Admixture in Humans of Two Divergent Plasmodium knowlesi Populations Associated with Different Macaque Host Species

Paul C. S. Divis1,2*, Balbir Singh2*, Fread Anderios3, Shamilah Hisam4, Asmad Matusop2,5, Clemens H. Kocken6, Samuel A. Assefa1, Craig W. Duffy1, David J. Conway1,2*

1 Pathogen Molecular Biology Department, London School of Hygiene and Tropical Medicine, London, United Kingdom, 2 Malaria Research Centre, Faculty of Medicine and Health Sciences, University Malaysia Sarawak, Kuching, Sarawak, Malaysia, 3 Sabah State Public Health Laboratory, Kota Kinabalu, Sabah, Malaysia, 4 Institute for Medical Research, Kuala Lumpur, Malaysia, 5 Sarawak State Health Department, Kuching, Sarawak, Malaysia, 6 Biomedical Primate Research Centre, Rijswijk, The Netherlands

* paul.divis@lshtm.ac.uk (PCS); bskhaira55@gmail.com (BS); david.conway@lshtm.ac.uk (DJC)

Abstract

Human malaria parasite species were originally acquired from other primate hosts and subsequently became endemic, then spread throughout large parts of the world. A major zoonosis is now occurring with Plasmodium knowlesi from macaques in Southeast Asia, with a recent acceleration in numbers of reported cases particularly in Malaysia. To investigate the parasite population genetics, we developed sensitive and species-specific microsatellite genotyping protocols and applied these to analysis of samples from 10 sites covering a range of >1,600 km within which most cases have occurred. Genotypic analyses of 599 P. knowlesi infections (552 in humans and 47 in wild macaques) at 10 highly polymorphic loci provide radical new insights on the emergence. Parasites from sympatric long-tailed macaques (Macaca fascicularis) and pig-tailed macaques (M. nemestrina) were very highly differentiated (FST = 0.22, and K-means clustering confirmed two host-associated subpopulations). Approximately two thirds of human P. knowlesi infections were of the long-tailed macaque type (Cluster 1), and one third were of the pig-tailed-macaque type (Cluster 2), with relative proportions varying across the different sites. Among the samples from humans, there was significant indication of genetic isolation by geographical distance overall and within Cluster 1 alone. Across the different sites, the level of multi-locus linkage disequilibrium correlated with the degree of local admixture of the two different clusters. The widespread occurrence of both types of P. knowlesi in humans enhances the potential for parasite adaptation in this zoonotic system.

Author Summary

Extraordinary phases of pathogen evolution may occur during an emerging zoonosis, potentially involving adaptation to human hosts, with changes in patterns of virulence and
transmission. In a large population genetic survey, we show that the malaria parasite *Plasmodium knowlesi* in humans is an admixture of two highly divergent parasite populations, each associated with different forest-dwelling macaque reservoir host species. Most of the transmission and sexual reproduction occurs separately in each of the two parasite populations. In addition to the reservoir host-associated parasite population structure, there was also significant genetic differentiation that correlated with geographical distance. Although both *P. knowlesi* types co-exist in the same areas, the divergence between them is similar to or greater than that seen between sub-species in other sexually reproducing eukaryotes. This may offer particular opportunities for evolution of virulence and host-specificity, not seen with other malaria parasites, so studies of ongoing adaptation and interventions to reduce transmission are urgent priorities.

**Introduction**

The epidemiological emergence of infections can be traced by genotypic analyses, with a high level of resolution when pathogens have a high mutation rate, as illustrated by recently emerged viruses that now have a massive impact on global public health [1,2]. Such analysis is more challenging for eukaryote pathogens with low mutation rate, although it is now clear that the major human malaria parasites *Plasmodium falciparum* and *P. vivax* have been endemic for many thousands of years after having been acquired as zoonotic infections from African apes [3,4]. In contrast, natural human infections by *P. knowlesi* were almost unknown [5] until a large focus of cases in Malaysian Borneo was described a decade ago [6]. Infections have since been reported from throughout southeast Asia, within the geographical range of the long-tailed and pig-tailed macaque reservoir hosts (*Macaca fascicularis* and *M. nemestrina*) and mosquito vectors (of the *Anopheles leucosphyrus* group) [7]. The most highly affected country is Malaysia, where there have been thousands of reported cases and *P. knowlesi* is now the leading cause of malaria in most areas [8,9].

It is vital to determine the causes of this apparent emergence, as *P. knowlesi* can cause severe clinical malaria with a potentially fatal outcome [10–12]. Increasing rates of case detection may reflect better diagnosis, increased transmission by mosquitoes from reservoir host macaques to humans, or parasite adaptation to humans. Molecular tools to discriminate *P. knowlesi* from other malaria parasite species were not widely applied until the zoonosis became known, but analysis of DNA in archived blood samples from Malaysia and Thailand shows that it was already widespread twenty years ago [13,14]. Sequences of parasite mitochondrial genomes and a few nuclear gene loci indicate ongoing zoonotic infection, as human *P. knowlesi* genotypes share most alleles identified in parasites sampled from wild macaques [15–17].

To understand this zoonosis, and to identify whether human-to-human mosquito transmission is occurring, analyses of parasite population genetic structure in humans and macaques should be performed by extensive population sampling and characterisation of multiple putatively neutral loci. This study presents a *P. knowlesi* microsatellite genotyping toolkit and its application to the analysis of a large sample of isolates from human cases at ten different sites, as well as from both species of wild macaque reservoir hosts. Results reveal a profound host-associated sympatric subdivision within this parasite species, as well as geographical differentiation indicating genetic isolation by distance. The existence of two divergent parasite subpopulations, and their admixture in human infections provides unparalleled opportunity for parasite hybridisation and adaptation. Observations of some clinical infections with parasite types that