PLOS ONE

Autophagy-Related Atg8 Localizes to the Apicoplast of the Human Malaria Parasite *Plasmodium falciparum*

Kei Kitamura[1,2], Chieko Kishi-Itakura[1], Takafumi Tsuboi[3], Shigeharu Sato[4], Kiyoshi Kita[5], Nobuo Ohta[2]*, Noboru Mizushima[1]*

1 Department of Physiology and Cell Biology, Tokyo Medical and Dental University, Tokyo, Japan, **2** Department of Environmental Parasitology, Tokyo Medical and Dental University, Tokyo, Japan, **3** Cell-Free Science and Technology Research Center and Venture Business Laboratory, Ehime University, Matsuyama, Ehime, Japan, **4** Division of Parasitology, MRC National Institute for Medical Research, London, United Kingdom, **5** Department of Biomedical Chemistry, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan

Abstract

Autophagy is a membrane-mediated degradation process, which is governed by sequential functions of Atg proteins. Although Atg proteins are highly conserved in eukaryotes, protozoa possess only a partial set of Atg proteins. Nonetheless, almost all protozoa have the complete factors belonging to the Atg8 conjugation system, namely, Atg3, Atg4, Atg7, and Atg8. Here, we report the biochemical properties and subcellular localization of the Atg8 protein of the human malaria parasite *Plasmodium falciparum* (PfAtg8). PfAtg8 is expressed during intra-erythrocytic development and associates with membranes likely as a lipid-conjugated form. Fluorescence microscopy and immunoelectron microscopy show that PfAtg8 localizes to the apicoplast, a four membrane-bound non-photosynthetic plastid. Autophagosome-like structures are not observed in the erythrocytic stages. These data suggest that, although *Plasmodium* parasites have lost most Atg proteins during evolution, they use the Atg8 conjugation system for the unique organelle, the apicoplast.

Citation: Kitamura K, Kishi-Itakura C, Tsuboi T, Sato S, Kita K, et al. (2012) Autophagy-Related Atg8 Localizes to the Apicoplast of the Human Malaria Parasite *Plasmodium falciparum*. PLoS ONE 7(8): e42977. doi:10.1371/journal.pone.0042977

Editor: Tobias Spielmann, Bernhard Nocht Institute for Tropical Medicine, Germany

Received January 30, 2012; **Accepted** July 16, 2012; **Published** August 10, 2012

Funding: This work was supported in part by Grant-in-aid from the Ministry of Education, Culture, Sports, Science and Technology (23117008) (to T.T.), the British Medical Research Council (U117584270) (to S.S.), Creative Scientific Research Grant 18GS0314 (to K. Kita), a grant for Ministry of Health, Labor and Welfare (H23-Kokui-004) (to N.O.) and the Funding Program for Next Generation World-Leading Researchers (to N.M.). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: matata.vip@tmd.ac.jp (NO); nmizu.phy2@tmd.ac.jp (NM)

Introduction

Macroautophagy (simply referred to as autophagy hereafter) is a fundamental cellular process, by which cytoplasmic components including proteins and organelles are delivered to the lysosome (or vacuole in yeasts and plants) for degradation. Autophagy is involved in many cellular functions such as adaptation to starvation, cell differentiation, quality control of proteins and organelles, aging, and degradation of invading microbes [1,2,3,4,5,6]. It is also implicated in human diseases such as cancer, inflammatory diseases, and neurodegeneration. Autophagy involves complex membrane dynamics; a membrane cisterna termed the isolation membrane (or phagophore) elongates on the endoplasmic reticulum (ER) and forms a double membrane-bound autophagosome, which contains cytoplasmic materials. Then, the autophagosome fuses with a lysosome to degrade the enclosed materials. Autophagosome formation is the central event of this process and is governed by autophagy-related (Atg) proteins, which were originally identified in yeast [7,8]. The genetic hierarchy of these Atg proteins has been determined and they are classified into at least six functional groups: the starvation-responsive Atg1 kinase complex (Atg1–Atg13–Atg17–Atg29–Atg31), the multi-membrane spanning protein Atg9, the class III phosphatidylinositol 3 (PtdIns 3)-kinase complex (Atg6–Atg14–Vps15–Vps34), the Atg2–Atg18 complex, the Atg12 – Atg5–Atg16 complex ("–" denotes a covalent attachment), and the Atg8–phosphatidylethanolamine (PE) conjugate (Figure 1A) [8,9,10].

These core Atg proteins are highly conserved in most eukaryotes including fungi, animals, and plants [11]. However, recent genome-wide analyses have revealed that they are only partially present in protozoa [12,13]. It is interesting that their conservation pattern is not random; the members belonging to the Atg8 conjugation systems are highly conserved in almost all protozoans, whereas potential homologs of other Atg proteins are only found sporadically (Figure 1A) [12]. The ubiquitin-like protein Atg8 can be covalently conjugated to PE through a sequential reaction that is mediated by a ubiquitin E1-like enzyme, Atg7, and an E2-like enzyme Atg3 [14]. Atg4 cleaves the C-terminal extension of the proform of Atg8 to expose a glycine residue, to which PE is conjugated. Atg4 also catalyzes deconjugation of the PE moiety from Atg8 – PE to release Atg8 from the membrane after completion of autophagosome formation [15]. Although the precise function of Atg8 and its PE conjugation in autophagy remains unclear, it is suggested that Atg8 – PE is important for membrane tethering and hemifusion [16], determination of the autophagosome size [17], and expansion and closure of the isolation membrane [18,19,20]. The partial conservation of the *ATG* genes in protozoans might imply that the smaller set of Atg proteins is sufficient to constitute the autophagosome in these

A

S.cerevisiae	*H.sapiens*	*P.falciparum*
Atg1	ULK1/2	-
Atg13	Atg13	-
Atg17	FIP200	-
Atg29	Atg101	
Atg31		
Atg9	Atg9A/B	-
Vps34	Vps34	Vps34 (PFE0765W)
Vps15	Vps15	-
Vps30/Atg6	Beclin 1	-
Atg14	Atg14(L)	-
-	AMBRA1	-
Atg2	Atg2A/B	-
Atg18	WIPI1/2/3/4	Atg18 (PF10_0126)
Atg12	Atg12	Atg12 (PF14_0667a)
Atg7	Atg7	Atg7 (PF11_0271)
Atg10	Atg10	-
Atg5	Atg5	Atg5? (PF14_0283)
Atg16	Atg16L1	-
Atg3	Atg3	Atg3 (PFI0280c)
Atg4	Atg4A/B/C/D	Atg4 (PF14_0171)
Atg7	Atg7	Atg7 (PF11_0271)
Atg8	LC3A/B/B2/C GABARAP/L1/L2/L3	Atg8 (PF10_0193)

B

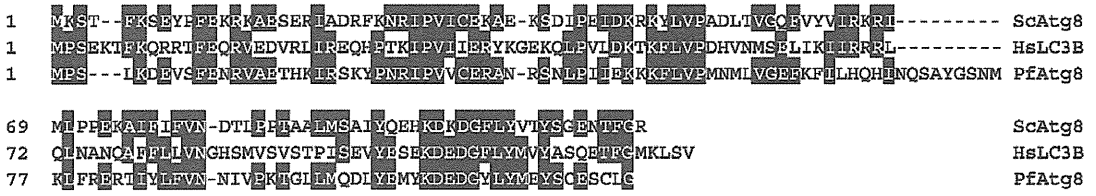

Figure 1. Atg protein sets are only partially conserved in *P. falciparum*. (A) List of Atg proteins in *S. cerevisiae*, *Homo sapiens* and *P. falciparum*. –, no ortholog found. It has been suggested that the mammalian FIP200–Atg101 complex and the yeast Atg17–29–31 complex are functional counterparts (dashed boxed) although they do not show significant sequence similarities. None of these factors seems to be conserved in *P. falciparum*. The tag of locus in the *P. falciparum* genome is indicated in parentheses. (B) Alignment of the full sequences of *S. cerevisiae* Atg8, *H. sapiens* LC3B (one of the Atg8 homologs), and *P. falciparum* Atg8. Identical amino acid residues are indicated with filled boxes.
doi:10.1371/journal.pone.0042977.g001

organisms. Alternatively, these organisms may use the Atg8 system for other purposes.

To date, several functional and morphological analyses of autophagy have been performed in protozoan parasites [13]. *Entamoeba invadens* possesses the Atg8 system, but lacks the Atg12 system. Atg8-positive vacuolar structures are generated in a PtdIns 3-kinase-dependent manner during encystation, but its ultrastructure is unknown [21]. In *Trypanosoma cruzi*, autophagosome-like double-membrane structures are formed in epimastigotes and implicated in differentiation into metacyclic trypomastigotes [22,23]. *Leishmania major* seems to have both Atg8 and Atg12 systems [24], and Atg8-positive punctate structures are observed during metacyclogenesis [25]. Accordingly, Atg4-deficient *L. major* shows a defect in differentiation into metacyclic promastigotes [25]. A more recent study performed in *Toxoplasma gondii* showed that genetic depletion of *TgAtg3*, which encodes an enzyme required for Atg8 – PE conjugation, causes growth inhibition and mitochondrial anomalies, which may be due to a defect in mitophagy [26].

In contrast, the nature of Atg proteins of the malaria parasite *Plasmodium* spp. remains largely unknown. *Plasmodium*, which belongs to phylum Apicomplexa together with *Toxoplasma*, possesses characteristic organelles such as the apicoplast, rhoptry, microneme, and dense granule. The *Plasmodium* sporozoite is transmitted by mosquito and first infects the hepatocyte which generates a large number of infectious merozoites. The merozoite infects erythrocytes and multiplies by schizogony to generate up to ~32 merozoites. Finally the infected erythrocytes rupture, and newly formed merozoites are released into the blood stream. An electron microscopy study of the rodent malaria parasite *P. berghei* demonstrated the presence of autophagosome-like double-membrane structures, which appeared to eliminate micronemes in liver-stage parasites [13,27]. Furthermore, *P. berghei* Atg8 appears to localize to abundant vesicles organized in a reticular network [27].

Because Atg8 – PE is present on both elongating isolation membranes and complete autophagosomes [28,29], Atg8 and its orthologs have been generally recognized as an autophagosome marker. Thus, in this study, we determined the biochemical properties and subcellular localization of Atg8 in *P. falciparum*, the major cause of human malaria. Contrary to our expectation, we found that *P. falciparum* Atg8 (PfAtg8) was specifically associated with the apicoplast, not autophagosomes, during the erythrocytic stage.

Results

Expression of PfAtg8 increases during the erythrocytic stage

Previously, it was reported that the *P. falciparum* genome has only a partial set of core Atg proteins [12,13]. We systematically searched for the orthologs of Atg proteins in the parasite genome and reached a similar conclusion (Figure 1A). We found genes encoding orthologs of a complete set of the Atg8 conjugation system (Atg3, Atg4, Atg7, and Atg8), although other *Atg* genes are only partially conserved (Figure 1A). Compared with Atg8 of the yeast *Saccharomyces cerevisiae*, PfAtg8 shows approximately 40% identity and 65% similarity and has the exposed C-terminal glycine residue, unlike Atg8 orthologs of other organisms (Figure 1B). *P. falciparum* possesses the class III PtdIns 3-kinase Vps34 [30]. In addition, although similarity is not high, there seem to be potential homologs of Atg5 (Figure S1), Atg12 (Figure S2), and Atg18 (Figure S3). Proteins encoded by *PF13_0116* and *PF14_0294* are partially similar to Atg2 and Vps15, respectively, but it remains unknown whether they are orthologs of Atg2 and Vps15.

To analyze endogenous PfAtg8, we generated two independent rabbit polyclonal antibodies against PfAtg8. Both antibodies specifically reacted with a band at approximately 14 kDa (Figure 2A). This size was close to that of Atg8 proteins of other species [28,29]. We used the anti-PfAtg8 antibody #1 in the following experiments unless otherwise specified. The expression level of PfAtg8 was low during early intra-erythrocytic development, but it increased as the parasite matured and reached the maximal level at the late schizont stage (Figure 2B). This pattern was similar to that of HSP70 serving as a cytosolic loading control, suggesting that the apparent increase of PfAtg8 corresponds to the increasing volume of the intra-erythrocytic parasites.

PfAtg8 is associated with membranes

In other organisms, Atg8 is present in two forms: free Atg8 and the membrane-associated form that conjugates with PE embedded in the lipid bilayer. The PE-conjugated and unconjugated forms of Atg8 can be separated by standard SDS-PAGE and urea-containing SDS-PAGE in mammals [29] and yeast [15], respectively. Although Atg8 gains molecular mass when conjugated to PE, its apparent mobility in SDS-PAGE increases probably because of the highly hydrophobic nature of PE. However, PfAtg8 was detected only as a single band in SDS-PAGE (Figure 2A), and no extra band was identified even in the presence of 6 M urea (data not shown). This suggests that the majority of PfAtg8 is present in either the conjugated or unconjugated form. When PfAtg8 was expressed in mammalian cells, PfAtg8 was also detected as a single band, although the mobility was lower than that of PfAtg8 expressed in parasites (Figure 2C). As it is unlikely that PfAtg8 conjugates with PE in mammalian cells, the band detected in mammalian cells is likely to represent the mobility of the unconjugated form. We therefore speculate that PfAtg8 is present primarily in a PE-conjugated form in *Plasmodium*.

We thus investigated whether or not PfAtg8 is membrane bound. The lysates of asynchronous parasites were fractionated by differential centrifugation. PfAtg8 was mainly collected in a low-speed (13,000×*g*) pellet (LSP) fraction (Figure 2D). PfAtg8 in the LSP fraction could be solubilized by treatment with 2% Triton X-100, but not with 2 M urea (Figure 2E). This behavior is characteristic of integral membrane proteins, and Atg8 – PE of other organisms have been known to behave in this manner [15,29]. Thus our data suggest that most PfAtg8 is membrane-associated in intra-erythrocytic *Plasmodium*.

In yeast and mammalian cells, Atg8/LC3 associating with the inner autophagosomal membrane is degraded upon fusion with lysosomes [31]. Chloroquine, a well-known anti-malarial drug, impairs lysosomal acidification, and thereby blocks degradation of LC3 in the lysosome [31]. If PfAtg8 is attached to the autophagosomal membrane in *P. falciparum* as are Atg8 or LC3 in yeasts and mammals, this protein should also be degraded within the lytic organelles such as the food vacuole. Chloroquine may affect this process and therefore we examined its effect on the level of PfAtg8. However, treatment of the parasite growing in erythrocytes with chloroquine did not increase the amount of PfAtg8 (Figure 2F), suggesting that PfAtg8 is not involved in the autophagic process at this stage of the parasite's life cycle.

PfAtg8 localizes to the apicoplast membrane

We next determined the subcellular localization of PfAtg8 by immunofluorescence microscopy. In segmented schizonts, the PfAtg8 signal was detected as a single punctate structure in each merozoite (Figure 3). The size of each punctate structure was approximately 200–400 nm. The fact that every parasite possessed one such structure precluded the possibility that the PfAtg8-positive structure is rapidly turned over in the same way as the autophagosome.

The nature of these PfAtg8-positive structures was further characterized by double staining with organelle markers. PfAtg8 did not colocalize with any markers for the merozoite apical organelles such as the microneme (Figure 3A), rhoptry body (Figure 3B), rhoptry neck (Figure 3C), and dense granule (Figure 3D). By contrast, PfAtg8 colocalized with the apicoplast-localizing green fluorescent protein (ACP-GFP) (some of the parasites displayed only weak ACP-GFP expression) [32,33] (Figure 3 F). We confirmed the colocalization between ACP-GFP and PfAtg8 using the independent anti-PfAtg8 antibody #2 (Figure 3G). All the organelles labeled with the ACP-GFP antibody also reacted with the antibody against the plastid-localizing PfHU (histone-like protein, heat unstable), an endogenous apicoplast marker (Figure 3H) [34]. The PfAtg8-positive structure was observed in close proximity to the mitochondrion, and even appeared to overlap part of the organelle (Figure 3E). This partial overlapping between mitochondria and PfAtg8 is consistent with the fact that the apicoplast and the mitochondrion are juxtaposed, probably maintaining physical contact, in the parasite cell [35,36]. To better dissect the localization of PfAtg8, we looked at the parasites at an earlier stage. Morphology of the apicoplast dramatically changes during development inside erythrocytes [32,37]. At the late trophozoite to early schizont stages, the apicoplast forms a tubular or branched shape. In fact, PfAtg8 localized to the tubular or branched apicoplasts, which were clearly distinct from the mitochondria (Figure 4). Taken together, these data suggest that PfAtg8 localizes to the apicoplast during normal development.

Even though the main localization of PfAtg8 is the apicoplast, PfAtg8 could be present on other structures such as autophagosomes. However, the localization pattern of PfAtg8 was not

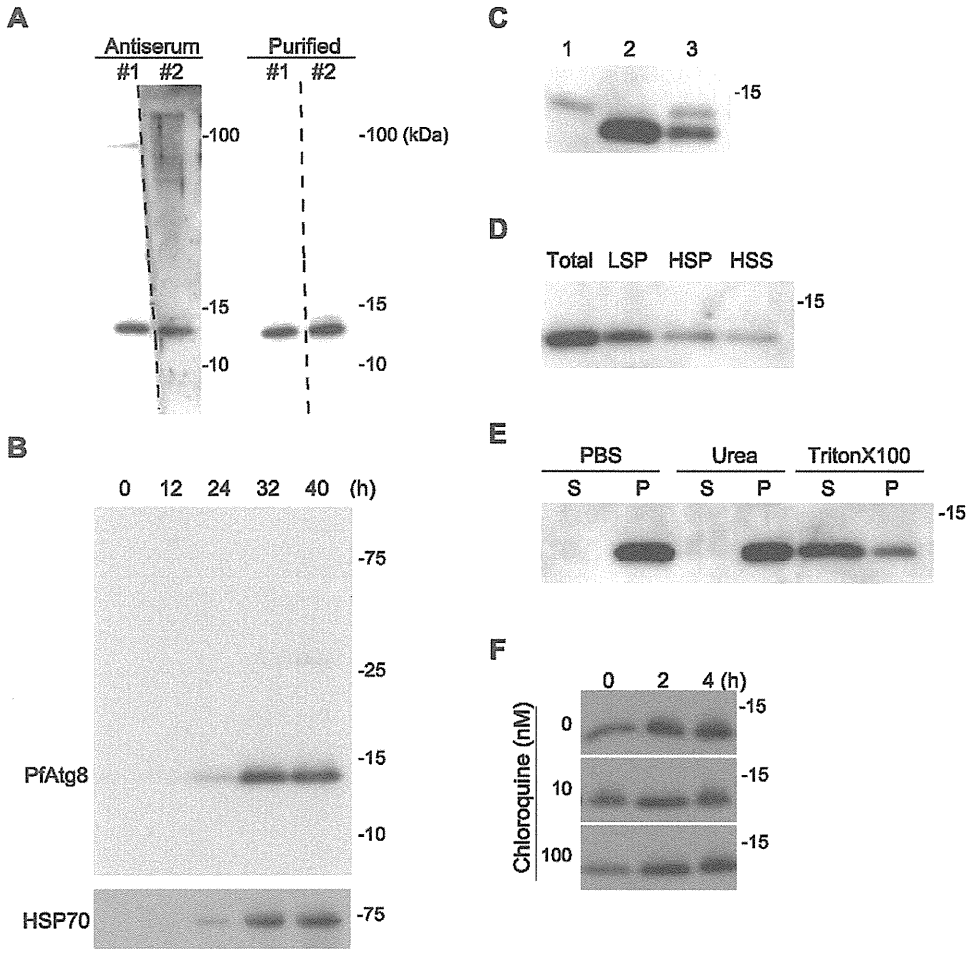

Figure 2. PfAtg8 is associated with membranes. (A) Specificity of the two independently generated anti-PfAtg8 antibodies (#1 and #2). Crude antisera and purified antibodies were used for immunoblotting of lysates of asynchronized *P. falciparum* parasites. (B) Expression of PfAtg8 increases during the erythrocytic stage of development. Highly synchronized *P. falciparum* parasites were collected at 0, 12, 24, 32, and 40 h after invasion. The duration of one cycle of the erythrocyte stage was approximately 42 h. Expression levels of PfAtg8 were analyzed by immunoblotting. An antibody against HSP70 was used as a loading control. (C) PfAtg8 exogenously expressed in mammalian cells (lane 1), endogenous PfAtg8 expressed in *P. falciparum* (lane 2), and the mixture of these two samples were subjected to immunoblot analysis using anti-PfAtg8 antibody. (D) Lysates of asynchronized *Plasmodium* were separated into low-speed (13,000×*g*) pellet (LSP), high-speed (100,000×*g*) pellet (HSP), and high-speed supernatant (HSS) fractions, and analyzed by immunoblotting using anti-PfAtg8 antibody. (E) The LSP fraction prepared in (D) was treated with 2 M urea or 2% Triton-X 100 and separated into 100,000×*g* pellet (P) and supernatant (S). (F) Infected erythrocytes were cultured in the presence of the indicated concentration of chloroquine and expression of PfAtg8 was analyzed.
doi:10.1371/journal.pone.0042977.g002

significantly changed by treatment of chloroquine (Figure 5A), which can typically accumulate autophagosomes/autolysosomes in mammalian cells (as we mentioned above, ACP-GFP was not uniformly expressed and some merozoites displayed only faint GFP signals) [38]. Thus, we could not conclude whether *P. falciparum* can generate PfAtg8-positive autophagosomes. Furthermore, treatment with wortmannin, a PtdIns 3-kinase inhibitor, did not affect PfAtg8 localization, suggesting that association of PfAtg8 with the apicoplast membrane is independent of PtdIns 3-phosphate (Figure 5B).

We further analyzed the localization of PfAtg8 by immunoelectron microscopy using the anti-PfAtg8 antibody. The silver-enhanced gold particles specifically associated with multiple membrane-bound organelles (Figure 6A). The inside of the organelles were filled with relatively low-density materials and a fiber-like structure, which are features of the apicoplast [35,39,40]. These characteristics suggest that the organelles surrounded by PfAtg8 were different from autophagosomes, which are defined as double membrane-bound organelles containing undigested cytoplasmic materials, and from autolysosomes, which contain degraded materials [41]. We did not observe such autophagic structures in the parasites in the erythrocytic stage. To confirm that the PfAtg8-positive multi-membrane structures were indeed apicoplasts, we performed immunoelectron microscopy of ACP-GFP-expressing parasites. Anti-GFP antibody specifically reacted with multi-membrane organelles that looked the same as the structures to which PfAtg8 localized (Figure 6A and 6B). The PfAtg8 signals were not detected on the mitochondrion that was in close proximity to the apicoplast (Figure 6A). This suggests that the fluorescence signal detected apparently in the mitochondrion (Figure 3E) was caused by the spatial overlap of the organelle and the apicoplast. Taken together, these data suggest that PfAtg8 specifically localizes to the membrane of the apicoplast in *P. falciparum*.

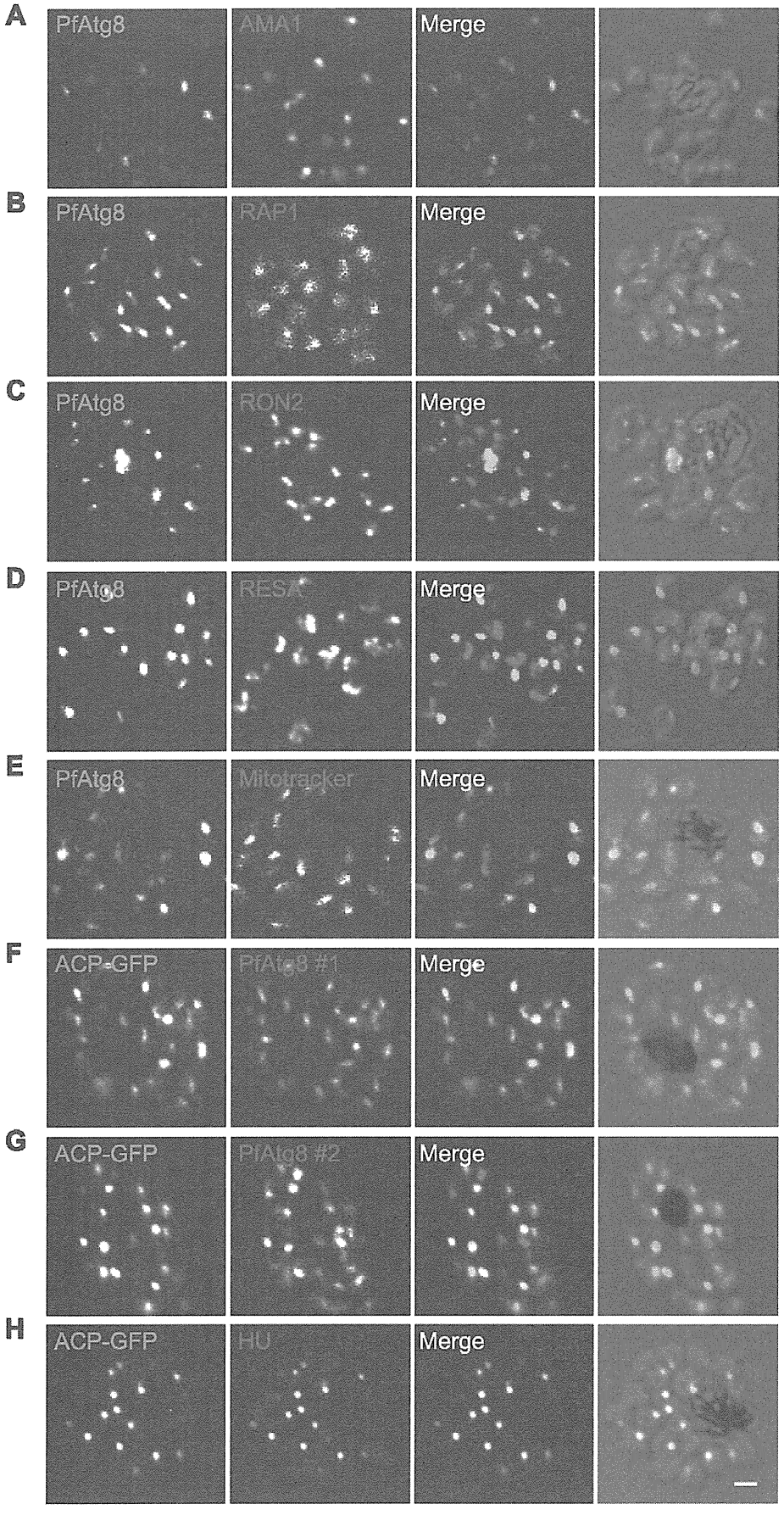
A
PfAtg8
AMA1
Merge
B
PfAtg8
RAP1
Merge
C
PfAtg8
RON2
Merge
D
PfAtg8
RESA
Merge
E
PfAtg8
Mitotracker
Merge
F
ACP-GFP
PfAtg8 #1
Merge
G
ACP-GFP
PfAtg8 #2
Merge
H
ACP-GFP
HU
Merge

Figure 3. PfAtg8 localizes to the apicoplast. *P. falciparum* FCR3 (A–E) and *P. falciparum* 3D7 transfected with ACP-GFP (F–H) were stained with the indicated organelle markers and visualized by confocal microscopy (because ACP-GFP was not uniformly expressed, some merozoites displayed only faint GFP signals). Anti-PfAtg8 antibody #1 was used in (A–F), and anti-PfAtg8 antibody #2 was used in (G). Apical membrane antigen 1 (AMA1) as a microneme marker (A), rhoptry-associated protein 1 (RAP1) as a rhoptry body marker (B), rhoptry neck protein 2 (RON2) as a rhoptry neck marker (C), the ring-infected erythrocyte surface antigen (RESA) as a dense granule marker (D), MitoTrackerRed CMXRos as a mitochondria marker (E), ACP-GFP (F–H) and the organellar histone-like protein PfHU (H) as an apicoplast marker were used. Scale bar, 1 μm.
doi:10.1371/journal.pone.0042977.g003

Discussion

We report here that PfAtg8 is associated with the apicoplast, probably with the outermost membrane likely in a lipid-conjugated form during the erythrocytic stage. Although we did not detect any autophagosome-like structures in *P. falciparum* even under chloroquine treatment, we do not rule out the possibility that Atg8 can localize to autophagosomes if such structures are formed in *Plasmodium*. The apicoplast is a non-photosynthetic plastid, which is an essential organelle possessing its own genome [40,42,43]. The apicoplast is involved in several metabolic pathways such as biosynthesis of fatty acids, haem, isoprenoid (required for tRNA modification, etc) and iron–sulphur clusters. *P. falciparum* has one plastid, and it physically contacts with the mitochondrion during intraerythrocyte stage [35,36], although the relationship between these two organelles seem to be more complicated in liver stages [44]. The apicoplast has four membranes and that is believed to explain the origin of the organelle from a secondary endosymbiotic alga, most likely a red alga [43,45,46]. The outer two membranes appear to be related to the ER. Nuclear-encoded apicoplast proteins possess an N-terminal signal peptide that is essential for their delivery to the apicoplast via Golgi-independent transport [32,40,47,48,49]. ER-associated protein degradation-like machinery exists in the second outermost membrane [50,51]. Although the autophagosome does not have any ER-related proteins and contribution of other organelles such as mitochondria [52] and the plasma membrane [53], autophagosomes are basically generated on or in close proximity to the ER [8,54]; even direct membrane continuity between the ER and autophagosome has been suggested [55,56]. Therefore, Atg8 may have a shared role in biogenesis of ER-related organelles.

To date, no particular contribution of PfAtg8 on the apicoplast or its biogenesis has been predicted. This protein was suggested to be essential because deletion of the *Atg8* gene in *P. berghei* causes a lethal phenotype [13,27]. Our ultrastructural analysis, as well as the fact that the parasite has one apicoplast throughout its cell cycle (except for a short period of organellar division), suggest that PfAtg8 is unlikely to be used for autophagic degradation of the apicoplast. Conditional targeting of *Atg3* in *T. gondii* resulted in a severe growth defect with altered mitochondrial morphology [26]. The observed mitochondrial defects might have been caused by a defect in mitochondrial autophagy (mitophagy) as suggested by Besteiro et al., but it may be due to an impaired apicoplast–mitochondria relationship. If the Atg8 conjugation system is involved in essential cellular activities of the malaria parasites, compounds that can inhibit the conjugation reaction (i.e. inhibitors of PfAtg7 or PfAtg3) would be promising therapeutic tools.

P. falciparum contains the class III PtdIns 3-kinase Vps34 (Figure 1A) and PtdIns 3-phosphate is present on both the food vacuole and the apiocoplast membrane [30]. As the Atg8 system functions downstream of the PtdIns 3-kinase complex in starvation-induced autophagy in both yeast and mammals [9,10], we speculated that association between PfAtg8 and the apicoplast membrane could depend on PtdIns 3-kinase activity. However, we observed that treatment of *P. falciparum* with wortmannin did not affect the localization of PfAtg8 (Figure 5B). Recently, it was reported that LC3 (a mammalian Atg8) can associate with membranes even in the absence of upstream Atg factors such as the ULK1/Atg1 complex, Atg9, and PtdIns 3-kinase activity in some types of selective autophagy such as xenophagy against *Salmonella* [57] and Parkin-mediated mitophagy [58]. Nonetheless, the membrane association with LC3 still depends on the two ubiquitin-like LC3 and Atg12 conjugation systems. Therefore, it is possible that the PfAtg8 conjugation system has a unique function, which is independent of most other Atg proteins.

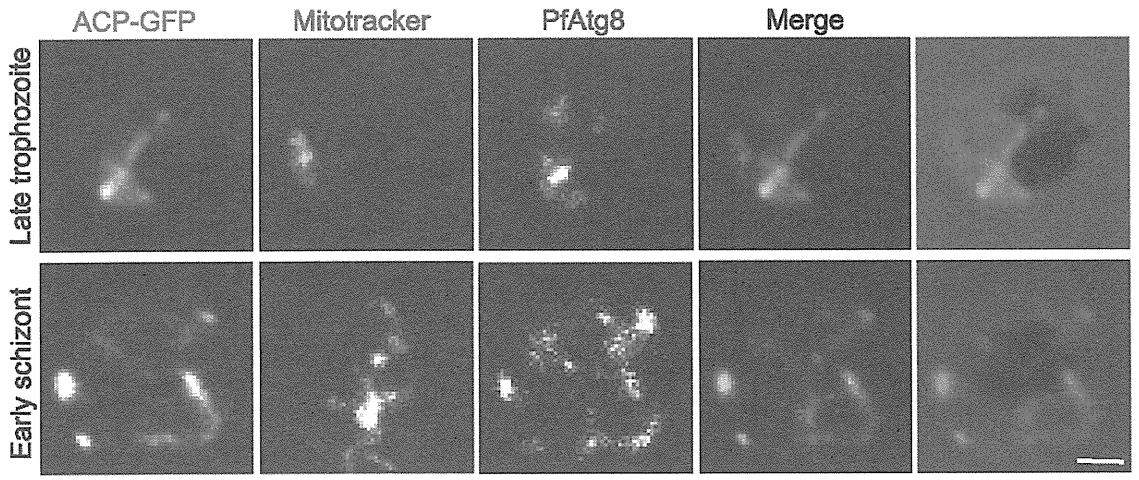

Figure 4. PfAtg8 localizes to tubular and branched apicoplasts. *P. falciparum* transfectant expressing ACP-GFP at late trophozoite and early schizont stages was stained with anti-GFP and anti-PfAtg8 antibodies and MitoTrackerRed CMXRos, and visualized by confocal microscopy. Scale bar, 1 μm.
doi:10.1371/journal.pone.0042977.g004

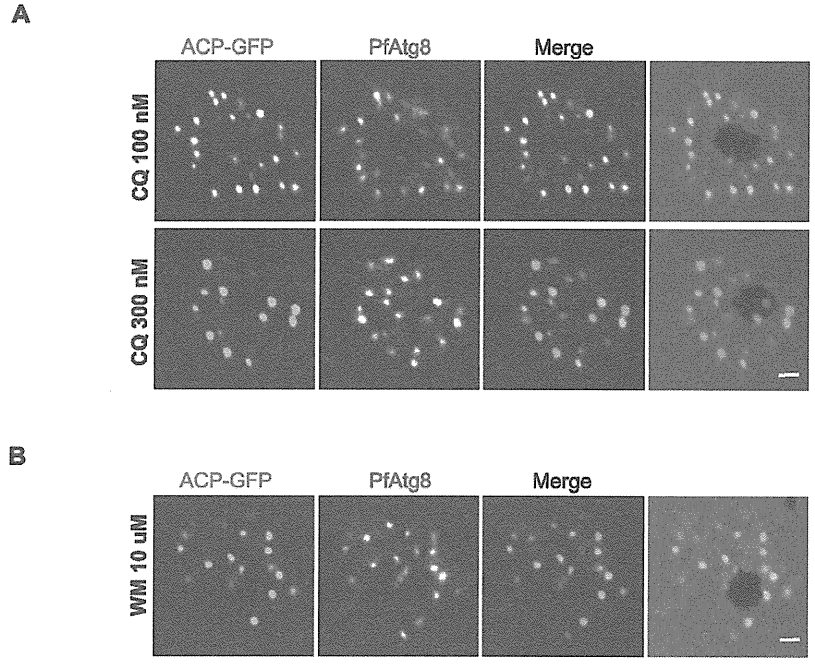

Figure 5. PfAtg8 localization is not affected by chloroquine or wortmannin treatment. *P. falciparum* transfectant expressing ACP-GFP was treated with chloroquine (100 or 300 nM) (A), or wortmannin (10 μM) (B) for 2 h. Scale bar, 1 μm.
doi:10.1371/journal.pone.0042977.g005

Another interesting issue is the requirement of the Atg12 conjugation system. In the autophagy pathway, the Atg8/LC3 conjugation reaction requires an E3-like activity of the Atg12–Atg5 conjugate both in yeast and mammals [59,60]. However, PfAtg12 lacks the C-terminal glycine residue (Figure S2), which is essential for formation of an isopeptide bond with Atg5. An attractive hypothesis is that PfAtg12 alone may have E3-like activity without Atg5, although the *P. falciparum* genome contains a

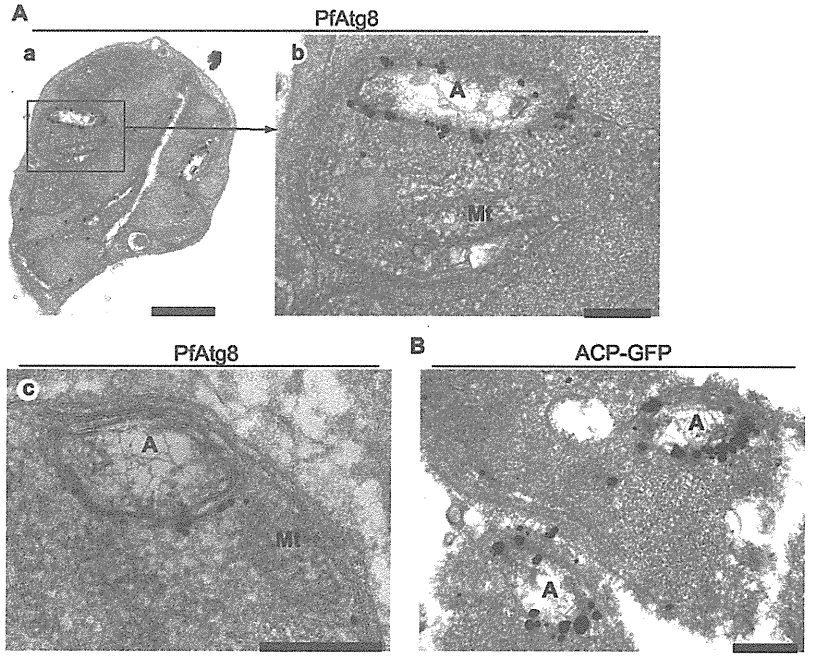

Figure 6. PfAtg8 is associated with the apicoplast membrane. (A) *P. falciparum* FCR3 parasites at the schizont stage were analyzed by immunoelectron microscopy (immunogold and silver enhancement method) with an antibody against PfAtg8 (#1). (a) A schizont in an erythrocyte. (b) A magnified image of the area indicated in (a). (c) Another typical image of a PfAtg8-positive structure. (B) *P. falciparum* transfectant expressing ACP-GFP was analyzed as in panel (A) with an antibody against GFP. A, apicoplast; Mt, mitochondrion. Scale bars, (A, a) 1 μm, (A, b and c, and B) 200 nm.
doi:10.1371/journal.pone.0042977.g006

gene encoding a potential Atg5 homolog (Figure S1). The gene, *PF14_0283* encodes an 863-amino acid (aa) protein that is much larger than yeast (294 aa) and human Atg5 (275 aa). Because of the presence of a number of insertion sequences, it is important to examine whether this Atg5 candidate is a functional Atg5 ortholog that should conjugate and collaborate with PfAtg12 in the organism. Further understanding the roles of Atg8 and Atg12 in *Plasmodium* will provide a general insight into the functions of Atg proteins even in the autophagy pathway.

Materials and Methods

Parasite culture

P. falciparum strain FCR3 was cultured in human B^+ erythrocytes as described [61]. In some subcellular localization experiments, the 3D7 parasite strain transfected with pSSPF2/GFP-ACP was used; the transfectant was cultured in the standard culture medium supplemented with 5 nM WR99210 [33]. Where indicated, chloroquine (Sigma-Aldrich) was added to the culture medium. For synchronizing the culture, the red blood cells infected by the late stage schizont were recovered from asynchronous culture by 60% Percoll (GE healthcare) density centrifugation at 2000×*g* for 20 min. After 4 h the culture was treated with 5% D-sorbitol [62], yielding parasites tightly synchronized in the early ring stage (0–4 h after parasite invasion of the erythrocyte).

Cloning of PfAtg8 cDNA and generation of anti-PfAtg8 antibodies

RNA extraction and cDNA synthesis were carried out as described previously [63,64]. GST-fused PfAtg8 recombinant protein was generated using a wheat germ cell-free system [65]. Two independent anti-PfAtg8 antisera (#1 and #2) were raised in two New Zealand white rabbits and the antibodies were purified using GST-PfAtg8 recombinant protein. Animal experimental protocols were approved by the Institutional Animal Care and Use Committee of Tokyo Medical and Dental University (No. 0110115A).

Immunoblotting

Parasites were collected from erythrocytes by treatment with 0.15% saponin (Sigma) in phosphate-buffered saline (PBS) with Complete Protease Inhibitor cocktail (Roche Applied Science), washed three times in PBS and lysed in sample buffer. Parasite extracts were loaded onto 13.5% SDS gel and transferred to a PVDF membrane. Blots were blocked with 5% skim milk in Tris-buffered saline with 0.01% Tween 20 (TBST) and stained with primary antibodies overnight at 4°C. The following primary antibodies were used: rabbit anti-PfAtg8 and mouse monoclonal anti-PfHSP70 antibodies (1:100) [66,67]. After washing with TBST, blots were stained with HRP-conjugated secondary antibodies and visualized with SuperSignal West Pico Chemiluminescent substrate (Thermo Fisher Scientific).

Subcellular fractionation

Asynchronous parasites were harvested as described above. Parasite pellets were disrupted by three cycles of freezing/thawing in MSE buffer (225 mM mannitol, 75 mM sucrose, 0.1 mM EDTA, and 3 mM Tris-HCl [pH 7.4]). Cell debris and intact erythrocytes were removed by centrifugation at 800×*g* for 5 min. The supernatant was spun at 13,000×*g* for 15 min to separate the LSP, and the supernatant was centrifuged again at 100,000×*g* for 60 min to generate the high-speed pellet (HSP) and high-speed supernatant (HSS). The LSP and HSP were resuspended in the same buffer. To analyze solubility, each sample was incubated with 2 M urea or 2% Triton X-100 on ice for 1 h, and then centrifuged at 100,000×*g* for 1 h. The samples were precipitated with ice-cold acetone, resuspended in SDS-PAGE sample buffer, and analyzed by SDS-PAGE.

Expression of PfAtg8 in mammalian cells

PfAtg8 cDNA was inserted into a pCI-neo mammalian expression plasmid (Promega) and transfected into HEK293T cells [68] using Lipofectamine 2000 reagent (Invitrogen). Total cell lysates were subjected to SDS-PAGE and immunoblot analysis.

Immunofluorescence microscopy

Parasite thin blood smears were fixed with 4% paraformaldehyde/PBS for 10 min and samples were permealized with 0.1% Triton X-100/PBS for 15 min. After blocking with 3% bovine serum albumin/PBS for 1 h, samples were incubated with primary and secondary antibodies for 2 h and 1 h, respectively. The smears were mounted with Prolong Gold (Invitrogen). All reactions were carried out at room temperature. Samples were observed with a confocal laser microscope (FV1000D IX81, Olympus) using a 60x PlanApoN oil immersion lens (1.42 NA; Olympus). The following primary antibodies were used: purified rabbit anti-PfAtg8 (1:200 for #1, and 1:100 for #2) antibody, mouse anti-apical membrane antigen (AMA)1 (1:500) [67], anti-rhoptry-associated protein 1 (RAP1, 1:200) [67], anti-rhoptry neck protein 2 (RON2, 1:200) [69], and anti-ring-infected erythrocyte surface antigen (RESA) (23/9, 1:200) [70] antibodies, rabbit anti-PfHU (organellar histone-like protein) antibody [34], and rat anti-GFP antibody (Nacalai Tesque). For visualizing the mitochondrion, parasites were preincubated for 30 min with complete culture medium containing 100 nM MitoTracker Red CMXRos (Molecular Probes).

Immunoelectron microscopy

Mature schizont stage parasites were enriched from synchronous culture using MACS 25LD columns (MiltenyiBiotec) as previously described [71]. For immunoelectron microscopy of *P. falciparum*, the previously described pre-embedding silver enhancement immunogold method [72] was used with slight modifications. The parasitized erythrocytes were fixed in 4% paraformaldehyde and 0.0075% glutaraldehyde dissolved in 0.1 M sodium phosphate buffer (PB) (pH 7.4) for 2 h and then washed three times with PB. Then the cells were permeabilized in liquid nitrogen and incubated in a blocking buffer containing 0.005% saponin, 10% goat serum, 0.1% cold water fish gelatin, and 10% bovine serum albumin for 30 min, and reacted with rabbit anti-PfAtg8 (#1) or rat monoclonal anti-GFP (IgG2a, Nacalai Tesque #04404-84) in blocking buffer at 4°C overnight. Next the cells were washed in PB containing 0.005% saponin and incubated with goat anti-rabbit IgG or anti-rat IgG conjugated with colloidal gold (1.4-nm diameter, Nanogold, Nanoprobes) in blocking buffer for 2 h at room temperature. Cells were washed five times with PB containing 0.005% saponin for 10 min, washed with PB for 5 min, and fixed with 1% glutaraldehyde for 10 min. After washing, the gold partiles were intensified using a silver enhancement kit (HQ silver, Nanoprobes) for 6 min at 20°C in the dark. After washing in distilled water, the cells were post-fixed with 0.03% OsO_4 for 15 min at 4°C. After washing with PB, cells were resuspended in 2% gelatin (Sigma) and pelleted again. Microcentrifuge tubes were plunged into ice-cold water to quickly solidify the gelatin with the cells. The tip of the tube was cut open and the cell pellets were retrieved into 15% ethanol, and cut into 1-mm^3 blocks. The blocks were suspended and dehydrated with a graded series of ethanol concentrations, and embedded in epoxy

resin. Ultrathin sections were doubly stained with uranyl acetate and lead citrate and observed using a Hitachi H7100 electron microscope.

Supporting Information

Figure S1 Sequence alignment of Atg5 homologs. Alignment of the sequences of *S. cerevisiae* Atg5, *H. sapiens* Atg5 and *P. falciparum* Atg5. Asterisk (*) shows the position of the Lys residue that receives Atg12 conjugation in yeast and human. This Lys is conserved in PfAtg5.
(TIF)

Figure S2 Sequence alignment of Atg12 homologs. Alignment of the sequences of *S. cerevisiae* Atg12, *H. sapiens* Atg12 and *P. falciparum* Atg12. Asterisk (*) shows the C-terminal Gly residue essential for conjugation with Atg5 in yeast and human. PfAtg12 lacks this Gly residue.
(TIF)

Figure S3 Sequence alignment of Atg18 homologs. Alignment of the sequences of *S. cerevisiae* Atg18, *H. sapiens* WIPI1 and *P. falciparum* Atg18. Asterisk (*) shows the motif required for PtdIns 3-phosphate binding.
(TIF)

Acknowledgments

We thank Ms. Yuriko Sakamaki for help with electron microscopy, and Dr. Nobuo N. Noda for discussion. We are also grateful to Dr. Narie Sasaki for providing the anti-PfHU antibody, and Dr. Robin F. Anders for anti-RESA antibody.

Author Contributions

Conceived and designed the experiments: K. Kitamura K. Kita NO NM. Performed the experiments: K. Kitamura CK-I TT. Analyzed the data: K. Kitamura CK-I TT SS K. Kita NO NM. Contributed reagents/materials/analysis tools: K. Kitamura TT SS K. Kita NO. Wrote the paper: K. Kitamura CK-I TT SS K. Kita NO NM.

References

1. Cecconi F, Levine B (2008) The role of autophagy in mammalian development: cell makeover rather than cell death. Dev Cell 15: 344–357.
2. Levine B, Kroemer G (2008) Autophagy in the pathogenesis of disease. Cell 132: 27–42.
3. Levine B, Mizushima N, Virgin HW (2011) Autophagy in immunity and inflammation. Nature 469: 323–335.
4. Cuervo AM (2011) Chaperone-mediated autophagy: Dice's 'wild' idea about lysosomal selectivity. Nat Rev Mol Cell Biol 12: 535–541.
5. Rubinsztein DC, Marino G, Kroemer G (2011) Autophagy and aging. Cell 146: 682–695.
6. Mizushima N, Komatsu M (2011) Autophagy: renovation of cells and tissues. Cell 147: 728–741.
7. Nakatogawa H, Suzuki K, Kamada Y, Ohsumi Y (2009) Dynamics and diversity in autophagy mechanisms: lessons from yeast. Nat Rev Mol Cell Biol 10: 458–467.
8. Mizushima N, Yoshimori T, Ohsumi Y (2011) The Role of Atg Proteins in Autophagosome Formation. Annu Rev Cell Dev Biol 27: 107–132.
9. Suzuki K, Kubota Y, Sekito T, Ohsumi Y (2007) Hierarchy of Atg proteins in pre-autophagosomal structure organization. Genes Cells 12: 209–218.
10. Itakura E, Mizushima N (2010) Characterization of autophagosome formation site by a hierarchical analysis of mammalian Atg proteins. Autophagy 6: 764–776.
11. Meijer WH, van der Klei IJ, Veenhuis M, Kiel JAKW (2007) *ATG* genes involved in non-selective autophagy are conserved from yeast to man, but the selective Cvt and pexophagy pathways also require organism-specific genes. Autophagy 3: 106–116.
12. Rigden DJ, Michels PA, Ginger ML (2009) Autophagy in protists: Examples of secondary loss, lineage-specific innovations, and the conundrum of remodeling a single mitochondrion. Autophagy 5: 784–794.
13. Duszenko M, Ginger ML, Brennand A, Gualdron-Lopez M, Colombo MI, et al. (2011) Autophagy in protists. Autophagy 7: 127–158.
14. Ichimura Y, Kirisako T, Takao T, Satomi Y, Shimonishi Y, et al. (2000) A ubiquitin-like system mediates protein lipidation. Nature 408: 488–492.
15. Kirisako T, Ichimura Y, Okada H, Kabeya Y, Mizushima N, et al. (2000) The reversible modification regulates the membrane-binding state of Apg8/Aut7 essential for autophagy and the cytoplasm to vacuole targeting pathway. J Cell Biol 151: 263–275.
16. Nakatogawa H, Ichimura Y, Ohsumi Y (2007) Atg8, a ubiquitin-like protein required for autophagosome formation, mediates membrane tethering and hemifusion. Cell 130: 165–178.
17. Xie Z, Nair U, Klionsky DJ (2008) Atg8 controls phagophore expansion during autophagosome formation. Mol Biol Cell 19: 3290–3298.
18. Sou YS, Waguri S, Iwata J, Ueno T, Fujimura T, et al. (2008) The Atg8 conjugation system is indispensable for proper development of autophagic isolation membranes in mice. Mol Biol Cell 19: 4762–4775.
19. Fujita N, Hayashi-Nishino M, Fukumoto H, Omori H, Yamamoto A, et al. (2008) An Atg4B mutant hampers the lipidation of LC3 paralogues and causes defects in autophagosome closure. Mol Biol Cell 19: 4651–4659.
20. Weidberg H, Shvets E, Shpilka T, Shimron F, Shinder V, et al. (2010) LC3 and GATE-16/GABARAP subfamilies are both essential yet act differently in autophagosome biogenesis. EMBO J 29: 1792–1802.
21. Picazarri K, Nakada-Tsukui K, Nozaki T (2008) Autophagy during proliferation and encystation in the protozoan parasite *Entamoeba invadens*. Infect Immun 76: 278–288.
22. Alvarez VE, Kosec G, Sant Anna C, Turk V, Cazzulo JJ, et al. (2008) Blocking autophagy to prevent parasite differentiation: a possible new strategy for fighting parasitic infections? Autophagy 4: 361–363.
23. Alvarez VE, Kosec G, Sant'Anna C, Turk V, Cazzulo JJ, et al. (2008) Autophagy is involved in nutritional stress response and differentiation in *Trypanosoma cruzi*. J Biol Chem 283: 3454–3464.
24. Williams RA, Woods KL, Juliano L, Mottram JC, Coombs GH (2009) Characterization of unusual families of ATG8-like proteins and ATG12 in the protozoan parasite *Leishmania major*. Autophagy 5: 159–172.
25. Besteiro S, Williams RA, Morrison LS, Coombs GH, Mottram JC (2006) Endosome sorting and autophagy are essential for differentiation and virulence of *Leishmania major*. J Biol Chem 281: 11384–11396.
26. Besteiro S, Brooks CF, Striepen B, Dubremetz JF (2011) Autophagy Protein Atg3 is Essential for Maintaining Mitochondrial Integrity and for Normal Intracellular Development of *Toxoplasma gondii* Tachyzoites. PLoS Pathog 7: e1002416.
27. Brennand A, Gualdron-Lopez M, Coppens I, Rigden DJ, Ginger ML, et al. (2011) Autophagy in parasitic protists: unique features and drug targets. Mol Biochem Parasitol 177: 83–99.
28. Kirisako T, Baba M, Ishihara N, Miyazawa K, Ohsumi M, et al. (1999) Formation process of autophagosome is traced with Apg8/Aut7p in yeast. J Cell Biol 147: 435–446.
29. Kabeya Y, Mizushima N, Ueno T, Yamamoto A, Kirisako T, et al. (2000) LC3, a mammalian homologue of yeast Apg8p, is localized in autophagosome membranes after processing. EMBO J 19: 5720–5728.
30. Tawk L, Chicanne G, Dubremetz JF, Richard V, Payrastre B, et al. (2010) Phosphatidylinositol 3-phosphate, an essential lipid in *Plasmodium*, localizes to the food vacuole membrane and the apicoplast. Eukaryot Cell 9: 1519–1530.
31. Mizushima N, Yoshimori T, Levine B (2010) Methods in mammalian autophagy research. Cell 140: 313–326.
32. Waller RF, Reed MB, Cowman AF, McFadden GI (2000) Protein trafficking to the plastid of *Plasmodium falciparum* is via the secretory pathway. EMBO J 19: 1794–1802.
33. Sato S, Rangachari K, Wilson RJ (2003) Targeting GFP to the malarial mitochondrion. Mol Biochem Parasitol 130: 155–158.
34. Sasaki N, Hirai M, Maeda K, Yui R, Itoh K, et al. (2009) The *Plasmodium* HU homolog, which binds the plastid DNA sequence-independent manner, is essential for the parasite's survival. FEBS Lett 583: 1446–1450.
35. Hopkins J, Fowler R, Krishna S, Wilson I, Mitchell G, et al. (1999) The plastid in *Plasmodium falciparum* asexual blood stages: a three-dimensional ultrastructural analysis. Protist 150: 283–295.
36. Kobayashi T, Sato S, Takamiya S, Komaki-Yasuda K, Yano K, et al. (2007) Mitochondria and apicoplast of *Plasmodium falciparum*: behaviour on subcellular fractionation and the implication. Mitochondrion 7: 125–132.
37. van Dooren GG, Marti M, Tonkin CJ, Stimmler LM, Cowman AF, et al. (2005) Development of the endoplasmic reticulum, mitochondrion and apicoplast during the asexual life cycle of Plasmodium falciparum. Mol Microbiol 57: 405–419.
38. Chen D, Fan W, Lu Y, Ding X, Chen S, et al. (2012) A mammalian autophagosome maturation mechanism mediated by TECPR1 and the Atg12-Atg5 conjugate. Mol Cell 45: 629–641.
39. McFadden GI, Roos DS (1999) Apicomplexan plastids as drug targets. Trends Microbiol 7: 328–333.
40. McFadden GI (2011) The apicoplast. Protoplasma 248: 641–650.
41. Eskelinen E-L (2005) Maturation of autophagic vacuoles in mammalian cells. Autophagy 1: 1–10.
42. Ralph SA, van Dooren GG, Waller RF, Crawford MJ, Fraunholz MJ, et al. (2004) Tropical infectious diseases: metabolic maps and functions of the *Plasmodium falciparum* apicoplast. Nat Rev Microbiol 2: 203–216.

43. Kalanon M, McFadden GI (2010) Malaria, *Plasmodium falciparum* and its apicoplast. Biochem Soc Trans 38: 775–782.
44. Stanway RR, Mueller N, Zobiak B, Graewe S, Froehlke U, et al. (2011) Organelle segregation into Plasmodium liver stage merozoites. Cell Microbiol 13: 1768–1782.
45. Kohler S, Delwiche CF, Denny PW, Tilney LG, Webster P, et al. (1997) A plastid of probable green algal origin in Apicomplexan parasites. Science 275: 1485–1489.
46. Janouskovec J, Horak A, Obornik M, Lukes J, Keeling PJ (2010) A common red algal origin of the apicomplexan, dinoflagellate, and heterokont plastids. Proc Natl Acad Sci U S A 107: 10949–10954.
47. Tonkin CJ, Struck NS, Mullin KA, Stimmler LM, McFadden GI (2006) Evidence for Golgi-independent transport from the early secretory pathway to the plastid in malaria parasites. Mol Microbiol 61: 614–630.
48. Tonkin CJ, Kalanon M, McFadden GI (2008) Protein targeting to the malaria parasite plastid. Traffic 9: 166–175.
49. Lim L, Kalanon M, McFadden GI (2009) New proteins in the apicoplast membranes: time to rethink apicoplast protein targeting. Trends Parasitol 25: 197–200.
50. Kalanon M, Tonkin CJ, McFadden GI (2009) Characterization of two putative protein translocation components in the apicoplast of *Plasmodium falciparum*. Eukaryot Cell 8: 1146–1154.
51. Spork S, Hiss JA, Mandel K, Sommer M, Kooij TW, et al. (2009) An unusual ERAD-like complex is targeted to the apicoplast of *Plasmodium falciparum*. Eukaryot Cell 8: 1134–1145.
52. Hailey DW, Rambold AS, Satpute-Krishnan P, Mitra K, Sougrat R, et al. (2010) Mitochondria supply membranes for autophagosome biogenesis during starvation. Cell 141: 656–667.
53. Ravikumar B, Moreau K, Jahreiss L, Puri C, Rubinsztein DC (2010) Plasma membrane contributes to the formation of pre-autophagosomal structures. Nat Cell Biol 12: 747–757.
54. Kovács AL, Pálfia Z, Réz G, Vellai T, Kávacs J (2007) Sequestration revisited: integrating traditional electron microscopy, de novo assembly and new results. Autophagy 3: 655–662.
55. Yla-Anttila P, Vihinen H, Jokitalo E, Eskelinen EL (2009) 3D tomography reveals connections between the phagophore and endoplasmic reticulum. Autophagy 5: 1180–1185.
56. Hayashi-Nishino M, Fujita N, Noda T, Yamaguchi A, Yoshimori T, et al. (2009) A subdomain of the endoplasmic reticulum forms a cradle for autophagosome formation. Nat Cell Biol 11: 1433–1437.
57. Kageyama S, Omori H, Saitoh T, Sone T, Guan JL, et al. (2011) The LC3 recruitment mechanism is separate from Atg9L1-dependent membrane formation in the autophagic response against *Salmonella*. Mol Biol Cell.
58. Itakura E, Kishi-Itakura C, Koyama-Honda I, Mizushima N (2012) Structures containing Atg9A and the ULK1 complex independently target depolarized mitochondria at initial stages of Parkin-mediated mitophagy. J Cell Sci 125: 1488–1499.
59. Suzuki K, Kirisako T, Kamada Y, Mizushima N, Noda T, et al. (2001) The pre-autophagosomal structure organized by concerted functions of *APG* genes is essential for autophagosome formation. EMBO J 20: 5971–5981.
60. Mizushima N, Yamamoto A, Hatano M, Kobayashi Y, Kabeya Y, et al. (2001) Dissection of autophagosome formation using Apg5-deficient mouse embryonic stem cells. J Cell Biol 152: 657–667.
61. Trager W, Jensen JB (1976) Human malaria parasites in continuous culture. Science 193: 673–675.
62. Lambros C, Vanderberg JP (1979) Synchronization of *Plasmodium falciparum* erythrocytic stages in culture. J Parasitol 65: 418–420.
63. Kyes S, Pinches R, Newbold C (2000) A simple RNA analysis method shows var and rif multigene family expression patterns in *Plasmodium falciparum*. Mol Biochem Parasitol 105: 311–315.
64. Dzikowski R, Frank M, Deitsch K (2006) Mutually exclusive expression of virulence genes by malaria parasites is regulated independently of antigen production. PLoS Pathog 2: e22.
65. Tsuboi T, Takeo S, Iriko H, Jin L, Tsuchimochi M, et al. (2008) Wheat germ cell-free system-based production of malaria proteins for discovery of novel vaccine candidates. Infect Immun 76: 1702–1708.
66. Tsuji M, Mattei D, Nussenzweig RS, Eichinger D, Zavala F (1994) Demonstration of heat-shock protein 70 in the sporozoite stage of malaria parasites. Parasitol Res 80: 16–21.
67. Ito D, Han ET, Takeo S, Thongkukiatkul A, Otsuki H, et al. (2011) Plasmodial ortholog of *Toxoplasma gondii* rhoptry neck protein 3 is localized to the rhoptry body. Parasitol Int 60: 132–138.
68. DuBridge RB, Tang P, Hsia HC, Leong PM, Miller JH, et al. (1987) Analysis of mutation in human cells by using an Epstein-Barr virus shuttle system. Mol Cell Biol 7: 379–387.
69. Cao J, Kaneko O, Thongkukiatkul A, Tachibana M, Otsuki H, et al. (2009) Rhoptry neck protein RON2 forms a complex with microneme protein AMA1 in *Plasmodium falciparum* merozoites. Parasitol Int 58: 29–35.
70. Culvenor JG, Day KP, Anders RF (1991) *Plasmodium falciparum* ring-infected erythrocyte surface antigen is released from merozoite dense granules after erythrocyte invasion. Infect Immun 59: 1183–1187.
71. Ribaut C, Berry A, Chevalley S, Reybier K, Morlais I, et al. (2008) Concentration and purification by magnetic separation of the erythrocytic stages of all human *Plasmodium* species. Malar J 7: 45.
72. Hayashi M, Taniguchi S, Ishizuka Y, Kim HS, Wataya Y, et al. (2001) A homologue of N-ethylmaleimide-sensitive factor in the malaria parasite *Plasmodium falciparum* is exported and localized in vesicular structures in the cytoplasm of infected erythrocytes in the brefeldin A-sensitive pathway. J Biol Chem 276: 15249–15255.

nature genetics

Plasmodium cynomolgi genome sequences provide insight into *Plasmodium vivax* and the monkey malaria clade

Shin-Ichiro Tachibana[1,13], Steven A Sullivan[2], Satoru Kawai[3], Shota Nakamura[4], Hyunjae R Kim[2], Naohisa Goto[4], Nobuko Arisue[5], Nirianne M Q Palacpac[5], Hajime Honma[1,5], Masanori Yagi[5], Takahiro Tougan[5], Yuko Katakai[6], Osamu Kaneko[7], Toshihiro Mita[8], Kiyoshi Kita[9], Yasuhiro Yasutomi[10], Patrick L Sutton[2], Rimma Shakhbatyan[2], Toshihiro Horii[5], Teruo Yasunaga[4], John W Barnwell[11], Ananias A Escalante[12], Jane M Carlton[2,14] & Kazuyuki Tanabe[1,5,14]

P. cynomolgi, a malaria-causing parasite of Asian Old World monkeys, is the sister taxon of _P. vivax_, the most prevalent malaria-causing species in humans outside of Africa. Because _P. cynomolgi_ shares many phenotypic, biological and genetic characteristics with _P. vivax_, we generated draft genome sequences for three _P. cynomolgi_ strains and performed genomic analysis comparing them with the _P. vivax_ genome, as well as with the genome of a third previously sequenced simian parasite, _Plasmodium knowlesi_. Here, we show that genomes of the monkey malaria clade can be characterized by copy-number variants (CNVs) in multigene families involved in evasion of the human immune system and invasion of host erythrocytes. We identify genome-wide SNPs, microsatellites and CNVs in the _P. cynomolgi_ genome, providing a map of genetic variation that can be used to map parasite traits and study parasite populations. The sequencing of the _P. cynomolgi_ genome is a critical step in developing a model system for _P. vivax_ research and in counteracting the neglect of _P. vivax_.

Human malaria is transmitted by anopheline mosquitoes and is caused by four species in the genus *Plasmodium*. Of these, *P. vivax* is the major malaria agent outside of Africa, annually causing 80–100 million cases[1]. Although *P. vivax* infection is often mistakenly regarded as benign and self-limiting, *P. vivax* treatment and control present challenges distinct from those of the more virulent *Plasmodium falciparum*. Biological traits, including a dormant (hypnozoite) liver stage responsible for recurrent infections (relapses), early infective sexual stages (gametocytes) and transmission from low parasite densities in the blood[2], coupled with emerging antimalarial drug resistance[3], render *P. vivax* resilient to modern control strategies. Recent evidence indicates that *P. falciparum* derives from parasites of great apes in Africa[4], whereas *P. vivax* is more closely related to parasites of Asian Old World monkeys[5–7], although not itself infective of these monkeys.

P. vivax cannot be cultured *in vitro*, and the small New World monkeys capable of hosting it are rare and do not provide an ideal model system. *P. knowlesi*, an Asian Old World monkey parasite recently recognized as a zoonosis for humans[8], has had its genome sequenced[9], but the species is distantly related to *P. vivax* and is phenotypically dissimilar. In contrast, *P. cynomolgi*, a simian parasite that can infect humans experimentally[10], is the closest living relative (a sister taxon) to *P. vivax* and possesses most of the same genetic, phenotypic and biological characteristics—notably, periodic relapses caused by dormant hypnozoites, early infectious gametocyte formation and invasion of Duffy blood group–positive reticulocytes. *P. cynomolgi* thus offers a robust model for *P. vivax* in a readily available laboratory host, the Rhesus monkey, whose genome was recently sequenced[11]. Here, we report draft genome sequences of three *P. cynomolgi* strains and comparative genomic analyses of *P. cynomolgi*, *P. vivax*[12] and *P. knowlesi*[9], three members of the monkey malaria clade.

We sequenced the genome of *P. cynomolgi* strain B, isolated from a monkey in Malaysia and grown in splenectomized monkeys (Online Methods). A combination of Sanger, Roche 454 and Illumina chemistries was employed to generate a high-quality reference assembly at 161-fold coverage, consisting of 14 supercontigs (corresponding to the 14 parasite chromosomes) and ~1,649 unassigned contigs, comprising

[1]Laboratory of Malariology, Research Institute for Microbial Diseases, Osaka University, Suita, Japan. [2]Department of Biology, Center for Genomics and Systems Biology, New York University, New York, New York, USA. [3]Laboratory of Tropical Medicine and Parasitology, Institute of International Education and Research, Dokkyo Medical University, Shimotsuga, Japan. [4]Genome Information Research Center, Research Institute for Microbial Diseases, Osaka University, Suita, Japan. [5]Department of Molecular Protozoology, Research Institute for Microbial Diseases, Osaka University, Suita, Japan. [6]The Corporation for Production and Research of Laboratory Primates, Tsukuba, Japan. [7]Department of Protozoology, Institute of Tropical Medicine (NEKKEN) and Global COE (Centers of Excellence) Program, Nagasaki University, Nagasaki, Japan. [8]Department of Molecular and Cellular Parasitology, Graduate School of Medicine, Juntendo University, Tokyo, Japan. [9]Department of Biomedical Chemistry, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan. [10]Tsukuba Primate Research Center, National Institute of Biomedical Innovation, Tsukuba, Japan. [11]Center for Global Health, Centers for Disease Control and Prevention, Divison of Parasitic Diseases and Malaria, Atlanta, Georgia, USA. [12]Center for Evolutionary Medicine and Informatics, The Biodesign Institute, Arizona State University, Tempe, Arizona, USA. [13]Present address: Career-Path Promotion Unit for Young Life Scientists, Kyoto University, Kyoto, Japan. [14]These authors jointly directed this work. Correspondence should be addressed to K.T. (kztanabe@biken.osaka-u.ac.jp) or J.M.C. (jane.carlton@nyu.edu).

Received 25 January; accepted 9 July; published online 5 August 2012; doi:10.1038/ng.2375

a total length of ~26.2 Mb (**Supplementary Table 1**). Comparing genomic features of *P. cynomolgi*, *P. knowlesi* and *P. vivax* reveals many similarities, including GC content (mean GC content of 40.5%), 14 positionally conserved centromeres and the presence of intrachromosomal telomeric sequences (ITSs; GGGTT(T/C)A), which were discovered in the *P. knowlesi* genome[9] but are absent in *P. vivax* (**Fig. 1**, **Table 1** and **Supplementary Table 2**).

We annotated the *P. cynomolgi* strain B genome using a combination of *ab initio* gene prediction programs trained on high-quality data sets and sequence similarity searches against the annotated *P. vivax* and *P. knowlesi* genomes. Not unexpectedly for species from the same monkey malaria clade, gene synteny along the 14 chromosomes is highly conserved, although numerous microsyntenic breaks are present in regions containing multigene families (**Fig. 2** and **Table 2**). This genome-wide view of synteny in six species of *Plasmodium* also identified two apparent errors in existing public sequence databases: an inversion in chromosome 3 of *P. knowlesi* and an inversion in chromosome 6 of *P. vivax*. The *P. cynomolgi* genome contains 5,722 genes, of which approximately half encode conserved hypothetical proteins of unknown function, as is the case in all the *Plasmodium* genomes sequenced to date. A maximum-likelihood phylogenetic tree constructed using 192 conserved ribosomal and translation- and transcription-related genes (**Supplementary Fig. 1**) confirms the close relationship of *P. cynomolgi* to *P. vivax* compared to five other *Plasmodium* species. Approximately 90% of genes (4,613) have reciprocal best-match orthologs in all three species (**Fig. 3**), enabling refinement of the existing *P. vivax* and *P. knowlesi* annotations (**Supplementary Table 3**). The high degree of gene orthology enabled us to identify specific examples of gene duplication (an important vehicle for genome evolution), including a duplicated homolog of *P. vivax Pvs28*—which encodes a sexual stage surface antigen that is a transmission-blocking vaccine candidate[13]—in *P. cynomolgi* (**Supplementary Table 4**). Genes common only to *P. cynomolgi* and *P. vivax* (n = 214) outnumber those that are restricted to *P. cynomolgi* and *P. knowlesi* (n = 100) or *P. vivax* and *P. knowlesi* (n = 17). Such figures establish the usefulness of *P. cynomolgi* as a model species for studying the more intractable *P. vivax*.

Notably, most of the genes specific to a particular species belong to multigene families (excluding hypothetical genes; **Table 2** and **Supplementary Table 5**). This suggests repeated lineage-specific gene duplication and/or gene deletion in multigene families within the three monkey malaria clade species. Moreover, copy numbers of the genes composing multigene families were generally greater in the *P. cynomolgi*–*P. vivax* lineage than in *P. knowlesi*, suggesting repeated gene duplication in the ancestral lineage of *P. cynomolgi* and *P. vivax* (or repeated gene deletion in the *P. knowlesi* lineage). Thus, the genomes of *P. cynomolgi*, *P. vivax* and *P. knowlesi* can largely be distinguished by variations in the copy number of multigene family members. Examples of such families include those that encode proteins involved in evasion of the human immune system (*vir*, *kir* and *SICAvar*) and invasion of host red blood cells (*dbp* and *rbp*).

In malaria-causing parasites, invasion of host erythrocytes, mediated by specific interactions between parasite ligands and erythrocyte receptors, is a crucial component of the parasite lifecycle. Of great interest are the *ebl* and *rbl* gene families, which encode parasite ligands required for the recognition of host erythrocytes. The *ebl* genes encode erythrocyte binding–like (EBL) ligands such as the Duffy-binding proteins (DBPs) that bind to Duffy antigen receptor for chemokines (DARC) on human and monkey erythrocytes. The *rbl* genes encode the reticulocyte binding–like (RBL) protein family, including reticulocyte-binding proteins (RBPs) in *P. cynomolgi* and *P. vivax*, and normocyte-binding proteins (NBPs) in *P. knowlesi*, which bind to unknown erythrocyte receptors[14]. We confirmed the presence of two *dbp* genes in *P. cynomolgi*[15] (**Supplementary Table 6**), in contrast to the one *dbp* and three *dbp* genes identified in *P. vivax* and *P. knowlesi*, respectively. This raises an intriguing hypothesis that *P. vivax* lost one *dbp* gene, and thus its infectivity of Old World monkey erythrocytes, after divergence from a common *P. vivax*–*P. cynomolgi* ancestor. This hypothesis is also supported by our identification of single-copy *dbp* genes in two other closely related Old World monkey malaria-causing parasites, *Plasmodium fieldi* and *Plasmodium simiovale*, which are incapable of infecting humans[16]. These two Old World monkey species lost one or more *dbp* genes during divergence that confer infectivity to humans, whereas *P. cynomolgi* and *P. knowlesi* retained *dbp* genes that allow invasion of human erythrocytes (**Supplementary Fig. 2**).

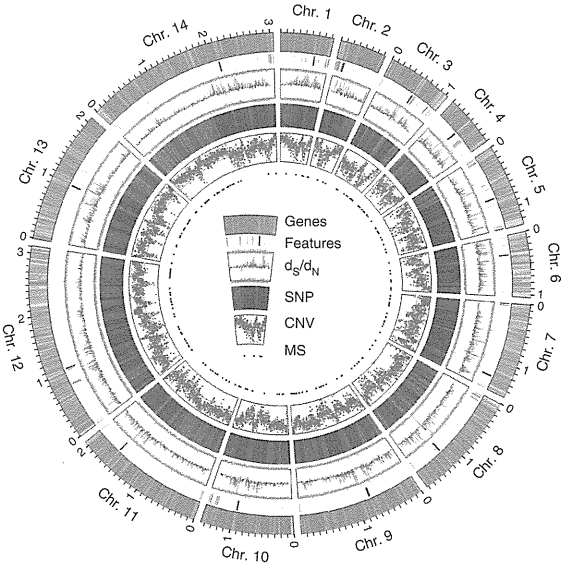

Figure 1 Architecture of the *P. cynomolgi* genome and associated genome-wide variation data. Data are shown for each of the 14 *P. cynomolgi* chromosomes. The six concentric rings, from outermost to innermost, represent (i) the location of 5,049 *P. cynomolgi* genes, excluding those on small contigs (cyan lines); (ii) genome features, including 14 centromeres (thick black lines), 43 telomeric sequence repeats (short red lines), 43 tRNA genes (red lines), 10 rRNAs (dark blue lines) and several gene family members, including 53 *cyir* (dark green lines), 8 *rbp* (brown lines), 13 *sera* (serine-rich antigen; pink lines), 25 *trag* (tryptophan-rich antigen; purple lines), 12 *msp3* (merozoite surface protein 3; light gray lines), 13 *msp7* (merozoite surface protein 7; gray lines), 25 *rad* (silver lines), 8 *etramp* (orange lines), 16 *Pf-fam-b* (light blue lines) and 7 *Pv-fam-d* (light green lines); (iii) plot of d_S/d_N for 4,605 orthologs depicting genome-wide polymorphism within *P. cynomolgi* strains B and Berok (black line) and divergence between *P. cynomolgi* strains B and Berok and *P. vivax* Salvador I (red line); a track above the plot indicates *P. cynomolgi* genes under positive selection (red) and purifying selection (blue), and a track below the plot indicates *P. cynomolgi*–*P. vivax* orthologs under positive selection (red) and purifying selection (blue); (iv) heatmap indicating SNP density of 3 *P. cynomolgi* strains plotted per 10-kb windows: red, 0–83 SNPs per 10 kb (regions of lowest SNP density); blue, 84–166 SNPs per 10 kb; green, 167–250 SNPs per 10 kb; purple, 251–333 SNPs per 10 kb; orange, 334–416 SNPs per10 kb; yellow, 417–500 SNPs per10 kb (regions of highest SNP density); (v) $\log_2$ ratio plot of CNVs identified from a comparison of *P. cynomolgi* strains B and Berok; and (vi) map of 182 polymorphic intergenic microsatellites (MS, black dots). The figure was generated using Circos software (see URLs).

Table 1 Comparison of genome features between *P. cynomolgi*, *P. vivax* and *P. knowlesi*, three species of the monkey malaria clade

Feature	*P. cynomolgi*	*P. vivax*[12]	*P. knowlesi*[9]
Assembly			
Size (Mb)	26.2	26.9	23.7
Number of scaffolds[a]	14 (1,649)	14 (2,547)	14 (67)
Coverage (fold)	161	10	8
GC content (%)	40.4	42.3	38.8
Genes			
Number of genes	5,722	5,432	5,197
Mean gene length (bp)	2,240	2,164	2,180
Gene density (bp per gene)[b]	4,428.2	4,950.5	4,416.1
Percentage coding[b]	51.0	47.1	49.0
Structural RNAs			
Number of tRNA genes	43	44	41
Number of 5S rRNA genes	3	3	0[c]
Number of 5.8S, 18S and 28S rRNA units	7	7	5
Nuclear genome			
Number of chromosomes	14	14	14
Number of centromeres	14	14	14
Isochore structure[d]	+	+	–
Mitochondrial genome			
Size (bp)[e]	5,986 (AB444123)	5,990 (AY598140)	5,958 (AB444108)
GC content (%)	30.3	30.5	30.5
Apicoplast genome			
Size (bp)	29,297[f]	5,064[g]	N/A
GC content (%)	13.0	17.1	N/A

N/A, not available.
[a]Small unassigned contigs indicated in parentheses. [b]Sequence gaps excluded. [c]Not present in *P. knowlesi* assembly version 4.0. [d]Regions of the genome that differ in density and are separable by CsCl centrifugation; isochores correspond to domains differing in GC content. [e]Identified in other studies (GenBank accessions given in parentheses). [f]Partial sequence (~86% complete) generated during this project. [g]Partial sequence of reference genome only published[12]; actual size is ~35 kb.

We found multiple *rbp* genes, some truncated or present as pseudogenes, in the *P. cynomolgi* genome (**Fig. 1** and **Table 2**). Phylogenetic analysis showed that *rbl* genes from *P. cynomolgi*, *P. vivax* and *P. knowlesi* can be classified into three distinct groups, RBP/NBP-1, RBP/NBP-2 and RBP/NBP-3 (**Supplementary Fig. 3**), and suggests that these groups existed before the three species diverged. All three groups of RBP/NBP are represented in *P. cynomolgi*, whereas *P. vivax* and *P. knowlesi* lack functional genes from the RBP/NBP-3 and RBP/NBP-1 groups, respectively. Thus, *rbl* gene family expansion seems to have occurred after speciation, indicating that the three species have multiple species-specific erythrocyte invasion mechanisms. Notably, we found an ortholog of *P. vivax rbp1b* in some strains of *P. cynomolgi* but not in others (**Supplementary Table 6**). To our knowledge, this is the first example of a CNV for a *rbp* gene between strains of a single *Plasmodium* species, highlighting how repeated creation and destruction of *rbl* genes, a signature of adaptive evolution, may have enabled species of the monkey malaria clade to expand or switch between monkey and human hosts.

The largest gene family in *P. cynomolgi*, consisting of 256 *cyir* (cynomolgi-interspersed repeat) genes, is part of the *pir* (plasmodium-interspersed repeat) superfamily that includes *P. vivax vir* genes (n = 319) and *P. knowlesi kir* genes (n = 70) (**Table 2**). *Pir*-encoded proteins reside on the surface of infected erythrocytes and have an important role in immune evasion[17]. Most *cyir* genes have sequence similarity to *P. vivax vir* genes (n = 254; **Supplementary Table 7**) and are found in subtelomeric regions (**Fig. 1**), but, notably, 11 *cyir* genes have sequence similarity to *P. knowlesi kir* genes (**Supplementary Table 7**) and occur more internally in the chromosomes, as do the *kir* genes in *P. knowlesi*. As with 'molecular mimicry' in *P. knowlesi* (mimicry of host sequences by pathogen sequences)[9], one CYIR protein (encoded by PCYB_032250) has a region of 56 amino acids that is highly similar to the extracellular domain of primate CD99 (**Supplementary Fig. 4**), a molecule involved in the regulation of T-cell function. A new finding is that *P. cynomolgi* has two genes whose sequences are similar to *P. knowlesi SICAvar* genes (**Supplementary Table 7**) that are expressed on the surfaces of schizont-infected macaque erythrocytes and are involved in antigenic variation[18].

The ability to form a dormant hypnozoite stage is common to both *P. cynomolgi* and *P. vivax* and was first shown in laboratory infections of monkeys by mosquito-transmitted *P. cynomolgi*[19]. In a search for candidate genes involved in the hypnozoite stage, we identified nine coding for 'dormancy-related' proteins that had the upstream ApiAP2 motifs[20] necessary for stage-specific transcriptional regulation at the sporozoite (pre-hypnozoite) stage (**Supplementary Table 8**). The candidates include kinases that are involved in cell cycle transition; hypnozoite formation may be regulated by phosphorylation of proteins specifically expressed at the pre-hypnozoite stage. Our list of *P. cynomolgi* candidate genes represents an informed starting point for experimental studies of this elusive stage.

We sequenced *P. cynomolgi* strains Berok (from Malaysia) and Cambodian (from Cambodia) to 26× and 17× coverage, respectively, to characterize *P. cynomolgi* genome-wide diversity through analysis of SNPs, CNVs and microsatellites. A comparison of the three *P. cynomolgi* strains identified 178,732 SNPs (**Supplementary Table 9**) at a frequency of 1 SNP per 151 bp, a polymorphism level somewhat

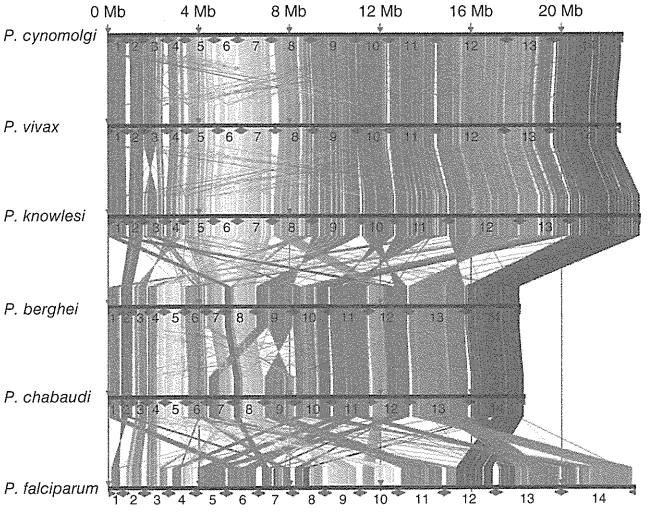

Figure 2 Genome synteny between six species of *Plasmodium* parasite. Protein-coding genes of *P. cynomolgi* are shown aligned with those of five other *Plasmodium* genomes: two species belonging to the monkey malaria clade, *P. vivax* and *P. knowlesi*; two species of rodent malaria, *P. berghei* and *P. chabaudi*; and *P. falciparum*. Highly conserved protein-coding regions between the genomes are colored in order from red (5′ end of chromosome 1) to blue (3′ end of chromosome 14) with respect to genomic position of *P. cynomolgi*.

Table 2 Components of multigene families of *P. cynomolgi*, *P. vivax* and *P. knowlesi* differ in copy number

Family	Multigene family	Localization	Arrangement	*P. cynomolgi*	*P. vivax*	*P. knowlesi*	Putative function and other information
1	*pir* (*vir*-like)	Subtelomeric	Scattered and clustered	254	319[a]	4	Immune evasion
2	*pir* (*kir*-like)	Subtelomeric and central	Scattered and clustered	11	2	66[a]	Immune evasion
3	*SICAvar*	Subtelomeric and central	Scattered and clustered	2	1	242[a]	Antigenic variation, immune evasion
4	*msp3*	Central	Clustered	12	12	3	Merozoite surface protein
5	*msp7*	Central	Clustered	13	13	5	Merozoite surface protein
6	*dbl* (*dbp/ebl*)	Subtelomeric	Scattered	2	1	3	Host cell recognition
7	*rbl* (*rbp/nbp/rh*)	Subtelomeric	Scattered	8[a]	10[a]	3[a]	Host cell recognition
8	*Pv-fam-a* (*trag*)	Subtelomeric	Scattered and clustered	36	36	26[a]	Tryptophan-rich
9	*Pv-fam-b*	Central	Clustered	3	6	1	Unknown
10	*Pv-fam-c*	Subtelomeric	Unknown[b]	1	7	0	Unknown
11	*Pv-fam-d* (*hypb*)	Subtelomeric	Scattered	18	16	2	Unknown
12	*Pv-fam-e* (*rad*)	Subtelomeric	Clustered	27	44	16	Unknown
13	*Pv-fam-g*	Central	Clustered	3	3	3	Unknown
14	*Pv-fam-h* (*hyp16*)	Central	Clustered	6	4	2	Unknown
15	*Pv-fam-i* (*hyp11*)	Subtelomeric	Scattered	6	6	5	Unknown
16	*Pk-fam-a*	Central	Scattered	0	0	12[a]	Unknown
17	*Pk-fam-b*	Subtelomeric	Scattered	0	0	9	Unknown
18	*Pk-fam-c*	Subtelomeric	Scattered	0	0	6[a]	Unknown
19	*Pk-fam-d*	Central	Scattered	0	0	3[a]	Unknown
20	*Pk-fam-e*	Subtelomeric	Scattered	0	0	3[a]	Unknown
21	*PST-A*	Subtelomeric and central	Scattered	9[a]	11[a]	7	αβ hydrolase
22	*ETRAMP*	Subtelomeric	Scattered	9	9	9	Parasitophorous vacuole membrane
23	*CLAG* (*RhopH-1*)	Subtelomeric	Scattered	2	3	2	High-molecular-wieght rhoptry antigen complex
24	*PvSTP1*	Subtelomeric	Unknown[b]	3	10[a]	0	Unknown
25	*PHIST* (Pf-*fam-b*)	Subtelomeric	Scattered and clustered	21	20	15	Unknown
26	*SERA*	Central	Clustered	13[a]	13[a]	8[a]	Cysteine protease

[a]Pseudogenes, truncated genes and gene fragments included. [b]Gene arrangement could not be determined due to localization on unassigned contigs.

similar to that found when *P. falciparum* genomes are compared[21,22]. We calculated the pairwise nucleotide diversity (π) as 5.41×10^{-3} across the genome, which varies little between the chromosomes. We assessed genome-wide CNVs between the *P. cynomolgi* B and Berok strains, using a robust statistical model in the CNV-seq program[23], by which we identified 1,570 CNVs (1 per 17 kb), including 1 containing the *rbp1b* gene on chromosome 7 (**Supplementary Fig. 5**). Finally, mining of the *P. cynomolgi* B and Berok strains identified 182 polymorphic intergenic microsatellites (**Supplementary Table 10**), the first set of genetic markers developed for this species. These provide a toolkit for studies of genetic diversity and population structure of laboratory stocks or natural infections of *P. cynomolgi*, many of which have recently been isolated from screening hundreds of wild monkeys for the zoonosis *P. knowlesi*[24].

We estimated the difference between the number of synonymous changes per synonymous site (d_S) and the number of nonsynonymous changes per nonsynonymous site (d_N) over 4,563 pairs of orthologs within *P. cynomolgi* strains B and Berok and 4,601 pairs of orthologs between these two *P. cynomolgi* strains and *P. vivax* Salvador I, using a simple Nei-Gojobori model[25]. We found 63 genes with $d_N > d_S$ within the two *P. cynomolgi* strains and 3,265 genes with $d_S > d_N$ (**Supplementary Table 11**). Genes with relatively high d_N/d_S ratios include those encoding transmembrane proteins, such as antigens and transporters, among which is a transmission-blocking target antigen, Pcyn230 (encoded by PCYB_042090). Notably, the *P. vivax* ortholog (PVX_003905) does not show evidence for positive selection[26], suggesting species-specific positive selection. We explored the degree to which evolution of orthologs has been constrained between *P. cynomolgi* and *P. vivax* and found 83 genes under possible accelerated evolution but 3,739 genes under possible purifying selection (**Supplementary Table 12**). This conservative estimate indicates that at least 81% of loci have diverged under strong constraint, compared with 1.8% of loci under less constraint or positive selection (**Fig. 1**), indicating that, overall, the genome of *P. cynomolgi* is highly conserved in single-locus genes compared to *P. vivax* and emphasizing the value of *P. cynomolgi* as a biomedical and evolutionary model for studying *P. vivax*.

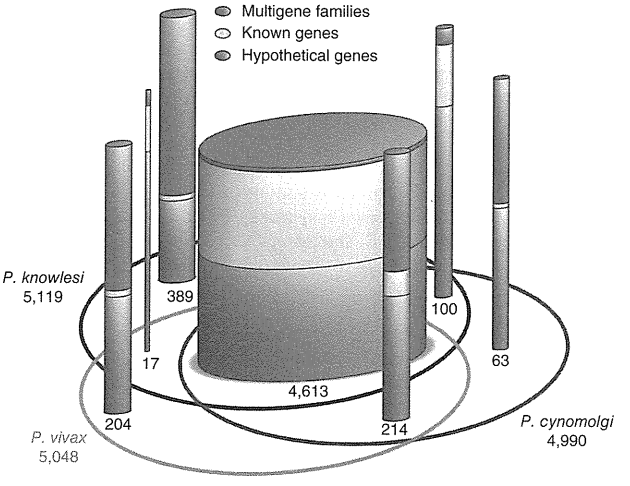

Figure 3 Comparison of the genes of *P. cynomolgi*, *P. vivax* and *P. knowlesi*. The Venn ellipses represent the three genomes, with the total number of genes assigned to the chromosomes indicated under the species name. Cylinders depict orthologous and non-orthologous genes between the three genomes, with the number of genes in each indicated and represented graphically by cylinder relative width. In each cylinder, genes are divided into three categories whose thickness is represented by colored bands proportional to category percentage.

Our generation of the first *P. cynomolgi* genome sequences is a critical step in the development of a robust model system for the intractable and neglected *P. vivax* species[27]. Comparative genome analysis of *P. vivax* and the Old World monkey malaria-causing parasites *P. cynomolgi* and *P. knowlesi* presented here provides the foundation for further insights into traits such as host specificity that will enhance prospects for the eventual elimination of vivax-caused malaria and global malaria eradication.

URLs. PlasmoDB, http://plasmodb.org/; Circos, http://circos.ca/; MIcroSAtelite Identification tool (MISA), http://pgrc.ipk-gatersleben.de/misa/; dbSNP, http://www.ncbi.nlm.nih.gov/projects/SNP/snp_viewBatch.cgi?sbid=1056645.

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. Sequence data for the *P. cynomolgi* B, Cambodian and Berok strains have been deposited in the DNA Data Bank of Japan (DDBJ), the European Molecular Biology Laboratory (EMBL) and the GenBank databases under the following accessions: B strain sequence reads DRA000196, genome assembly BAEJ01000001–BAEJ01003341 and annotation DF157093–DF158755; Cambodian strain sequence reads DRA000197; and Berok strain sequence reads SRA047950. SNP calls have been submitted to dbSNP (NYU_CGSB_BIO; 1056645) and may also be downloaded from the dbSNP website (see URLs). Sequences of the *dbp* genes from *P. cynomolgi* (Cambodian strain), *P. fieldi* (A.b.i. strain) and *P. simiovale* (AB617788–AB617791) and the *P. cynomolgi* Berok strain (JQ422035–JQ422036) and *rbp* gene sequences from the *P. cynomolgi* Berok and Cambodian strains (JQ422037–JQ422050) have been deposited. A partial apicoplast genome of the *P. cynomolgi* Berok strain has been deposited (JQ522954). The *P. cynomolgi* B reference genome is also available through PlasmoDB (see URLs).

Note: Supplementary information is available in the online version of the paper.

ACKNOWLEDGMENTS

We thank H. Sawai for suggestions on genome analysis, D. Fisher for help with genome-wide evolutionary analyses and the NYU Langone Medical Center Genome Technology Core for access to Roche 454 sequencing equipment (funded by grant S10 RR026950 to J.M.C. from the US National Institutes of Health (NIH)). Genome and phylogenetic analyses used the Genome Information Research Center in the Research Institute of Microbial Diseases at Osaka University. This work was supported by grants from the Ministry of Education, Culture, Sports, Science and Technology of Japan (18073013, 18GS03140013, 20390120 and 22406012) to K.T., an NIH grant (R01 GM080586) to A.A.E. and a Burroughs Wellcome Fund grant (1007398) and an NIH International Centers of Excellence for Malaria Research grant (U19 AI089676-01) to J.M.C. The content is soley the responsibility of the authors and does not necessarily represent the official views of the NIH.

AUTHOR CONTRIBUTIONS

K.T., J.M.C., A.A.E. and J.W.B. conceived and conducted the study. S.K., Y.K., Y.Y., S.-I.T. and J.W.B. provided *P. cynomolgi* material. S.N., N.G., T.Y. and H.R.K. constructed a computing system for data processing, and S.-I.T., H.H., P.L.S., S.A.S. and H.R.K. performed scaffolding of contigs and manual annotation of the predicted genes. S.N. performed sequence correction of supercontigs and gene prediction. S.-I.T., S.N., N.G., N.A., M.Y., O.K., K.T., H.R.K., R.S., S.A.S. and J.M.C. analyzed data. S.-I.T., N.M.Q.P., T.T., T.M., K.K., J.M.C., T.H., A.A.E., J.W.B. and K.T. wrote the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/doifinder/10.1038/ng.2375.

1. Mendis, K., Sina, B.J., Marchesini, P. & Carter, R. The neglected burden of *Plasmodium vivax* malaria. *Am. J. Trop. Med. Hyg.* **64**, 97–106 (2001).
2. Mueller, I. *et al.* Key gaps in the knowledge of *Plasmodium vivax*, a neglected human malaria parasite. *Lancet Infect. Dis.* **9**, 555–566 (2009).
3. Baird, J.K. Resistance to chloroquine unhinges vivax malaria therapeutics. *Antimicrob. Agents Chemother.* **55**, 1827–1830 (2011).
4. Rayner, J.C., Liu, W., Peeters, M., Sharp, P.M. & Hahn, B.H. A plethora of *Plasmodium* species in wild apes: a source of human infection? *Trends Parasitol.* **27**, 222–229 (2011).
5. Cornejo, O.E. & Escalante, A.A. The origin and age of *Plasmodium vivax*. *Trends Parasitol.* **22**, 558–563 (2006).
6. Escalante, A.A. *et al.* A monkey's tale: the origin of *Plasmodium vivax* as a human malaria parasite. *Proc. Natl. Acad. Sci. USA* **102**, 1980–1985 (2005).
7. Mu, J. *et al.* Host switch leads to emergence of *Plasmodium vivax* malaria in humans. *Mol. Biol. Evol.* **22**, 1686–1693 (2005).
8. Singh, B. *et al.* A large focus of naturally acquired *Plasmodium knowlesi* infections in human beings. *Lancet* **363**, 1017–1024 (2004).
9. Pain, A. *et al.* The genome of the simian and human malaria parasite *Plasmodium knowlesi*. *Nature* **455**, 799–803 (2008).
10. Eyles, D.E., Coatney, G.R. & Getz, M.E. Vivax-type malaria parasite of macaques transmissible to man. *Science* **131**, 1812–1813 (1960).
11. Gibbs, R.A. *et al.* Evolutionary and biomedical insights from the rhesus macaque genome. *Science* **316**, 222–234 (2007).
12. Carlton, J.M. *et al.* Comparative genomics of the neglected human malaria parasite *Plasmodium vivax*. *Nature* **455**, 757–763 (2008).
13. Saxena, A.K., Wu, Y. & Garboczi, D.N. *Plasmodium* p25 and p28 surface proteins: potential transmission-blocking vaccines. *Eukaryot. Cell* **6**, 1260–1265 (2007).
14. Iyer, J., Gruner, A.C., Renia, L., Snounou, G. & Preiser, P.R. Invasion of host cells by malaria parasites: a tale of two protein families. *Mol. Microbiol.* **65**, 231–249 (2007).
15. Okenu, D.M., Malhotra, P., Lalitha, P.V., Chitnis, C.E. & Chauhan, V.S. Cloning and sequence analysis of a gene encoding an erythrocyte binding protein from *Plasmodium cynomolgi*. *Mol. Biochem. Parasitol.* **89**, 301–306 (1997).
16. Coatney, G.R., Collins, W.E., Warren, M. & Contacos, P.G. *The Primate Malarias* (US Department of Health, Education and Welfare, Washington, DC, 1971).
17. Cunningham, D., Lawton, J., Jarra, W., Preiser, P. & Langhorne, J. The *pir* multigene family of *Plasmodium*: antigenic variation and beyond. *Mol. Biochem. Parasitol.* **170**, 65–73 (2010).
18. al-Khedery, B., Barnwell, J.W. & Galinski, M.R. Antigenic variation in malaria: a 3′ genomic alteration associated with the expression of a *P. knowlesi* variant antigen. *Mol. Cell* **3**, 131–141 (1999).
19. Krotoski, W.A. The hypnozoite and malarial relapse. *Prog. Clin. Parasitol.* **1**, 1–19 (1989).
20. Campbell, T.L., De Silva, E.K., Olszewski, K.L., Elemento, O. & Llinas, M. Identification and genome-wide prediction of DNA binding specificities for the ApiAP2 family of regulators from the malaria parasite. *PLoS Pathog.* **6**, e1001165 (2010).
21. Mu, J. *et al.* Genome-wide variation and identification of vaccine targets in the *Plasmodium falciparum* genome. *Nat. Genet.* **39**, 126–130 (2007).
22. Volkman, S.K. *et al.* A genome-wide map of diversity in *Plasmodium falciparum*. *Nat. Genet.* **39**, 113–119 (2007).
23. Xie, C. & Tammi, M.T. CNV-seq, a new method to detect copy number variation using high-throughput sequencing. *BMC Bioinformatics* **10**, 80 (2009).
24. Lee, K.S. *et al.* *Plasmodium knowlesi*: reservoir hosts and tracking the emergence in humans and macaques. *PLoS Pathog.* **7**, e1002015 (2011).
25. Nei, M. & Gojobori, T. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol. Biol. Evol.* **3**, 418–426 (1986).
26. Doi, M. *et al.* Worldwide sequence conservation of transmission-blocking vaccine candidate Pvs230 in *Plasmodium vivax*. *Vaccine* **29**, 4308–4315 (2011).
27. Carlton, J.M., Sina, B.J. & Adams, J.H. Why is *Plasmodium vivax* a neglected tropical disease? *PLoS Negl. Trop. Dis.* **5**, e1160 (2011).

ONLINE METHODS

Parasite material. Details of the origin of the *P. cynomolgi* B, Berok and Cambodian strains, their growth in macaques and isolation of parasite material are given in the **Supplementary Note.**

Genome sequencing and assembly. *P. cynomolgi* B strain was sequenced using the Roche 454 GS FLX (Titanium) and Illumina/Solexa Genome Analyzer IIx platforms to 161× coverage. In addition, 2,784 clones (6.8 Mb) of a ~40-kb insert fosmid library in pCC1FOS (EpiCentre Biotechnologies) was sequenced by the Sanger method. A draft assembly of strain B was constructed using a combination of automated assembly and manual gap closure. We first generated *de novo* contigs by assembling Roche 454 reads using GS *De novo* Assembler version 2.0 with default parameters. Contigs of >500 bp were mapped to the *P. vivax* Salvador I reference assembly[12] (PlasmoDB; see URLs). *P. cynomolgi* contigs were iteratively arrayed through alignment to *P. vivax*–assembled sequences with manual corrections. A total of 1,264 aligned contigs were validated by mapping paired-end reads from fosmid clones using blastn ($e < 1 \times 10^{-15}$; identity > 90%; coverage > 200 bp) implemented in GenomeMatcher software version 1.65 (ref. 28). Additional linkages (699 regions) were made using PCR across the intervening sequence gaps with primers designed from neighboring contigs. The length of sequence gaps was estimated from insert lengths of the fosmid paired-end reads, the size of PCR products and homologous sequences of the *P. vivax* genome. Supercontigs were then manually constructed from the aligned contigs. Eventually, we obtained 14 supercontigs corresponding to the 14 chromosomes of the parasite, with a total length of ~22.73 Mb, encompassing ~80% of the predicted *P. cynomolgi* genome. A total of 1,651 contigs (>1 kb) with a total length of 3.45 Mb was identified as unassigned subtelomeric sequences by searching against the *P. vivax* genome using blastn. Additionally, to improve sequence accuracy, we constructed a mapping assembly of Illumina paired-end reads and the 14 supercontigs and unassigned contigs as reference sequences using CLC Genomics Workbench version 3.0 with default settings (CLC Bio). Comparison of the draft *P. cynomolgi* B sequence with 23 *P. cynomolgi* protein-coding genes (64 kb) obtained by Sanger sequencing showed 99.8% sequence identity (**Supplementary Table 13**). The *P. cynomolgi* Berok and Cambodian strains were sequenced to 26× and 17× coverage, respectively, using the Roche 454 GS FLX platform, with single-end and 3-kb paired-end libraries made for the former and a single-end library only made for the latter. For phylogenetic analyses of specific genes, sequences were independently verified by Sanger sequencing (**Supplementary Table 14** and **Supplementary Note**).

Prediction and annotation of genes. Gene prediction for the 14 supercontigs and 1,651 unassigned contigs was performed using the MAKER genome annotation pipeline[29] with *ab initio* gene prediction programs trained on proteins and ESTs from PlasmoDB Build 7.1. For gene annotation, blastn ($e < 1 \times 10^{-15}$; identity > 70%; coverage > 100 bp) searches of *P. vivax* (PvivaxAnnotated Transcripts_PlasmoDB-7.1.fasta) and *P. knowlesi* (PknowlesiAnnotatedTranscripts_PlasmoDB-7.1.fasta) predicted proteomes were run, and the best hits were identified. All predicted genes were manually inspected at least twice for gene structure and functional annotation, and orthologous relationships between *P. cynomolgi*, *P. vivax* and *P. knowlesi* were determined on synteny. A unique identifier, PCYB_######, was assigned to *P. cynomolgi* genes, where the first two of the six numbers indicate chromosome number. Paralogs of genes that seemed to be specific to either *P. cynomolgi*, *P. vivax* or *P. knowlesi* were searched using blastp with default parameters, using a cutoff *e* value of 1×10^{-16}.

Multiple genome sequence alignment. Predicted proteins of *P. cynomolgi* B strain were concatenated and aligned with those from the 14 chromosomes of 5 other *Plasmodium* genomes: *P. vivax*, *P. knowlesi*, *P. falciparum*, *P. berghei* and *P. chabaudi*, using Murasaki software version 1.68.6 (ref. 30).

Search for sequence showing high similarity to host proteins. Eleven *P. cynomolgi* CYIR proteins (with sequence similarity to *P. knowlesi* KIR) were subjected to blastp search for regions having high similarity to host *Macacca mulatta* CD99 protein, with cutoff *e* value of 1×10^{-12} and compositional adjustment (no adjustment) against the nonredundant protein sequence data set of the *M. mulatta* proteome in NCBI.

Phylogenetic analyses. Genes were aligned using ClustalW version 2.0.10 (ref. 31) with manual corrections, and unambiguously aligned sites were selected for phylogenetic analyses. Maximum-likelihood phylogenetic trees were constructed using PROML programs in PHYLIP version 3.69 (ref. 32) under the Jones-Taylor-Thornton (JTT) amino-acid substitution model. To take the evolutionary rate heterogeneity across sites into consideration, the R (hidden Markov model rates) option was set for discrete γ distribution, with eight categories for approximating the site-rate distribution. CODEML programs in PAML 4.4 (ref. 33) were used for estimating the γ shape parameter, α values. For bootstrap analyses, SEQBOOT and CONSENSE programs in PHYLIP were applied.

Candidate genes for hypnozoite formation. We undertook two approaches. First, genes unique to *P. vivax* and *P. cynomolgi* (hypnozoite-forming parasites) and not found in other non-hypnozoite–forming *Plasmodium* species were identified. We used the 147 unique genes identified in the *P. vivax* genome[12] to search the *P. cynomolgi* B sequence. For the orthologs identified in both species, ~1 kb of sequence 5′ to the coding sequence was searched for four specific ApiAP2 motifs[20]—PF14 0633, GCATGC; PF13_0235_D1, GCCCCG; PFF0670w_D1, TAAGCC; and PFD0985w_D2, TGTTAC—which are involved in sporozoite stage–specific regulation and expression (corresponding to the pre-hypnozoite stage). Second, dormancy-related proteins were retrieved from GenBank and used to search for *P. vivax* homologs. Candidate genes ($n = 128$) and orthologs of *P. cynomolgi* and five other parasite species were searched in the region ~1 kb upstream of the coding sequence for the presence of the four ApiAP2 motifs. Data for *P. vivax*, *P. knowlesi*, *P. falciparum*, *P. berghei*, *Plasmodium chabaudi* and *Plasmodium yoelii* were retrieved from PlasmoDB Build 7.1.

Genome-wide screen for polymorphisms. For SNP identification, alignment of Roche 454 data from strains B, Berok and Cambodian was performed using SSAHA2 (ref. 34), with 0.1 mismatch rate and only unique matches reported. Potential duplicate reads generated during PCR amplification were removed, so that when multiple reads mapped at identical coordinates, only the reads with the highest mapping quality were retained. We used a statistical method[35] implemented in SAMtools version 0.1.18 to call SNPs simultaneously in the case of duplicate runs of the same strain. SNPs with high read depth (>100) were filtered out, as were SNPs in poor alignment regions at the ends of chromosomes (**Supplementary Note**).

Nucleotide diversity (π) was calculated as follows. For each site being compared, we calculated allele frequency by counting the two alleles and measured the proportion of nucleotide differences. Letting π be the genetic distance between allele *i* and allele *j*, then the nucleotide diversity within the population is

$$\pi = \sum_{i,j} P_i P_j \pi_{ij}$$

where P_i and P_j are the overall allele frequencies of *i* and *j*, respectively. Mean π was calculated by averaging over sites, weighting each by $\sum_{i=1}^{n-1} \frac{1}{i}$, where *n* is the number of aligned sites. Average d_N/d_S ratios were estimated using the modified Nei-Gojobori/Jukes-Cantor method in MEGA 4 (ref. 36).

CNV-seq[23] was used to identify potential CNVs in *P. cynomolgi*. Briefly, this method is based on a statistical model that allows confidence assessment of observed copy-number ratios from next-generation sequencing data. Roche 454 sequences from *P. cynomolgi* strain B assembly were used as the reference genome, and the *P. cynomolgi* Berok strain was used as a test genome; the sequence coverage of the Cambodian strain was considered too low for analysis. The test reads were mapped to the reference genome, and CNVs were detected by computing the number of reads for each test strain in a sliding window. The validity of the observed ratios was assessed by the computation of a probability of a random occurrence, given no copy-number variation.

Polymorphic microsatellites (defined as repeat units of 1–6 nucleotides) between *P. cynomolgi* strains B and Berok were identified by aligning contigs

from a *de novo* assembly of Berok (generated using Roche GS Assembler version 2.6, with 40-bp minimum overlap, 90% identity) to the B strain using the Burrows-Wheeler Aligner (BWA)[37] and allowing for gaps. Using the Phred-scaled probability of the base being misaligned by SAMtools[35], indel candidates were called from the alignment. In-house Python scripts were used to then cross-reference with the microsatellites found in the reference strain B assembly identified by MISA (see URLs). All homopolymer microsatellites were discarded to account for potential sequence errors introduced by 454 sequencing.

Selective constraint analysis of 4,563 orthologs between *P. cynomolgi* strains B and Berok and 4,601 orthologs between these strains and *P. vivax* Salvador I used MUSCLE[38] alignments with stringent removal of gaps and missing data (*P. cynomolgi* Berok orthologs were identified through a reciprocal best-hit BLAST search against strain B genes). Analyses were conducted using the Nei-Gojobori model[25]. To detect values that could not be explained by chance, we estimated the standard error by a bootstrap procedure with 200 pseudoreplicates for each gene. The expected value for d_S/d_N is 0 if a given pair of sequences is diverging without obvious effects on fitness. In the case of the comparison within *P. cynomolgi*, values with a difference of ± 2 s.e.m. from 0 were considered indicative of an excess of synonymous ($d_S/d_N > 0$) or nonsynonymous ($d_S/d_N < 0$) changes. In the case of the comparison between *P. cynomolgi* and *P. vivax*, we used a more stringent criterion of ± 3 s.e.m. from 0.

28. Ohtsubo, Y., Ikeda-Ohtsubo, W., Nagata, Y. & Tsuda, M. GenomeMatcher: a graphical user interface for DNA sequence comparison. *BMC Bioinformatics* **9**, 376 (2008).
29. Cantarel, B.L. *et al.* MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Res.* **18**, 188–196 (2008).
30. Popendorf, K., Tsuyoshi, H., Osana, Y. & Sakakibara, Y. Murasaki: a fast, parallelizable algorithm to find anchors from multiple genomes. *PLoS ONE* **5**, e12651 (2010).
31. Thompson, J.D., Higgins, D.G. & Gibson, T.J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680 (1994).
32. Felsenstein, J. *PHYLIP, Phylogeny Inference Package*, 3.6a3 edn (University of Washington, Seattle, 2005).
33. Yang, Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591 (2007).
34. Ning, Z., Cox, A.J. & Mullikin, J.C. SSAHA: a fast search method for large DNA databases. *Genome Res.* **11**, 1725–1729 (2001).
35. Li, H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **27**, 2987–2993 (2011).
36. Tamura, K., Dudley, J., Nei, M. & Kumar, S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* **24**, 1596–1599 (2007).
37. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
38. Edgar, R.C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).

OPEN ACCESS Freely available online

PLoS NEGLECTED TROPICAL DISEASES

Population Structure and Transmission Dynamics of *Plasmodium vivax* in the Republic of Korea Based on Microsatellite DNA Analysis

Moritoshi Iwagami[1], Megumi Fukumoto[1,2], Seung-Young Hwang[3], So-Hee Kim[4], Weon-Gyu Kho[3,4]*, Shigeyuki Kano[1,2]*

1 Department of Tropical Medicine and Malaria, Research Institute, National Center for Global Health and Medicine, Shinjuku, Tokyo, Japan, **2** Graduate School of Comprehensive Human Sciences, University of Tsukuba, Tsukuba, Ibaraki, Japan, **3** Department of Parasitology, College of Medicine, Inje University, Busanjin-gu, Busan, Korea, **4** Department of Malariology, College of Medicine, Paik Institute of Clinical Research, Inje University, Busanjin-gu, Busan, Korea

Abstract

Background: In order to control malaria, it is important to understand the genetic structure of the parasites in each endemic area. *Plasmodium vivax* is widely distributed in the tropical to temperate regions of Asia and South America, but effective strategies for its elimination have yet to be designed. In South Korea, for example, indigenous vivax malaria was eliminated by the late 1970s, but re-emerged from 1993. We estimated the population structure and temporal dynamics of transmission of *P. vivax* in South Korea using microsatellite DNA markers.

Methodology/Principal Findings: We analyzed 255 South Korean *P. vivax* isolates collected from 1994 to 2008, based on 10 highly polymorphic microsatellite DNA loci of the *P. vivax* genome. Allelic data were obtained for the 87 isolates and their microsatellite haplotypes were determined based on a combination of allelic data of the loci. In total, 40 haplotypes were observed. There were two predominant haplotypes: H16 and H25. H16 was observed in 9 isolates (10%) from 1996 to 2005, and H25 in 27 (31%) from 1995 to 2003. These results suggested that the recombination rate of *P. vivax* in South Korea, a temperate country, was lower than in tropical areas where identical haplotypes were rarely seen in the following year. Next, we estimated the relationships among the 40 haplotypes by eBURST analysis. Two major groups were found: one composed of 36 isolates (41%) including H25; the other of 20 isolates (23%) including H16. Despite the low recombination rate, other new haplotypes that are genetically distinct from the 2 groups have also been observed since 1997 (H27).

Conclusions/Significance: These results suggested a continual introduction of *P. vivax* from other population sources, probably North Korea. Molecular epidemiology using microsatellite DNA of the *P. vivax* population is effective for assessing the population structure and transmission dynamics of the parasites - information that can assist in the elimination of vivax malaria in endemic areas.

Citation: Iwagami M, Fukumoto M, Hwang S-Y, Kim S-H, Kho W-G, et al. (2012) Population Structure and Transmission Dynamics of *Plasmodium vivax* in the Republic of Korea Based on Microsatellite DNA Analysis. PLoS Negl Trop Dis 6(4): e1592. doi:10.1371/journal.pntd.0001592

Editor: Mehmet Ali Ozcel, Ege University, Turkey

Received October 12, 2011; **Accepted** February 17, 2012; **Published** April 3, 2012

Funding: This work was financially supported by a Grant-in-Aid for Scientific Research (B) (19406013) from the Ministry of Education, Culture, Sports, Science, and Technology of Japan (http://www.mext.go.jp/english/), the Grant for National Center for Global Health and Medicine (21A107), and a Grant for Global Health Promotion (H22-KUNII-SHITEI-004) from Ministry of Health, Labour and Welfare of Japan (http://www.mhlw.go.jp/english/). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: wgkho@inje.ac.kr (WGK); kano@ri.ncgm.go.jp (SK)

Introduction

Plasmodium vivax, the second most prevalent species of the human malaria parasite, is widely distributed around the world, especially in Asia and South America; it ranges from tropical to temperate areas [1,2]. In these countries, the proportion of *P. falciparum* cases is gradually decreasing due to the impact of global malaria control programs such as "The Roll Back Malaria Partnership" and "The Global Fund to Fight AIDS, Tuberculosis and Malaria" as well as local control programs. In contrast, the proportion of *P. vivax* cases is gradually increasing [1], and therefore deserves more attention than it has previously received [3].

Understanding the genetic characteristics of the malaria parasite population is important for monitoring the transmission pattern and evaluating the effectiveness of malaria control in endemic areas [4–7]. Recently, the population structure and transmission dynamics of *P. vivax* have been reported in some tropical and subtropical areas where the parasites are prevalent throughout the year or seasonally prevalent but not discontinuous during the year [8–13]. However, little is known about these characteristics in temperate areas where vivax malaria is only seasonally prevalent and discontinuous during the year.

In the Republic of Korea (South Korea), which is in the temperate zone of the continent of Asia, indigenous vivax malaria had been successfully eliminated by the late 1970s thanks to an effective program conducted by the National Malaria Eradication Service of the South Korean government with the support of the WHO [14–16], but has re-emerged since 1993 [17]. At the

Author Summary

Vivax malaria is widely prevalent, mainly in Asia and South America with 390 million reported cases in 2009. Worldwide, in the same year, 2.85 billion people were at risk. *Plasmodium vivax* is prevalent not only in tropical and subtropical areas but also in temperate areas where there are no mosquitoes in cold seasons. While most malaria researchers are focusing their studies on the parasite in tropical areas, we examined the characteristics of *P. vivax* in South Korea (temperate area) temporally, using 10 highly polymorphic microsatellite DNA (a short tandem repeat DNA sequence) in the parasite genome, and highlighted the differences between the tropical and temperate populations. We found that the South Korean *P. vivax* population had low genetic diversity and low recombination rates in comparison to tropical *P. vivax* populations that had been reported. We also found that some of the parasite clones in the population were changing from 1994 to 2008, evidence suggesting the continual introduction of the parasite from other populations, probably from North Korea. Polymorphic DNA markers of the *P. vivax* parasite are useful tools for estimating the situation of its transmission in endemic areas.

beginning of the re-emergence, the patients were only South Korean soldiers, veterans, and soldiers from the US military who were serving in the border area between North and South Korea in the western Demilitarized Zone (DMZ) [18–20]. Gradually, however, the number of infected civilians who lived in or near the area increased [18], suggesting local transmission of *P. vivax* between humans and *Anopheles* mosquitoes in the country. The number of vivax malaria cases increased steadily until 2000 (4,183 cases), then began to decrease gradually until 2004 (864 cases) (Fig. 1) [1,16]. In spite of continuous malaria control measures implemented by the South Korean government, the numbers of reported cases fluctuated between 1,000 and 2,000 cases per year from 2005 to 2009 [1]. The WHO reports that vivax malaria was more prevalent in the Democratic People's Republic of Korea (North Korea), where there were 296,540 cases in 2001 and 14,845 cases in 2009 [1,19].

We previously conducted genetic epidemiological surveys of the *P. vivax* population in South Korea using DNA sequences of some antigenic molecules of the parasite (circumsporozoite protein, Duffy binding protein, apical membrane antigen 1, merozoite surface protein-1) and found that there were 2 genotypes in the country's parasite population [21–25]. The advantage of using antigenic molecules of the parasites for genetic epidemiology is that they could be vaccine candidates; however such antigenic molecules are under strong selective pressure from the host immune system, so the variation in the molecules might be biased due to this [26]. In previous studies, the isolates that were used were collected from vivax malaria patients in a single year so temporal changes in the parasite population could not be examined. In the present study, we examined the population structure and the transmission dynamics of *P. vivax* in South Korea temporally using 10 highly polymorphic neutral DNA markers of the parasite collected from 1994 to 2008 and compared these characteristics with those reported in tropical and subtropical areas. Based on these data, we provide a possible explanation as to why it has not been possible to eliminate vivax malaria in South Korea in spite of a continuous governmental effort.

Methods

Materials

A total of 255 *P. vivax* samples isolated from South Korean soldiers or veterans who had served in the DMZ from 1994 to 2008 were used in this study. These patients were also diagnosed by microscopic examination of peripheral blood smears when they contracted malaria. The patient blood samples were collected and preserved at −30°C until use. This study was performed according to the Ethical Guidelines for Clinical Research issued by the Ministry of Health, Labour and Welfare of Japan on July 31, 2008, and the Ethical Guidelines for Epidemiological Research issued by the Ministries of Health, Labour and Welfare, and of Education, Science, Culture, and Sports of Japan on December 1, 2008. Because of the long-term prior collection of widely distributed samples, written or oral informed consent from the patients for the specific purpose of this study could not be obtained at each sample collection. However, no author of the study was involved in gathering patient samples and the individual information of the donors was disconnected from the authors. Thus, all the samples were anonymized, and indeed it is most unlikely that the results obtained from the analysis of the isolated parasites would result in a breach of donor privacy.

DNA extraction

Parasite DNA was extracted from frozen whole blood samples by phenol-chloroform extraction after proteinase K digestion [27] or by QIAamp DNA Mini Kit (Qiagen, Valencia, CA, USA).

Genotyping by polymerase chain reaction (PCR)

Ten microsatellite DNA loci were amplified by PCR. The loci were as follows: MS1 (chromosome 3), MS4 (chromosome 6), MS5 (chromosome 6), MS6 (chromosome 11), MS7 (chromosome 12), MS8 (chromosome 12), MS9 (chromosome 8), MS12 (chromosome 5), MS15 (chromosome 5) and MS20 (chromosome 10). The PCR primer sets and amplification conditions were consistent with the protocol of Karunaweera et al. [28]. Sizes of fluorescently-labeled PCR products were measured on an Applied Biosystems Prism Genetic Analyzer 3130*xl* using GeneMapper(R) version 4.1 with a 500 ROX size standard (Applied Biosystems, CA, USA).

Amplified different-sized PCR products using the same primer sets were considered to be individual alleles within a locus, as size variation among isolates is consistent with the repeat number in a microsatellite locus [5]. The electropherogram shows peak profiles for the microsatellite loci, based on the fluorescence intensity of the labeled PCR products in this analysis. Multiple alleles per locus were scored if minor peaks were taller than at least one-third the height of the predominant allele for each locus. Multiple-genotype infections (MGIs) were defined as those in which at least one of the 10 loci contained more than one allele [5].

Population genetic analyses

Population genetic analyses were performed based on allele frequencies of the 10 microsatellite loci of the population. The level of genetic diversity of the *P. vivax* population in South Korea was assessed by allele number per locus (A) and expected heterozygosity (H_E). H_E values for each locus were calculated using $H_E = [n/(n-1)]\ [1-\Sigma p_i^2]$, where n corresponds to the number of isolates examined and p_i is the frequency of the i^{th} allele. The statistical differences among those values were evaluated by Welch's t-test.

Multilocus linkage disequilibrium (LD) was assessed using the standardized index of association (I_A^S) [29,30]. This analysis was performed using the LIAN 3.5 Web interface [31]. I_A^S was

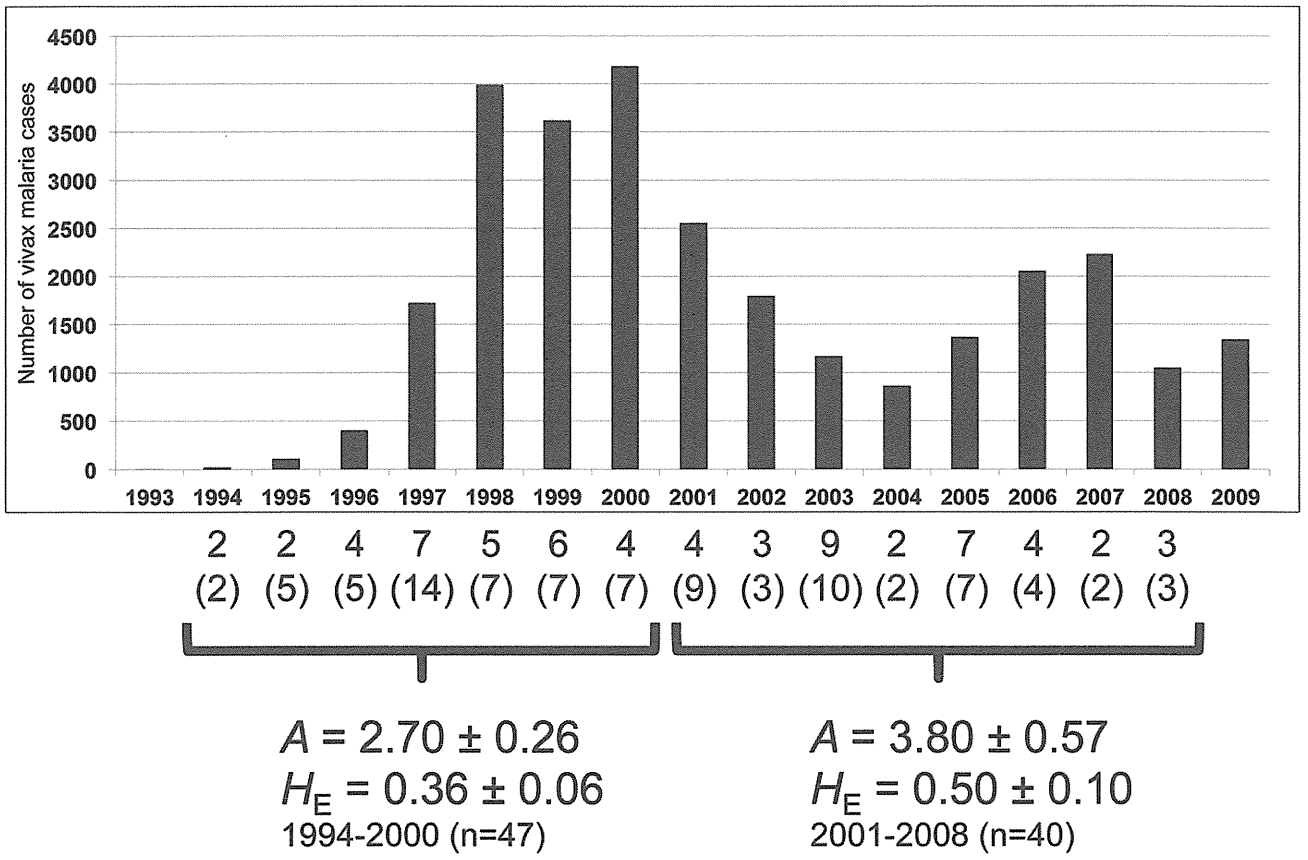

Figure 1. Genetic diversity of the *P. vivax* population in South Korea. *A*: Average number of the alleles ± SE, H_E: Average expected heterozygosity ± SE. Numbers and numbers in parentheses represent number of haplotypes and isolates observed each year, respectively. The graph was made based on numbers of reported vivax malaria cases in South Korea. n: represents the number of isolates. The data were obtained from the *World Malaria Report 2010* (WHO) [1].
doi:10.1371/journal.pntd.0001592.g001

calculated using the formula $I_A^S = (V_D/Ve-1)/(l-1)$ with permutation testing of the null hypothesis of complete linkage equilibrium ($I_A^S = 0$), where V_D is the observed mismatch variance, Ve is the expected mismatch variance, and l is the number of examined loci. Significances of the observed I_A^S values were calculated by Monte-Carlo simulation, using 10,000 random permutations of the data. This statistic is a variation of the method proposed by Maynard-Smith et al [29]. The results were standardized by the number of loci, to enable a comparison of different data sets [30]. This test was applied to the data sets from each population in two ways. First, the mixed-clone infections were excluded so that only the single-clone infections were analyzed, giving absolute confidence in the haplotype profile. Second, any multilocus genotype found in more than one isolate was only counted once in the analysis, i.e. unique haplotypes only, reducing the sample size slightly and thereby removing the possible effect of recent epidemic expansion of particular clones [5].

Microsatellite haplotypes of the isolates were determined based on a combination of the allelic data of the 10 loci. The relationships among the haplotypes were estimated by eBURST analysis [32].

Results

The allelic data of the 10 microsatellite loci were obtained from 87 of the 255 (34%) isolates that were used in the study. They were not available in the remaining 168 isolates (66%) due to failure in acquiring PCR products of some loci by PCR-based genotyping. Failure was possibly due to there being only a small amount of DNA for PCR amplification or to the DNA being of low quality after multiple times of freeze-thaw.

When different sizes of alleles were observed in one locus, we regarded this as multiple genotype infections (MGIs). MGIs were observed in some of the 10 microsatellite loci in 85 of the 87 isolates (97.7%). The frequencies of MGIs varied among the 10 loci (0.00 to 0.84; average: 0.29) (Table 1). We also examined the number of MGI loci per isolate. In the 87 isolates with 10 loci, the highest frequency of MGI loci per isolate was 2 (25 isolates) and the frequencies decreased gradually according to the increase in the number of MGI loci (Fig. 2). The highest number of MGI loci per isolate was 8 (one isolate). The major alleles in each locus were used for population genetic analysis.

Genetic diversity

In the 10 loci, the number of alleles (A) for each locus was 2 to 7 (average: 4.3). The expected heterozygosity (H_E) for each of these loci was 0.05 to 0.66 (average: 0.43) (Table 2).

Next, the *P. vivax* population was divided into 2 groups: one comprised of the 47 isolates collected from 1994 to 2000 when the numbers of vivax malaria cases increased; the other comprised of the 40 isolates collected from 2001 to 2008, when the numbers of cases decreased until 2004 and then increased slightly. The level of