HHS Public Access

Author manuscript

*Cell Host Microbe*. Author manuscript; available in PMC 2018 February 21.

Published in final edited form as:

*Cell Host Microbe*. 2017 March 08; 21(3): 310–312. doi:10.1016/j.chom.2017.02.019.

# A Feast of Malaria Parasite Genomes

**Jane M. Carlton**[1,*] and **Steven A. Sullivan**[1]

[1]Center for Genomics and Systems Biology, Department of Biology, New York University, New York, NY 10003, USA

## Abstract

The *Plasmodium* genus has evolved over time and across hosts, complexifying our understanding of malaria. In a recent *Nature* paper, Rutledge et al. (2017) describe the genome sequences of three major human malaria parasite species, providing insight into *Plasmodium* evolution and raising the question of how many species there are.

A budding malariologist in the 1990s was steeped in the dogma that there were four species of malaria parasite responsible for millions of cases of malaria in tropical and sub-tropical regions of the world. *Plasmodium falciparum* was the deadliest, causing two to three million deaths each year due to the development of cerebral malaria in children under the age of 5 in sub-Saharan countries. *Plasmodium vivax*, on the other hand, was the comparatively "benign" species found in South American and Central and Southeast Asian countries, little studied because of the greater threat posed by *P. falciparum* and the lack of an in vitro culture system. *Plasmodium malariae* and *Plasmodium ovale* were species on the periphery, perceived as causing so few infections as to be almost negligible in impact. A clade of monkey malaria species, primarily studied because of their potential as in vivo models for *P. falciparum* and *P. vivax*, were also known to infect Old World primates in Southeast Asia and included *Plasmodium knowlesi*, *Plasmodium cynomolgi*, and *Plasmodium inui*.

Fast forward 20 years and how things have changed! The number of global human malaria infections has dramatically declined, and there now exists a greater understanding of the importance and threat that *P. vivax* presents and acknowledgment that this species requires its own research and elimination agenda (WHO, 2015). Compelling evidence also exists that *P. ovale* is in fact two species, *P. ovale curtisi* and *P. ovale wallikeri* (Sutherland et al., 2010), morphologically similar and inhabiting sympatric ranges but genetically distinct. And most astonishing of all has been an evaluation of the zoonotic (host-switching) potential of *Plasmodium* species, due to pioneering work by Balbir Singh and colleagues identifying *P. knowlesi* as a species causing a heavy burden of clinical infections in Malaysia (Singh et al., 2004) and several groups providing evidence for the existence of additional *Plasmodium* species naturally infecting wild-living chimpanzees and gorillas (reviewed in Loy et al., 2017). Indeed, among the latter findings include evidence that *P. vivax* found in African apes has the potential to infect humans, suggesting that human and ape *P. vivax* parasites

*Correspondence: jane.carlton@nyu.edu.

represent a single species. *P. malariae*-like and *P. ovale*-like sequences too have been found in African great apes, although their zoonotic potential remains unclear.

While reference genome sequences of *P. falciparum*, *P. vivax*, and *P. knowlesi* have been published, the community has been waiting anxiously for high-quality genomes of *P. malariae* and the two species of *P. ovale*. Excitingly, the paper by Rutledge and colleagues in the January 2017 issue of *Nature* (Rutledge et al., 2017) does not disappoint and presents reference/high-quality drafts, as well as shotgun sequences, for five *P. malariae*, two *P. malariae*-like, three *P. ovale curtisi*, and two *P. ovale wallikeri* isolates (Table 1). For several of these, a means to identify pre-existing sequence data ("co-infection mining") involved screening the sequencing reads of ~2,500 samples from the MalariaGen Pf3K global collection generated at the same genome-sequencing center, the Wellcome Trust Sanger Institute, where many of the authors are based. Fortuitously (although not for the patients), four samples of *P. malariae* were obtained from travelers returning to Australia from Uganda, Papua Indonesia, Malaysia, and Guinea, and one of them contained sufficiently high parasite DNA with low human host contamination for it to be used to generate a high-quality, manually curated *P. malariae* reference genome.

What are the major findings of this work? First, the *P. malariae* genome appears to be indistinguishable from sequences of a species, *Plasmodium brasilianum*, found in New World primates, confirming a recent report describing them as a single anthrozoonotic species circulating freely between monkeys and humans in the Venezuelan Amazon (Lalremruata et al., 2015). Second, using 1,000 manually curated orthlogs in 12 *Plasmodium* species, Rutledge et al. (2017) constructed maximum likelihood trees that appear to show rodent malaria parasites as an outgroup to *P. ovale*. This intriguing finding suggests that there was an ancestral host switch from primates to the ancestor of African thicket rats that host the rodent malaria species. If this finding holds true, it may well be that rodent malaria parasites could be a good model system for the study of *P. ovale* compared to other human-infective species.

Finally, one goal of "malariomics" is to relate the preference of different *Plasmodium* species for specific hosts and developmental stages of red blood cells (RBCs) to genomic differences. *P. ovale* and *P. vivax*, for instance, infect immature RBCs of humans, while *P. malariae* prefers mature RBCs in humans and monkeys, and *P. knowlesi* infects both RBC stages and both hosts. Understanding the genomic bases of these distinctions would doubtless shed light on the mechanisms of zoonoses as well. As *Plasmodium* genomes have become available, it has become clear that what mainly distinguishes them are their multigene families. Such families can comprise a few or hundreds of genes, typically encoding proteins expressed on the parasite surface or exported to the cytoplasm and surface of infected RBCs, with putative functions that include cell adhesion and invasion, and evasion of the host's immune system. A species or strain's unique genomic repertoire of such proteins might determine its host and host cell range. But multigene families tend to cluster in the repetitive, difficult-to-assemble subtelomeric regions of chromosomes, making it hard to tally gene numbers accurately. Rutledge et al. (2017)'s use of long reads using the Pacific Biosciences next-generation sequencing platform, in combination with extensive manual curation, allows them to overcome that common obstacle and upgrade earlier

analyses that were based on much lower-quality assemblies (Ansari et al., 2016). They quantitate large differences between *P. ovale* and *P. malariae* in gene complements of the *pir*, *STP1*, and *surfin* families of surface/exported proteins and, most notably, discover two large, previously undistinguished gene families in *P. malariae* that they relate to a known *P. falciparum* RBC invasion protein using 3D protein modeling. They also identify, from surveying the *rbp* family of RBC adhesion proteins across *Plasmodium* species, one member that may be required specifically for invading mature RBCs—an intriguing finding given that the same protein (RBP3, also called NBPXa) in *P. knowlesi* has been shown to be necessary for invading human RBCs (Moon et al., 2016).

While the study by Rutledge et al. (2017) provides missing pieces in the puzzle of the complex evolution and phylogeography of the *Plasmodium* genus, it also illustrates the benefits and drawbacks of sequencing laboratory-adapted strains versus extant clinical or field isolates. Whereas previous human malaria reference strains were chosen for sequencing because of their availability to researchers as pure frozen stocks from the Malaria Research and Reference Reagent Resource (https://www.beiresources.org/MR4Home.aspx) and the body of research describing them, this came at a cost. For example, many were isolated decades ago and may not be representative of today's genetic diversity, and others were adapted to growth in a novel system, such as in vitro culture, a procedure now known to induce loss-of-function mutants that are not normally seen in natural isolates (Claessens et al., 2017). The alternative—to generate high-quality genome sequences from clinical isolates that have not been previously studied and for which there may be limited biological material for others to use—also has constraints. It's wise to remember, then, that there really is no such thing as *the* reference genome, a one-size-fits-all, for a *Plasmodium* species, especially in the light of extensive genetic diversity and differences in gene copy number (Hupalo et al., 2016) that can exist even within a single species of this fascinating lineage.

## References


Ansari HR, Templeton TJ, Subudhi AK, Ramaprasad A, Tang J, Lu F, Naeem R, Hashish Y, Oguike MC, Benavente ED, et al. Int J Parasitol. 2016; 46:685–696. [PubMed: 27392654]

Claessens A, Affara M, Assefa SA, Kwiatkowski DP, Conway DJ. Sci Rep. 2017; 7:41303. [PubMed: 28117431]

Hupalo DN, Luo Z, Melnikov A, Sutton PL, Rogov P, Escalante A, Vallejo AF, Herrera S, Arévalo-Herrera M, Fan Q, et al. Nat Genet. 2016; 48:953–958. [PubMed: 27348298]

Lalremruata A, Magris M, Vivas-Martínez S, Koehler M, Esen M, Kempaiah P, Jeyaraj S, Perkins DJ, Mordmüller B, Metzger WG. EBioMedicine. 2015; 2:1186–1192. [PubMed: 26501116]

Loy DE, Liu W, Li Y, Learn GH, Plenderleith LJ, Sundararaman SA, Sharp PM, Hahn BH. Int J Parasitol. 2017; 47:87–97. [PubMed: 27381764]

Moon RW, Sharaf H, Hastings CH, Ho YS, Nair MB, Rchiad Z, Knuepfer E, Ramaprasad A, Mohring F, Amir A, et al. Proc Natl Acad Sci USA. 2016; 113:7231–7236. [PubMed: 27303038]

Rutledge GG, Böhme U, Sanders M, Reid AJ, Cotton JA, Maiga-Ascofare O, Djimdé AA, Apinjoh TO, Amenga-Etego L, Manske M, et al. Nature. 2017; 542:101–104. [PubMed: 28117441]

Singh B, Kim Sung L, Matusop A, Radhakrishnan A, Shamsul SS, Cox-Singh J, Thomas A, Conway DJ. Lancet. 2004; 363:1017–1024. [PubMed: 15051281]

Sutherland CJ, Tanomsing N, Nolder D, Oguike M, Jennison C, Pukrittayakamee S, Dolecek C, Hien TT, do Rosàrio VE, Arez AP, et al. J Infect Dis. 2010; 201:1544–1550. [PubMed: 20380562]

WHO. Control and Elimination of Plasmodium vivax: A Technical Brief. World Health Organization; 2015.

**Table 1**

The Current Satus of References, Draft Assemblies, and Shotgun Sequences for Six Human-Infective Species of Malaria Parasite

| | **Pf** | **Pv** | **Pk** | **Pm** | **Poc** | **Pow** |
|---|---|---|---|---|---|---|
| No. genomes sequenced | 1 reference *, 2 draft assemblies *, 1,000 s shotgun sequences | 1 reference *, 5 draft assemblies, 100 s shotgun sequences | 1 reference *, ~50 shotgun sequences | 1 reference, 7 draft assemblies */shotgun sequences | 1 reference, 3 draft assemblies */shotgun sequences | 1 reference, 2 draft/ shotgun sequences |
| Geographical location of sequenced genomes | Global | Global | Malaysia | Uganda, Papua Indonesia, Mali, Malaysia, Guinea | Ghana, Cameroon | Ghana, Cameroon |
| Zoonotic | No | Possibly | Yes | Yes | Not known | Not known |
| Notable multi-gene families | pir, var, tra, ETRAMP, PHIST, EBP | pir, STP1, tra, ETRAMP, PHIST, RBP, DBP | pir, SICAvar, tra, ETRAMP, PHIST, RBP, DBP | pir, STP1, tra, ETRAMP, PHIST, RBP, DBP, fam-l, fam-m | pir, STP1, tra, ETRAMP, PHIST, RBP, DBP | pir, STP1, tra, ETRAMP, PHIST, RBP, DBP |

Pf, *P. falciparum*; Pv, *P. vivax*; Pk, *P. knowlesi*; Pm, *P. malariae*; Poc, *P. ovale curtisi*; Pow, *P. ovale wallikeri*. Data compiled from Ansari et al. (2016) and Rutledge et al. (2017).

*indicates one or more assemblies generated from culture-adapted or chimpanzee/monkey-adapted strains.

*Cell Host Microbe*. Author manuscript; available in PMC 2018 February 21.