAGCCAGCACCAGGAAGGGAGGAACCAGCACC  500
R  E  Q  P  A  P  G  R  E  E  P  A  P  G  R  E  Q  P  A  P  G  R  E  Q  P  A  P  G  R  E  E  P  A  P

AGGAAGGGAACAACCAGCACCAGGAAGGGAGGAACCAGCACCAGGAAGGGAACAACCAGCACCAGGAAGGGAACAGCCAGCACCAGGAAGGGAACAGCCA  600
  G  R  E  Q  P  A  P  G  R  E  E  P  A  P  G  R  E  Q  P  A  P  G  R  E  Q  P  A  P  G  R  E  Q  P

GCACCGGGGGGTGAACAACCAGCACCAGGAAGGGAACAGCCAGCACCGGGTGGTGAACAACCAGCACCAGCACCAGGAGCAGGTGCGGGAGATGGAGCAC  700
 A  P  G  G  E  Q  P  A  P  G  R  E  Q  P  A  P  G  G  E  Q  P  A  P  A  P  G  A  G  A  G  D  G  A

GAGGAGGAAACGCAGGGGCAGGGAAAGGACAGGGACAAAACAATCAGGGTGCAAATGTCCCAAATGAAAAAGTTGTGAATGATTACCTACACAAAATTAG  800
R  G  G  N  A  G  A  G  K  G  Q  G  Q  N  N  Q  G  A  N  V  P  N  E  K  V  V  N  D  Y  L  H  K  I  R

ATCTAGCGTTACCACCGAGTGGACTCCATGCAGTGTAACCTGTGGAAATGGTGTAAGAATTAGAAGAAAAGGTCATGCAGGTAATAAAAAGGCAGAGGAC  900
  S  S  V  T  T  E  W  T  P  C  S  V  T  C  G  N  G  V  R  I  R  R  K  G  H  A  G  N  K  K  A  E  D

CTTACTATGGATGACCTTGAGGTGGAAGCTTGTGTAATGGATAAGTGCGCTGGCATATTTAACGTTGTGAGTAATTCATTAGGCTTAGTCATATTGTTAG  1000
 L  T  M  D  D  L  E  V  E  A  C  V  M  D  K  C  A  G  I  F  N  V  V  S  N  S  L  G  L  V  I  L  L

TCCTAGCATTATTCAATTAA
V  L  A  L  F  N  .
```

SG/EHI/WM33/Y09-39 *(P. knowlesi*; JQ219919)

```
ATGAAGAACTTCATTCTCTTGGCCGTCTCCTCCATCCTGCTGGTGGACTTGCTCCCCACACACTTCGAACATAATGTAGATCTCTCCAGGGCCATAAATG  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  L  P  T  H  F  E  H  N  V  D  L  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAATGTAGACACCAGTTCACTTGGCGCAGCACAGGTGAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAAGCCAAA  200
V  N  G  V  S  F  N  N  V  D  T  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  K  P  K

AGAAGGAGCTGATAAAGAAAAGAAAAAAGAAAAAGAAAAAGAAAAAGAAGAACCAAAGAAGCCAAATGAAAATAAGCTGAAACAACCGGAACAACCAGCA  300
  E  G  A  D  K  E  K  K  K  E  K  E  K  E  K  E  E  P  K  K  P  N  E  N  K  L  K  Q  P  E  Q  P  A

GCAGGAGCAGGGGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAAC  400
 A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E

AACCAGCGGCAGGAGCAGGGGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCAGCAGGAGCAGG  500
Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G

AGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCAGCAGGAGCAGGGGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCACCA  600
  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  P

GCACCAAGGAGGGAACAACCAGCAGCAGGAGCAGGGGGCGAACAACCAGCACCAGCACCAAGGAGGGAACAACCAGCACCAGGAGCAGGTGCGGGAGATG  700
 A  P  R  R  E  Q  P  A  A  G  A  G  G  E  Q  P  A  P  A  P  R  R  E  Q  P  A  P  G  A  G  A  G  D

GAGCACGAGGAGGAAACGCAGGGGCAGGTAAAGGACAGGGACAAAACAATCAGGGTGCGAATGTCCCAAATGAAAAAGTTGTGAATGATTACCTACACAA  800
G  A  R  G  G  N  A  G  A  G  K  G  Q  G  Q  N  N  Q  G  A  N  V  P  N  E  K  V  V  N  D  Y  L  H  K

AATTAGATCTAGCGTTACCACCGAGTGGACTCCATGCAGTGTAACCTGTGGAAATGGTGTAAGAATTAGAAGAAGACAGAATGCTGGTAATAAAAAGGCA  900
  I  R  S  S  V  T  T  E  W  T  P  C  S  V  T  C  G  N  G  V  R  I  R  R  R  Q  N  A  G  N  K  K  A

GAGGACCTTACTATGGATGACCTTGAGGTGGAAGCTTGTGTAATGGATAAGTGCGCTGGCATATTTAACGTTGTGAGTAATTCATTAGGCTTAGTCATAT  1000
 E  D  L  T  M  D  D  L  E  V  E  A  C  V  M  D  K  C  A  G  I  F  N  V  V  S  N  S  L  G  L  V  I

TGTTAGTCTTAGCATTATTCAATTAA
L  L  V  L  A  L  F  N  .
```

SG/EHI/WM33/Y09-47 *(P. cynomolgi*; JQ219927)

```
ATGAAGAACTTCATTCTCTTGGCCGTTTCTTCCATCCTGTTGGTGGACTTGTTCCGCACACACTGGGGACATAATGTAGATTTCTCCAAGGCCATAAATT  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  F  R  T  H  W  G  H  N  V  D  F  S  K  A  I  N

TAAATGGAGTAAGCTTCAGTAATGTAGACGCCAGTTCACTTGGCGCAGCACAGGTAAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAACCCAAA  200
L  N  G  V  S  F  S  N  V  D  A  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  N  P  K

AAACGAGGAAGGAGCTGATAAACCAAAAAAAAAGGACGAAAAACAAGTAGAACCAAAAAAGCCACGTGAAAATAAGCTGAAACAACCACAAGCAGGAGGT  300
  N  E  E  G  A  D  K  P  K  K  K  D  E  K  Q  V  E  P  K  K  P  R  E  N  K  L  K  Q  P  Q  A  G  G

GATGCAGGAAATGCACAAGCCGGAGGAGATGCGGGAAATGCACAAGCAGGAGGTGATGCAGGAAATGCACAAGCAGGAGGTGCACAAGCAGGAGGTGATG  400
 D  A  G  N  A  Q  A  G  G  D  A  G  N  A  Q  A  G  G  D  A  G  N  A  Q  A  G  G  A  Q  A  G  G  D

CAGGAAATGCACAAGCAGGAGGTGCACAAGCAGGAGGTGATGCAGGAAATGCACAAGCAGGAGGTGCGCAAGCAGGAGGTGATGCAGGAAATGCACAAGC  500
A  G  N  A  Q  A  G  G  A  Q  A  G  G  D  A  G  N  A  Q  A  G  G  A  Q  A  G  G  D  A  G  N  A  Q  A

GGGAGGTGCGCAAGCAGGAGGAGATGCAGGAAATGCACAAGCGGGAGGTGCGCAAGCAGGAGGAGATGCGGGAAATGCACAAGCAGGAGGTGCACAAGCA  600
  G  G  A  Q  A  G  G  D  A  G  N  A  Q  A  G  G  A  Q  A  G  G  D  A  G  N  A  Q  A  G  G  A  Q  A

GGAGGTGATGCAGGAAATGCACAAGCAGGAGGTGCACAAGCAGGAGGTGATGCAGGAAATGCACAAGCAGGAGGTGCGCAAGCAGGAGGTGCACAAGCAG  700
 G  G  D  A  G  N  A  Q  A  G  G  A  Q  A  G  G  D  A  G  N  A  Q  A  G  G  A  Q  A  G  G  A  Q  A

GAGGTGCACAAGCAGGAGGTGCACAAGCAGGAGGTGCACAAGCAGGAGGTGCGCAAGCAGGAGGTGCACAAGCAGGAGGTGCACAAGCAGGAGGTGCACA  800
G  G  A  Q  A  G  G  A  Q  A  G  G  A  Q  A  G  G  A  Q  A  G  G  A  Q  A  G  G  A  Q  A  G  G  A  Q

AGCAGGAGGAGCAAATGCGGGAAATAAAAAAAGCAGGAGACGCAGGAGCAGGACAGGGACAAAATAATGAAGGTGCGAATATGCCAAATGTAAAGCTTGTG  900
  A  G  G  A  N  A  G  N  K  K  A  G  D  A  G  A  G  Q  G  Q  N  N  E  G  A  N  M  P  N  V  K  L  V

AAAGAATACCTAGACAAAATTAGATCTACCATTGGCGTCGAGTGGAGTCCATGCAGTGTAACCTGTGGAAAGGGTGTAAGAATGAGAAGAAAAGTTAGTG  1000
 K  E  Y  L  D  K  I  R  S  T  I  G  V  E  W  S  P  C  S  V  T  C  G  K  G  V  R  M  R  R  K  V  S

CAGCTAACAAAAAACCAGAAGAGCTTGATGTGAATGACCTTGAGGCAGAAGTTTGTACAATGGATAAGTGCGCTGGTATATTTAACGTTGTGAGTAATTC  1100
A  A  N  K  K  P  E  E  L  D  V  N  D  L  E  A  E  V  C  T  M  D  K  C  A  G  I  F  N  V  V  S  N  S

ATTAGGGCTAGTCATATTGTTAGTCCTAGCATTATTCAATTAA
  L  G  L  V  I  L  L  V  L  A  L  F  N  .
```

SG/EHI/WM35/Y09-38 *(P. knowlesi*; JQ219920)

```
ATGAAGAACTTCATTCTCTTGGCCGTCTCCTCCATCCTGCTGGTGGACTTGCTCCCCACACACTTCGAACATAATGTAGATCTCTCCAGGGCCATAAATG  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  L  P  T  H  F  E  H  N  V  D  L  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAATGTAGACACCAGTTCACTTGGCGCAGCACAGGTGAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAAGCCAAA  200
V  N  G  V  S  F  N  N  V  D  T  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  K  P  K

AGAAGGAGCTGATAAAGAAAAGAAAAAAGAAAAAGAAAAAGAAAAAGAAGAACCAAAGAAGCCAAATGAAAATAAGCTGAAACAACCGGAACAACCAGCA  300
  E  G  A  D  K  E  K  K  K  E  K  E  K  E  K  E  E  P  K  K  P  N  E  N  K  L  K  Q  P  E  Q  P  A

GCAGGAGCAGGGGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAAC  400
 A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E

AACCAGCGGCAGGAGCAGGGGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCAGCAGGAGCAGG  500
Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G

AGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCAGCAGGAGCAGGGGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCACCA  600
  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  P

GCACCAAGGAGGGAACAACCAGCAGCAGGAGCAGGGGGCGAACAACCAGCACCAGCACCAAGGAGGGAACAACCAGCACCAGGAGCAGGTGCGGGAGATG  700
 A  P  R  R  E  Q  P  A  A  G  A  G  G  E  Q  P  A  P  A  P  R  R  E  Q  P  A  P  G  A  G  A  G  D

GAGCACGAGGAGGAAACGCAGGGGCAGGTAAAGGACAGGGACAAAACAATCAGGGTGCGAATGTCCCAAATGAAAAAGTTGTGAATGATTACCTACACAA  800
G  A  R  G  G  N  A  G  A  G  K  G  Q  G  Q  N  N  Q  G  A  N  V  P  N  E  K  V  V  N  D  Y  L  H  K

AATTAGATCTAGCGTTACCACCGAGTGGACTCCATGCAGTGTAACCTGTGGAAATGGTGTAAGAATTAGAAGAAGACAGAATGCTGGTAATAAAAAGGCA  900
  I  R  S  S  V  T  T  E  W  T  P  C  S  V  T  C  G  N  G  V  R  I  R  R  R  Q  N  A  G  N  K  K  A

GAGGACCTTACTATGGATGACCTTGAGGTGGAAGCTTGTGTAATGGATAAGTGCGCTGGCATATTTAACGTTGTGAGTAATTCATTAGGCTTAGTCATAT  1000
 E  D  L  T  M  D  D  L  E  V  E  A  C  V  M  D  K  C  A  G  I  F  N  V  V  S  N  S  L  G  L  V  I

TGTTAGTCTTAGCATTATTCAATTAA
L  L  V  L  A  L  F  N  .
```

SG/EHI/WM35/Y09-74 *(P. knowlesi*; JQ219921)

```
ATGAAGAACTTCATTCTCTTGGCCGTCTCCTCCATCCTGCTGGTGGACTTGCTCCCCACACACTTCGAACATAATGTAGATCTCTCCAGGGCCATAAATG  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  L  P  T  H  F  E  H  N  V  D  L  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAATGTAGACACCAGTTCACTTGGCGCAGCACAGGTGAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAAGCCAAA  200
V  N  G  V  S  F  N  N  V  D  T  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  K  P  K

AGAAGGAGCTGATAAAGAAAAGAAAAAAGAAAAAGAAAAAGAAAAAGAAGAACCAAAGAAGCCAAATGAAAATAAGCTGAAACAACCGGAACAACCAGCA  300
  E  G  A  D  K  E  K  K  K  E  K  E  K  E  K  E  E  P  K  K  P  N  E  N  K  L  K  Q  P  E  Q  P  A

GCAGGAGCAGGGGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAAC  400
 A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E

AACCAGCGGCAGGAGCAGGGGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCAGCAGGAGCAGG  500
Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G

AGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCAGCAGGAGCAGGGGGCGAACAACCAGCAGCAGGAGCAGGAGGCGAACAACCAGCACCA  600
  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  A  G  A  G  G  E  Q  P  A  P

GCACCAAGGAGGGAACAACCAGCAGCAGGAGCAGGGGGCGAACAACCAGCACCAGCACCAAGGAGGGAACAACCAGCACCAGGAGCAGGTGCGGGAGATG  700
 A  P  R  R  E  Q  P  A  A  G  A  G  G  E  Q  P  A  P  A  P  R  R  E  Q  P  A  P  G  A  G  A  G  D

GAGCACGAGGAGGAAACGCAGGGGCAGGTAAAGGACAGGGACAAAACAATCAGGGTGCGAATGTCCCAAATGAAAAAGTTGTGAATGATTACCTACACAA  800
G  A  R  G  G  N  A  G  A  G  K  G  Q  G  Q  N  N  Q  G  A  N  V  P  N  E  K  V  V  N  D  Y  L  H  K

AATTAGATCTAGCGTTACCACCGAGTGGACTCCATGCAGTGTAACCTGTGGAAATGGTGTAAGAATTAGAAGAAGACAGAATGCTGGTAATAAAAAGGCA  900
  I  R  S  S  V  T  T  E  W  T  P  C  S  V  T  C  G  N  G  V  R  I  R  R  R  Q  N  A  G  N  K  K  A

GAGGACCTTACTATGGATGACCTTGAGGTGGAAGCTTGTGTAATGGATAAGTGCGCTGGCATATTTAACGTTGTGAGTAATTCATTAGGCTTAGTCATAT  1000
 E  D  L  T  M  D  D  L  E  V  E  A  C  V  M  D  K  C  A  G  I  F  N  V  V  S  N  S  L  G  L  V  I

TGTTAGTCTTAGCATTATTCAATTAA
L  L  V  L  A  L  F  N  .
```

SG/EHI/WM42/Y10-1 *(P. cynomogi*; JQ219928)

```
ATGAAGAACTTCATTCTCTTAGCCGTTTCTTCCATCCTGTTGGTGGACTTGTTTCCCACAAACTGCGGGCATAATGTAGATTTCTCCAGGGCCATAAATT   100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  F  P  T  N  C  G  H  N  V  D  F  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAATGTAGACGCCAGTTCCCTTGGCGCAGCACAGGTAAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAACCCAAA   200
L  N  G  V  S  F  N  N  V  D  A  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  N  P  K

AAACGAGGAAGGAGCTGATAAACAAAAAAAGGACGAAAAAAAAGTAGAACCAAAAAAGCCACGTGAAAATAAGCTGAAACAACCAGCAGCAGCAAATGCG   300
  N  E  E  G  A  D  K  Q  K  K  D  E  K  K  V  E  P  K  K  P  R  E  N  K  L  K  Q  P  A  A  A  A  N  A

GGAGATGGACAACCAGCAGCAGCAAATGCAGGAGATGGACAACCAGCAGCAGCAAATGCGGGAGATGGACAACCAGCAGCAGCAAATGCGGGAGATGGAC   400
 G  D  G  Q  P  A  A  A  N  A  G  D  G  Q  P  A  A  A  N  A  G  D  G  Q  P  A  A  A  N  A  G  D  G

AACCAGCAGCAGCAAATGCGGGAGATGGACAACCAGCAGCAGCAAATGCGGGAGATGGACAACCAGCAGCAGCAAATGCAGGAGATGGACAACCAGCAGC   500
Q  P  A  A  A  N  A  G  D  G  Q  P  A  A  A  N  A  G  D  G  Q  P  A  A  A  N  A  G  D  G  Q  P  A  A

AGCAAATGCGGGAGATGGACAACCAGCAGCAGCAAATGCGGGAGATGGACAACCAGCAGCAGCAAATGCAGGAGATGGACAACCAGCAGCAGCAAATGCG   600
  A  N  A  G  D  G  Q  P  A  A  A  N  A  G  D  G  Q  P  A  A  A  N  A  G  D  G  Q  P  A  A  A  N  A

GGAGGTGCACAAGCAGGAGGTGCACAAGCAGGAGGTGCACAAGCAGGAGGAGCAAATGCGGGAAATAAAAAAGCAGGAGACGCAGGACAGGGACAAAATA   700
 G  G  A  Q  A  G  G  A  Q  A  G  G  A  Q  A  G  G  A  N  A  G  N  K  K  A  G  D  A  G  Q  G  Q  N

ATGGAGGTGCGAATGTGCCAAATGTAAAGCTTGTGCAAGAATACCTAGACAAAATTAGATCTACCATTGGCGTCGAGTGGAGTCCATGCAGTGTAACCTG   800
N  G  G  A  N  V  P  N  V  K  L  V  Q  E  Y  L  D  K  I  R  S  T  I  G  V  E  W  S  P  C  S  V  T  C

TGGAAAGGGTGTAAGAATGAGAAGAAAAGTTAATGCAGCTAACAAAAAACCAGAAGAGCTTGATGCGAATGACCTTGAGACTGAAGTTTGTACAATGGAT   900
  G  K  G  V  R  M  R  R  K  V  N  A  A  N  K  K  P  E  E  L  D  A  N  D  L  E  T  E  V  C  T  M  D

AAGTGCGCTGGTATATTTAACGTTGTGAGCAATTCATTAGGGCTAGTCATATTGTTAGTCCTAGCATTATTCAATTAA
K  C  A  G  I  F  N  V  V  S  N  S  L  G  L  V  I  L  L  V  L  A  L  F  N  .
```

SG/EHI/WM44/Y10-3 *(P. cynomogi*; JQ219929)

```
ATGAAGAACTTCATTCTCTTAGCCGTTTCTTCCATCCTGTTGGTGGACTTGTTTCCCACAAACTGCGGGCATAATGTAGATTTCTCCAGGGCCATAAATT   100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  F  P  T  N  C  G  H  N  V  D  F  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAATGTAGACGCCAGTTCACTTGGCGCAGCACAGGTAAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAACCCAAA   200
L  N  G  V  S  F  N  N  V  D  A  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  N  P  K

AAACGAGGAAGGAGCTGCTAAACAAAAAAAGGACGAAAAAAAAGTAGAACCAAAAAAGCCACGTGAAAATAAGCTGAAACAACCACCAGCAGCAGCAGAT   300
  N  E  E  G  A  A  K  Q  K  K  D  E  K  K  V  E  P  K  K  P  R  E  N  K  L  K  Q  P  P  A  A  A  D

GGAGCACCCGCAGCAGCAGATGGAGCACCCGCAGCAGCAGATGGAGCACCCGCAGCAGCAGATGGAGCACCAGCAGCAGCAGATGGAGCACCAGCAGCAG   400
 G  A  P  A  A  A  D  G  A  P  A  A  A  D  G  A  P  A  A  A  D  G  A  P  A  A  A  D  G  A  P  A  A

CAGATGGAGCACCAGCAGCAGCAGATGGAGCACCCGCAGCAGCAGATGGGGCACCAGCAGCAGCAGATGGAGCACCAGCAGCAGCAGATGGAGCACCAGC   500
A  D  G  A  P  A  A  A  D  G  A  P  A  A  A  D  G  A  P  A  A  A  D  G  A  P  A  A  A  D  G  A  P  A

AGCAGGAGGAAATCAGGCAGGTGCACAAGCAGGAGCAGGAGGAAATCAGGCAGGTGCACAAGCAGGAGCAGGAGGAAATCAGGCAGGTGCACAAGCAGGA   600
  A  G  G  N  Q  A  G  A  Q  A  G  A  G  G  N  Q  A  G  A  Q  A  G  A  G  G  N  Q  A  G  A  Q  A  G

GCAGGAGGAAATCAGGCAGGTGGACAGCCAGGAGCAGGAGGAAATCAGGCAGGTGGACAGCCAGGAGCAGGAGGAAATCAGGCAGGTGGACAGCCAGGAG   700
 A  G  G  N  Q  A  G  G  Q  P  G  A  G  G  N  Q  A  G  G  Q  P  G  A  G  G  N  Q  A  G  G  Q  P  G

CAGGAGGAAATCAGGCAGGTGGACAAGCAGGAGCAGGAGGAAATCAGGCAGGTGGACAAGCAGGAGGAGCAAATGCGGGAAATAAAAAAGCAGGAGACGC   800
A  G  G  N  Q  A  G  G  Q  A  G  A  G  G  N  Q  A  G  G  Q  A  G  G  A  N  A  G  N  K  K  A  G  D  A

AGGACAGGGACAAAATAATGGAGGTGCGAATGTGCCAAATGTAAAGCTTGTGCAAGAATACCTAGACAAAATTAGATCTACCATTAGCACCGAGTGGAGT   900
  G  Q  G  Q  N  N  G  G  A  N  V  P  N  V  K  L  V  Q  E  Y  L  D  K  I  R  S  T  I  S  T  E  W  S

CCATGCAGTGTAACCTGTGGAAAGGGTGTAAGAATGAGAAAAAAAGTTAATGCAGCTAACAAAAAACCAGAAGAGCTTGATGTGAATGACCTTGAGGCAG  1000
 P  C  S  V  T  C  G  K  G  V  R  M  R  K  K  V  N  A  A  N  K  K  P  E  E  L  D  V  N  D  L  E  A

AAGTTTGTACAATGGATAAGTGCGCTGGTATATTTAACGTTGTGAGTAATTCATTAGGGCTAGTCATATTGTTAGTCCTAGCATTATTCAATTAA
E  V  C  T  M  D  K  C  A  G  I  F  N  V  V  S  N  S  L  G  L  V  I  L  L  V  L  A  L  F  N  .
```

SG/EHI/WM44/Y10-30 *(P. cynomogi*; JQ219930)

```
ATGAAGAACTTCATTCTCTTAGCCGTTTCTTCCATCCTGTTGGTGGACTTGTTTCCCACAAACTGTGGGCATAATGTAGATTTCTCCAGGGCCATAAATT  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  F  P  T  N  C  G  H  N  V  D  F  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAATGTAGACGCCAGTTCACTTGGCGCAGCACAGGTAAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAACCCAAA  200
L  N  G  V  S  F  N  N  V  D  A  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  N  P  K

AAACGAGGAAGGAGCTGCTAAACAAAAAAAGGACGAAAAAAAAGTAGAACCAAAAAAGCCACGTGAAAATAAGCTGAAACAACCACCAGCAGCAGCAGAT  300
  N  E  E  G  A  A  K  Q  K  K  D  E  K  K  V  E  P  K  K  P  R  E  N  K  L  K  Q  P  P  A  A  A  D

GGAGCACCCGCAGCAGCAGATGGAGCACCCGCAGCAGCAGATGGAGCACCCGCAGCAGCAGATGGAGCACCAGCAGCAGCAGATGGAGCACCAGCAGCAG  400
 G  A  P  A  A  A  D  G  A  P  A  A  A  D  G  A  P  A  A  A  D  G  A  P  A  A  A  D  G  A  P  A  A

CAGATGGAGCACCAGCAGCAGCAGATGGAGCACCCGCAGCAGCAGATGGGGCACCAGCAGCAGCAGATGGAGCACCAGCAGCAGCAGATGGAGCACCAGC  500
A  D  G  A  P  A  A  A  D  G  A  P  A  A  A  D  G  A  P  A  A  A  D  G  A  P  A  A  A  D  G  A  P  A

AGCAGGAGGAAATCAGGCAGGTGCACAAGCAGGAGCAGGAGGAAATCAGGCAGGTGCACAAGCAGGAGCAGGAGGAAATCAGGCAGGTGCACAAGCAGGA  600
  A  G  G  N  Q  A  G  A  Q  A  G  A  G  G  N  Q  A  G  A  Q  A  G  A  G  G  N  Q  A  G  A  Q  A  G

GCAGGAGGAAATCAGGCAGGTGGACAGCCAGGAGCAGGAGGAAATCAGGCAGGTGGACAGCCAGGAGCAGGAGGAAATCAGGCAGGTGGACAGCCAGGAG  700
 A  G  G  N  Q  A  G  G  Q  P  G  A  G  G  N  Q  A  G  G  Q  P  G  A  G  G  N  Q  A  G  G  Q  P  G

CAGGAGGAAATCAGGCAGGTGGACAAGCAGGAGCAGGAGGAAATCAGGCAGGTGGACAAGCAGGAGGAGCAAATGCGGGAAATAAAAAAGCAGGAGACGC  800
A  G  G  N  Q  A  G  G  Q  A  G  A  G  G  N  Q  A  G  G  Q  A  G  G  A  N  A  G  N  K  K  A  G  D  A

AGGACAGGGACAAAATAATGGAGGTGCGAATGTGCCAAATGTAAAGCTTGTGCAAGAATACCTAGACAAAATTAGATCTACCATTAGCACCGAGTGGAGT  900
  G  Q  G  Q  N  N  G  G  A  N  V  P  N  V  K  L  V  Q  E  Y  L  D  K  I  R  S  T  I  S  T  E  W  S

CCATGCAGTGTAACCTGTGGAAAGGGTGTAAGAATGAGAAAAAAAGTTAATGCAGCTAACAAAAAACCAGAAGAGCTTGATGTGAATGACCTTGAGGCAG  1000
 P  C  S  V  T  C  G  K  G  V  R  M  R  K  K  V  N  A  A  N  K  K  P  E  E  L  D  V  N  D  L  E  A

AAGTTTGTACAATGGATAAGTGCGCTGGTATATTTAACGTTGTGAGTAATTCATTAGGGCTAGTCATATTGTTAGTCCTAGCATTATTCAATTAA
E  V  C  T  M  D  K  C  A  G  I  F  N  V  V  S  N  S  L  G  L  V  I  L  L  V  L  A  L  F  N  .
```

SG/EHI/WM44/Y10-64 *(P. fieldi*; JQ219935)

```
ATGAAGAACTTCATTCTCTTGGCCGTTTCTTCCATCCTGTTGGTGGACTTGTTCCCCACACACTGCGGGCATAATGTAGATCTCTCCAGGGCCATAAATT  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  F  P  T  H  C  G  H  N  V  D  L  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAATGTAGACGCCAGTTCACTTGGCGCAGCACAGGTAAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAACCCAAA  200
L  N  G  V  S  F  N  N  V  D  A  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  N  P  K

AGACGATGAAAAAGCTGATAAACCAAAAAAAAAGGACGAAAAAAAAGTAGAACCAAAAAAGCCACATGAAAATAAGCTGAAACAACCAGTCCCAGGAGCA  300
  D  D  E  K  A  D  K  P  K  K  K  D  E  K  K  V  E  P  K  K  P  H  E  N  K  L  K  Q  P  V  P  G  A

AATCAGGAAGGCGGAGCAGCAGCCCCAGGAGCAAATCAAGAAGGTGGAGCAGCAGCCCCAGGTGCAAACCAGGAAGGTGGAGCAGCAGCCCCAGGAGCAA  400
 N  Q  E  G  G  A  A  A  P  G  A  N  Q  E  G  G  A  A  A  P  G  A  N  Q  E  G  G  A  A  A  P  G  A

ACCAGGAAGGTGGAGCAGCAGCCCCAGGAGCAAACCAGGAAGGTGGAGCAGCAGCCCCAGGAGCAAATCAGGAAGGCGGAGCAGCAGCCCCAGGAGCAAA  500
N  Q  E  G  G  A  A  A  P  G  A  N  Q  E  G  G  A  A  A  P  G  A  N  Q  E  G  G  A  A  A  P  G  A  N

TCAGGAAGGTGGAGCAGCAGCCCCAGGTGCAAACCAGGAAGGTGGAGCAGCAGCCCCAGGAGCAAATCAGGAAGGCGGAGCAGCAGCCCCAGGAGCAAAT  600
  Q  E  G  G  A  A  A  P  G  A  N  Q  E  G  G  A  A  A  P  G  A  N  Q  E  G  G  A  A  A  P  G  A  N

CAGGAAGGCGGAGCAGCAGCCCCAGGAGCAAATCAGGAAGGTGGAGCAGCAGCCCCAGGAGCAAACCAGGGAGGTGGAGCAGCAGCACCAGGAGCAAACC  700
 Q  E  G  G  A  A  A  P  G  A  N  Q  E  G  G  A  A  A  P  G  A  N  Q  G  G  G  A  A  A  P  G  A  N

AGGGAGGTGGAGCAGCAGCACCAGGAGCAAACCAGGGAGGTGGAGCAGCAGCACCAGGAGCAAACCAGGGAGGTGGAGCAGCAGCACCAGGAGCAAACCA  800
Q  G  G  G  A  A  A  P  G  A  N  Q  G  G  G  A  A  A  P  G  A  N  Q  G  G  G  A  A  A  P  G  A  N  Q

GGAAGGTGGAGCAGCAGCACCAGGAGCAAACCAGGGAGGTGCAAAGCCAGCAGGAGGACAGGGACAAAATAATGAAGGTGCGAATAAGCCAGATGAAAAG  900
  E  G  G  A  A  A  P  G  A  N  Q  G  G  A  K  P  A  G  G  Q  G  Q  N  N  E  G  A  N  K  P  D  E  K

CATGTGAAAGAATACCTAGAGAAAATTAGATCTACCGTTGGCACCGAATGGACTCCATGCAGTGTAACCTGTGGAAAGGGTGTAAGAGTTAGAAGAAAAC  1000
 H  V  K  E  Y  L  E  K  I  R  S  T  V  G  T  E  W  T  P  C  S  V  T  C  G  K  G  V  R  V  R  R  K

TTAATGCAGGTGACAAAAAACCAGATAAGCTTACTCTGAATGACCTTGAGGCAGAAGTTTGTACAATGGATAAGTGCGCTGGCATATTTAACGTTGTGAG  1100
L  N  A  G  D  K  K  P  D  K  L  T  L  N  D  L  E  A  E  V  C  T  M  D  K  C  A  G  I  F  N  V  V  S

TAATTCATTAGGGCTAGTCATATTGTTAGTCCTAGCATTATTCAATTAA
  N  S  L  G  L  V  I  L  L  V  L  A  L  F  N  .
```

SG/EHI/WM91/Y11-61 *(P.inui*; JQ219936)

```
ATGAAGAACTTCATTCTCTTGGCCGTTTCGTCCATCCTGTTGGTGGACTTATTCCCCACACACTGCGGGCATAATGTAGATCTCTCCAGGGCCATAAATT  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  F  P  T  H  C  G  H  N  V  D  L  S  R  A  I  N

TAAATGGAGTAAGCTTCAATAATGTAGACGCCAGTTCACTTGGCGCCGCACAGGTAAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAAAAACCAAA  200
L  N  G  V  S  F  N  N  V  D  A  S  S  L  G  A  A  Q  V  R  Q  S  A  S  R  G  R  G  L  G  E  K  P  K

AGACCAGGAAGGAGATGCTAAGGGAAAAAAGAAGGGAAAAAAAGGAGAACCAAAGAATCCGGATGAAAAGAACCTGAAGCAACCAGCAGGGGACCCAGCA  300
  D  Q  E  G  D  A  K  G  K  K  K  G  K  K  G  E  P  K  N  P  D  E  K  N  L  K  Q  P  A  G  D  P  A

CCACCAGCAGGGGACCCAGCACCACCAGCAGGGGATCCAGCACCACCAGCAGGGGATCCAGCACCACCAGCAGGCGACCCAGCACCACCAGCAGGGGATC  400
 P  P  A  G  D  P  A  P  P  A  G  D  P  A  P  P  A  G  D  P  A  P  P  A  G  D  P  A  P  P  A  G  D

CAGCACCACCAGCAGGGGATCCAGCACCACCAGCAGGGGATCCAGCACCACCAGCAGGCGACCCAGCACCACCAGCAGGCGACCCAGCACCACCAGCAGG  500
P  A  P  P  A  G  D  P  A  P  P  A  G  D  P  A  P  P  A  G  D  P  A  P  P  A  G  D  P  A  P  P  A  G

GGATCCAGCACCACCAGCAGGGGATCCAGCACCACCAGCAGGGGGCCCAGCACCACCAGCAGGGGGCCCAGCACCACCAGCAGGGGGCCCAGCAGGACAG  600
  D  P  A  P  P  A  G  D  P  A  P  P  A  G  G  P  A  P  P  A  G  G  P  A  P  P  A  G  G  P  A  G  Q

GCACAAAACCCAGGAGGACCAGCACAAAACCCAGGAGGACCAGCACAAAACCCAGGAGGACCAGCACAAAAGCCAGGGGGACCAGCACAAAAGCCAGGAG  700
 A  Q  N  P  G  G  P  A  Q  N  P  G  G  P  A  Q  N  P  G  G  P  A  Q  K  P  G  G  P  A  Q  K  P  G

GACCAGCAAGAAACGCAGGAGAAGGAGCAGGAAATGCGGGAGGTGATAATGCCCCAGATGAAAAGGTTGTGAAAGATTACCTAGACAGAGTGAAATCTAA  800
G  P  A  R  N  A  G  E  G  A  G  N  A  G  G  D  N  A  P  D  E  K  V  V  K  D  Y  L  D  R  V  K  S  N

CCTTACCACCGAATGGAGTGTATGCAGTGTAAGCTGTGGACAGGGTGTAAGAGTTAGAAGAAAAGTTGGTGCATCTAACAAGAAACCAGAGGAACTTACT  900
  L  T  T  E  W  S  V  C  S  V  S  C  G  Q  G  V  R  V  R  R  K  V  G  A  S  N  K  K  P  E  E  L  T

CTGGATGACCTTGAGGTAGAAATTTGTAAAATGGAGAAGTGCGCTAGCATATTTAACGTTGTGAGTAATTCATTAGGGGTAGTAATATTGCTAGTCCTAG  1000
 L  D  D  L  E  V  E  I  C  K  M  E  K  C  A  S  I  F  N  V  V  S  N  S  L  G  V  V  I  L  L  V  L

CATTATTCAATTAA
A  L  F  N  .
```

SG/EHI/WM91/Y11-73 *(P.inui*; JQ219937)

```
ATGAAGAACTTCATTCTCTTGGCCGTTTCGTCCATCCTGTTGGTGGACTTATTCCCCACACACTGCGGGCATAATGTAGATCTCACCAGGGCCATAAATT  100
 M  K  N  F  I  L  L  A  V  S  S  I  L  L  V  D  L  F  P  T  H  C  G  H  N  V  D  L  T  R  A  I  N

TAAATGGAGTAAGCTTCAATAATGTAGACGCCAGTTCACTTGGCCCCGCACAGGTAAGACAAAGTGATAGACGAGGCAGAGGACTTGGTGAAAACCCAAA  200
L  N  G  V  S  F  N  N  V  D  A  S  S  L  G  P  A  Q  V  R  Q  S  D  R  R  G  R  G  L  G  E  N  P  K

AGACGAGGAAGGAGATGCTAAGGGAAAAAAGAAGGGAAAAAAAGGAGAACGAAAGAATCCAGATGAAAAGAACCTGAAGCAACCAGCAGGAGAGGCAGGA  300
  D  E  E  G  D  A  K  G  K  K  K  G  K  K  G  E  R  K  N  P  D  E  K  N  L  K  Q  P  A  G  E  A  G

GGAGCAGGACAGCCAGGAGCAGCAGGAGAGGCAGGAGGAGCAGGACAGCCAGGAGCAGCAGGAGAGGCAGGAGGAGCAGGACAGCCAGGAGCAGCAGGAG  400
 G  A  G  Q  P  G  A  A  G  E  A  G  G  A  G  Q  P  G  A  A  G  E  A  G  G  A  G  Q  P  G  A  A  G

AGGCAGGAGCAGCAGGACAGCCAGGAGCAGCAGGAGAGGCAGGAGGAGCAGGACAACCAGGAGCAGCAGGAGAGGCAGGAGGAGCAGGACAACCAGGAGC  500
E  A  G  A  A  G  Q  P  G  A  A  G  E  A  G  G  A  G  Q  P  G  A  A  G  E  A  G  G  A  G  Q  P  G  A

AGCAGGAGAGGCAGGAGGAGCAGGACAACCAGGAGCAGCAGGAGAGGCAGGAGGAGCAGGCCAACCAGGAGCAGCAGGAGAGGCAGGAGGAGCAGGACAA  600
  A  G  E  A  G  G  A  G  Q  P  G  A  A  G  E  A  G  G  A  G  Q  P  G  A  A  G  E  A  G  G  A  G  Q

CCAGGAGCAGCAGGAGAGGCAGGAGGAGCAGGACAGCCAGGAGCAGCAGGAGAGGCAGGAGGAGCAGGCCAACCAGGAGCAGCAGGAGAGGCAGGAGGAG  700
 P  G  A  A  G  E  A  G  G  A  G  Q  P  G  A  A  G  E  A  G  G  A  G  Q  P  G  A  A  G  E  A  G  G

CAGGACAACCAGGAGCAGCAGGACAACCAGGAGCAGCAGGACAGCCAGGAGCAGCAGGAGAGGCAGGAGGAGCAGGCAGGAGCAGCAGGACAACC  800
A  G  Q  P  G  A  A  G  Q  P  G  A  A  G  Q  P  G  A  A  G  Q  A  G  A  A  G  Q  A  G  A  A  G  Q  P

AGGAGCAGCAGGACAACCAGGAGCAGCAGGACAACCAGGAGGAGCAAAAGGAGCAAAAGGAGGAGACGCAGGACAACGAGGGCCAGCAGGACAAGGAGGA  900
  G  A  A  G  Q  P  G  A  A  G  Q  P  G  G  A  K  G  A  K  G  G  D  A  G  Q  R  G  P  A  G  Q  G  G

GCAGGAGGAAATTCGGGAGGTGAAAATATCCCAGATGCAAAGGTTGTGAAAGATTACCTAGACAAAGTGAAACCTACCCTTACCACCGAATGGAGTGCAT  1000
 A  G  G  N  S  G  G  E  N  I  P  D  A  K  V  V  K  D  Y  L  D  K  V  K  P  T  L  T  T  E  W  S  A

GCAGTGTAACCTGTGGGACGGGTGTAAGAGTTAGAAGAAAAGTTGGTGCATCAAACAAGAAACCAGAGGAACTTACTCCGGATGACCTTGAGGCAGAAAT  1100
C  S  V  T  C  G  T  G  V  R  V  R  R  K  V  G  A  S  N  K  K  P  E  E  L  T  P  D  D  L  E  A  E  I

TTGTAAAATGGATAAGTGCGCTAGCATATTTAACGTTGTGAGTAACTCATTAGGGGTAGTCATATTGTTGGTCATAGCATTATTCAATTAA
  C  K  M  D  K  C  A  S  I  F  N  V  V  S  N  S  L  G  V  V  I  L  L  V  I  A  L  F  N  .
```