OXFORD

Briefings in Functional Genomics, 17(6), 2018, 451–457

doi: 10.1093/bfgp/ely013
Advance Access Publication Date: 25 April 2018
Review paper

Host–parasite interactions and ecology of the malaria parasite—a bioinformatics approach

Dariusz Izak, Joanna Klim and Szymon Kaczanowski

Corresponding author: Szymon Kaczanowski, Department of Bioinformatics, Institute of Biochemistry and Biophysics of the Polish Academy of Sciences, Pawinskiego 5a, 02-106 Warsaw, Poland. E-mail: szymon@ibb.waw.pl

Abstract

Malaria remains one of the highest mortality infectious diseases. Malaria is caused by parasites from the genus *Plasmodium*. Most deaths are caused by infections involving *Plasmodium falciparum*, which has a complex life cycle. Malaria parasites are extremely well adapted for interactions with their host and their host's immune system and are able to suppress the human immune system, erase immunological memory and rapidly alter exposed antigens. Owing to this rapid evolution, parasites develop drug resistance and express novel forms of antigenic proteins that are not recognized by the host immune system. There is an emerging need for novel interventions, including novel drugs and vaccines. Designing novel therapies requires knowledge about host–parasite interactions, which is still limited. However, significant progress has recently been achieved in this field through the application of bioinformatics analysis of parasite genome sequences. In this review, we describe the main achievements in 'malarial' bioinformatics and provide examples of successful applications of protein sequence analysis. These examples include the prediction of protein functions based on homology and the prediction of protein surface localization via domain and motif analysis. Additionally, we describe PlasmoDB, a database that stores accumulated experimental data. This tool allows data mining of the stored information and will play an important role in the development of malaria science. Finally, we illustrate the application of bioinformatics in the development of population genetics research on malaria parasites, an approach referred to as reverse ecology.

Key words: malaria; genomics; bioinformatics; sequence analysis; population genetics

Introduction

Malaria remains one of the most important global health problems [1].

In malaria-endemic countries in Africa, a large proportion of child deaths is directly or indirectly attributable to infection with *Plasmodium falciparum*, which is estimated to be responsible for the deaths of at least half a million children annually. However, classical epidemiological analysis of the impact of malaria on child mortality has suggested that this number is significantly higher. One such study was performed on Bioko Island (Equatorial Guinea). Multiple successful malaria interventions were implemented on this island in 2004 [2]. During that time, the mortality of children <5 years of age fell from 152 per 1000 births to 55 per 1000, indicating that the mortality caused by the disease was much higher than previously expected in areas where malaria was not well controlled. Malaria is caused by parasites from genus *Plasmodium*. The majority of deaths are caused by infections involving *P. falciparum*, which exhibits a complex life cycle [3] (see Figure 1). The life cycle

Dariusz Izak is a PhD student in the Department of Bioinformatics at the Institute of Biochemistry and Biophysics of the Polish Academy of Sciences. He received his MSc (2013) in Microbiology from the Autonomous Department of Microorganisms' Biology, Warsaw University of Life Sciences. Currently, his main research interests are data mining associated with evolution and genetic interactions in yeast.

Joanna Klim is a PhD student in the Department of Microbial Chemistry at the Institute of Biochemistry and Biophysics of the Polish Academy of Sciences. She received her MSc (2015) in general microbiology from the Faculty of Biology, University of Warsaw. Currently, her main research interests are evolutionary processes in yeast.

Szymon Kaczanowski is an evolutionary biologist in the Department of Bioinformatics at the Institute of Biochemistry and Biophysics of the Polish Academy of Sciences. He was a visiting scholar at Harvard University in the Hartl Lab in 2005. He has conducted a number of bioinformatics studies concerning the biology of malaria.

involves two hosts (human and mosquito) and two types of ploidy. For most of the human host stage, the *Plasmodium* cells are haploidic and asexual. Sexual reproduction occurs in the mosquito. The symptoms of the disease are caused by part of the haploidic cycle called erythrocytic schizogony. During the erythrocytic schizogony cycle, parasite cells multiply inside erythrocytes. As indicated in Figure 1, it is also the parasite development stage in which diagnosis is possible. Malarial anaemia is not induced simply by red blood cell rupture; instead, this disease is multifactorial and disturbs erythropoiesis [4].

The parasites are haploid cells for most of their life cycle (a single asexual parasite is capable of giving rise to all life cycle stages, including both male and female gametes). The parasites replicate asexually in the circulating blood of the host and must undergo a round of sexual reproduction in the mosquito vector to be transmitted to a new host. A small percentage of the asexual cells in the blood differentiate into sexual cells, which are referred to as gametocytes. Male and female gametocytes are the precursor cells for male and female gametes, which form immediately after the gametocytes are taken up by a mosquito during a blood meal. Within 10–12 min after activation, each male gametocyte produces up to eight haploid gametes, and each female gametocyte differentiates into a single haploid gamete.

Rapid evolution of parasites is the source of the main difficulties in malaria control. Parasites have developed resistance against the drugs chloroquine [5] and, recently, artemisinin [6]. A recent study showed a real danger that drug treatment may promote the selection of more virulent parasites [7]. Additionally, mosquitoes have developed resistance to insecticides [8]. Therefore, the need for a malaria vaccine is increasing. Modern malaria vaccine development stems from studies in the 1960s involving immunization of mice with irradiated sporozoites [9]. The early vaccination experiments were promising [10]. Both human protection and transmission-blocking vaccines were tested. Transmission-blocking vaccines induce immunity against the stages of the parasites that infect mosquitoes [11] and therefore protect mosquitoes from infection via the blood of infected people. Based on these approaches, an effective malaria vaccine has seemed to be just around the corner for >30 years. However, a successful vaccine has never emerged from these investigations.

Currently, whole-parasite vaccines are being tested. The US biotech company Sanaria I is working on the development of a pre-erythrocytic vaccine composed of attenuated sporozoites [12]. Trap antigen [13, 14] and circumsporozoite (CS) [14–16] proteins are also candidates for vaccines that would induce protective immunity via antibodies against sporozoites. Current vaccine candidates for transmission-blocking vaccines involve targeting proteins that are essential for male fertility [11, 17–19], zygote development [20] and female development [21]. Progress in the development of blood-stage vaccines has been relatively slower [22]. The most promising current vaccine candidate is RTS,S/AS01 [23], which is a human host protection vaccine (immunizing against the *Plasmodium* stage present in the human host). RTS is a hybrid polypeptide consisting of a portion of the CS protein, which is a sporozoite surface antigen of *P. falciparum* malaria parasite strain NF54, fused to the amino-terminal end of the hepatitis B virus S protein. The S protein is the surface antigen of hepatitis B virus and is the antigen used in GSK Biological's licenced hepatitis B vaccines.

The development of an effective malaria vaccine has appeared imminent for >30 years. Unfortunately, even the most promising current vaccine candidate (RTS,S/AS01) has been found to provide only limited (approximately 30%) protection in the key trial age group of infants [23]. The use of this type of vaccine has the potential to make the situation even worse because repeated experimental evidence has shown that vaccines that reduce pathogen replication may select for more virulent pathogens, thereby eroding the benefits of vaccination [24]. Thus, understanding how such adaptations are selected in the *Plasmodium* population has critical significance for the design of new effective therapies. Parasites appear to be extremely well adapted to host immune pressure. Additionally, recognition of vaccine candidates by the host immune system is limited [15]. The best-described adaptation of the parasite for immune evasion is referred to as immune escape (i.e. selection of mutations in antigens that are not recognized by the host immune system) [15, 25, 26]. There is limited immunological recognition of critical malaria vaccine antigens by cytotoxic T lymphocytes [15]. This observation indicates that immune escape is particularly important in the case of the epitopes recognized by T cells. Another immune evasion mechanism is host immune suppression. Parasites have been shown to modulate the maturation of dendritic cells and as a result suppress T-cell activity [27, 28]. *Plasmodium* infection has also been demonstrated to induce regulatory T cells to suppress experimental autoimmune encephalomyelitis [29]. However, the plasmodial factors involved in host immune suppression have not been well described.

In conclusion, malaria remains one of the most important global health problems. Malaria parasites are extremely well adapted for interactions with their host and their host's immune system.

There is an emergent need for novel interventions, including novel drugs and vaccines. Designing novel therapies requires knowledge about host–parasite interactions, which is still limited. Studying this problem and performing the relevant experiments are difficult because these experiments are time-consuming and expensive.

Bioinformatics provides a cheap and affordable alternative to experimental approaches and leads to testable predictions. In this review, we present the main achievements of malarial bioinformatics to date.

First, we illustrate how protein function can be predicted using orthology.

Next, we show that domain analysis may suggest which proteins are involved in interactions with the host. Then, we briefly describe tools applied for the prediction of protein surface localization. These proteins are expected to be involved in interactions with the host and the host immune system. Subsequently, we describe PlasmoDB, which is a database where accumulated experimental data are stored. This tool allows data mining using the stored information.

Finally, we describe the main achievements in the field of reverse ecology of malaria, such as the prediction of ecological events using sequence data and population genetics algorithms, including our own results indicating that host immune pressure accelerates the speed of protein evolution and that host immune pressure differs during different stages. For summary of these methods, see Table 1.

Prediction of the mechanisms of host–parasite interactions—tools for homology searches

The first sequenced malaria genome was the human malaria parasite *P. falciparum* [30]. This genome has been analysed by different groups via homology searches with standard Blast searches. The authors of these studies discovered that

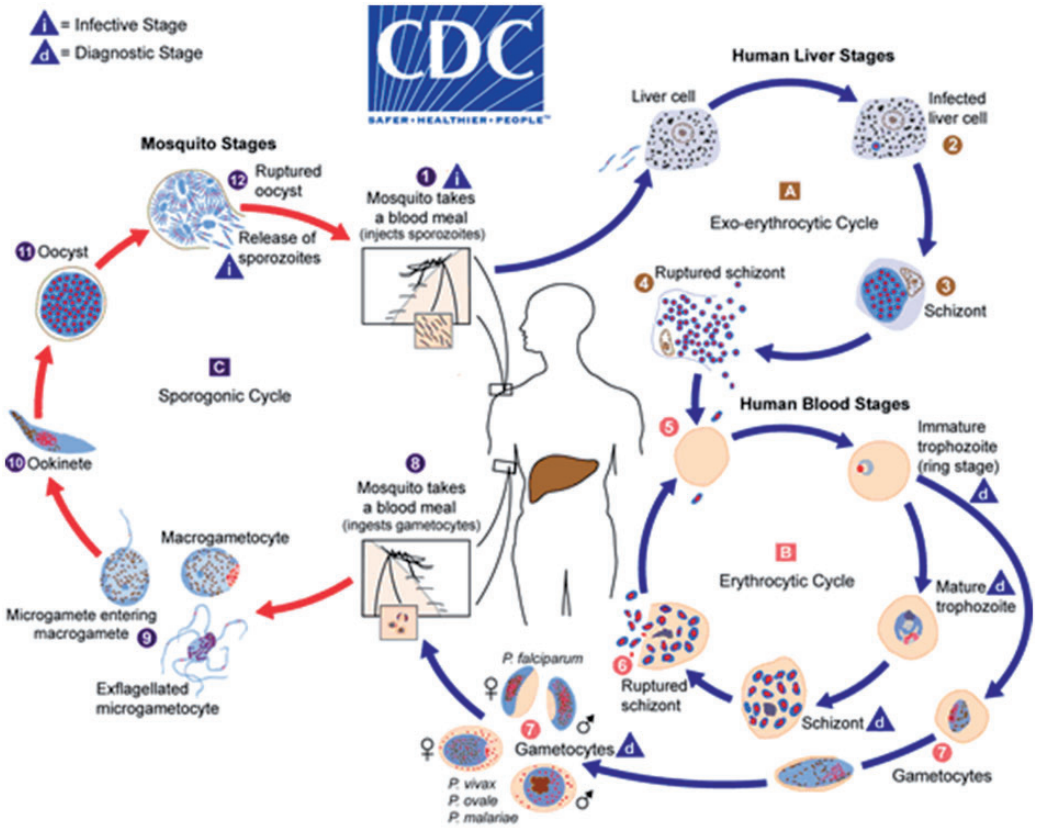

Figure 1: Plasmodium life cycle. Detailed description in text.
Source: Centers for Disease Control and Prevention's Public Health Image Library, https://phil.cdc.gov/phil/home.asp.

Table 1: Short summary of the methods, databases, associated publications and easily applicable tools mentioned in this study

	Homology searches	Domain analysis	Localization prediction	Reverse ecology	GWASs
Studies mentioned in this review	31–55	56–62	51, 52, 64–67	15, 16, 18, 23, 26, 27, 60, 69	4, 70–73
Useful databases	PlasmoDB (http://plasmodb.org), KEGG (http://www.genome.jp/kegg/), MPMP (http://mpmp.huji.ac.il/)	PlasmoDB (http://plasmodb.org), Pfam (http://pfam.xfam.org/)	PlasmoDB (http://plasmodb.org)	PlasmoDB (http://plasmodb.org)	PlasmoDB (http://plasmodb.org)
Ready-to-use tools	BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi)	PFAM server (http://pfam.xfam.org/search#tabview=tab1)	THMM2 (http://www.cbs.dtu.dk/services/TMHMM/), SignalP (http://www.cbs.dtu.dk/services/SignalP/)		

approximately 40% of the genes in this genome had obvious orthologues in other systematic groups (e.g. yeasts), justifying the prediction that the genes exhibited similar functions. Knowledge about these genes has been used to make such predictions. This approach leads to testable hypotheses describing putative basic mechanisms of host–parasite interactions. For example, this approach has been applied in studies on the expression of the antigenic erythrocyte membrane protein 1 (PfEMP1) family proteins, which are encoded by the *var* genes. These proteins are expressed on the surfaces of parasite-infected erythrocytes and are responsible for cytoadherence to receptors in the microvasculature, brain and placenta. The PfEMP1 proteins are important targets of the human immune system. Parasites avoid the induction of an immune response to PfEMP1 by switching between the expression of different members of the ~60-member *var* gene family encoding these proteins [31]. Only one *var* gene is expressed in any single parasite, whereas the other members of this family are essentially silenced [32]. Thus, the host cannot develop an

efficient response because the parasites escape by switching between PfEMP1 antigenic types [33, 34]. The observation of an absence of the programmed DNA rearrangements involved in regulating the expression of *var* genes suggested that epigenetic mechanisms of histone modification were involved in this process [32]. These mechanisms have been described in yeast [35]. Homology searches have been used for the prediction of putative factors involved in epigenetic mechanisms of histone modification in *Plasmodium*. Knocking out two orthologues of the yeast sir2 (silent information regulator) histone deacetylase was found to derepress a subset of *var* genes, and the two orthologues were shown to have complementary effects [36, 37]. Hence, bioinformatics analysis revealed that the basic mechanism underlying *var* gene expression involved epigenetic modification of histones. Although this mechanism is not well understood, further studies have revealed many important details (see as a review [38]).

Bioinformatics has also been applied in studies examining immunosuppression of the host immune system by malaria parasites. As noted in the 'Introduction' section, the plasmodial factors involved in host immune suppression are not well known. However, homology searches have indicated that two putative proteins are involved in this process [putative macrophage migration inhibitory factor (MIF) and T-cell immunomodulatory protein of *Plasmodium* (TIP)]. Although knockout experiments involving MIF were inconclusive and suggested that MIF was predominantly involved in efficient liver stage development [39], recent data indicated that this protein was involved in T-cell suppression [40]. We identified a putative homologue of TIP in the malaria genome [41]. Murine and human TIPs suppress host-versus-graft disease [42].

We also performed homology searches to predict the cell death machinery of unicellular parasites, including *Plasmodium*. Mitochondrial apoptosis (also known as apoptosis-like cell death in older articles) has been described in various unicellular organisms, including *Plasmodium* [43–45]. Apoptosis can be easily distinguished from other types of cell death because it is initiated by the release of apoptotic factors from the mitochondria. Studies have suggested that the apoptosis of parasitic protists is likely to constitute 'altruistic suicide' for the good of the entire population [44–48]. Experimental studies have shown that caspases, such as the protease metacaspase, induce apoptosis in *Plasmodium* [49].

Through homology searches, we identified additional elements of the apoptosis machinery. We showed that the *Plasmodium* genome encoded homologues of the key plant apoptotic nuclease ZEN1 but not the animal apoptotic nucleases EndoG and NUC1. Additionally, we demonstrated that apoptotic induction factors were encoded by the *Plasmodium* genome. In conclusion, we suggest that the apoptotic machinery of *Plasmodium* includes AIF, ZEN1 nucleases and metacaspases [48].

Homology searches can be impeded by low-complexity regions composed of a single amino acid residue. Asparagine-rich regions can constitute up to 30% of the *P. falciparum* proteome [50, 51]. This problem is usually solved by using algorithms that mask these regions and thus remove the bias introduced by repeats [52, 53].

Prediction of host–parasite interactions—tools for domain analysis

Proteins involved in host–parasite interactions often contain 'adhesive' domains. These domains can be detected using the PFAM server (http://pfam.xfam.org/) [54]. This analysis can predict which proteins are involved in interactions with the host simply through domain analysis, even if the proteins exhibit no homologues with known functions.

The existence of 'adhesive' domains was recognized before genome sequencing was conducted in the case of epidermal growth factor (EGF). Kaslow and co-workers cloned the transmission-blocking vaccine candidate Pfs25. Antibodies against this protein protected mosquitoes from infection by human blood. The Pfs25 protein contains an EGF domain [55]. Subsequently, a number of different surface proteins were shown to interact with hosts containing this domain, including various merozoite surface proteins, such as MSP1 [56], MSP4 [57] and MSP5 [58]. When the genome was subsequently sequenced, additional proteins containing EGF were described based on domain analysis and experimentally characterized, such as MSP8 [59] and MSP10 [60].

Other examples of adhesive domains include limulus clotting factor C (LCCL) and the ricin B lectin-related, carbohydrate-binding ricin domain. These domains are present in a number of parasitic proteins involved in interactions with the host [61].

Prediction of host–parasite interactions—tools for localization predictions

The prediction of protein subcellular localization and antigenicity is critical for any targeted drug, including malaria vaccines. In this context, surface and exported proteins are clearly of particular importance, and bioinformatics tools are useful.

The programs THMM2 [62] and SignalP [63] are widely used to predict the localization of proteins in different organisms. These programs have also been shown to be useful for malaria science.

Surface proteins usually contain transmembrane segments, which are detected with a high degree of accuracy by the THMM2 program [62]. Because transmembrane proteins generally contain many of these segments, these predictions are highly valuable. Secreted proteins usually contain signal peptide to target the proteins, which are slightly different in eukaryotes and prokaryotes. These peptides are generally removed during secretion. The SignalP software predicts the presence of these peptides with high confidence [63].

Methods for predicting the localization of *Plasmodium* proteins have also been developed based on 'addressing' motifs described only in malaria parasites. The HT [64] and PEXEL [65] motifs play a central role in the export of parasitic proteins to host erythrocytes. The ExportPred algorithm predicts proteins that are exported by parasites to the infected erythrocytes [66] based on signal peptides and the presence of PEXEL motifs.

PlasmoDB—an experimental data source

The progress of parasitology has led to the accumulation of experimental data from malaria parasites. These data may be analysed using bioinformatics approaches. Currently, most of the large-scale experimental data are available in the PlasmoDB database (http://plasmodb.org/plasmo/) [67]. PlasmoDB is a free online genome database for genus *Plasmodium*. To date, this database comprises almost 113 000 gene records from malaria parasites. This database belongs to the EuPathDB resource linking databases for different eukaryotic pathogens, mainly including pathogens considered emerging or re-emerging.

PlasmoDB has developed into a data bank containing various forms of data, ranging from basic information about *Plasmodium* genes, such as their nomenclature, chromosomal localization, gene products and features, to more complex topics, such as

genetic variation, epigenetic modifications, transcriptomics, proteomics and metabolomics, and even data from studies on evolutionary processes ranging from genetic to environmental scales and phylogenomics aspects.

In addition to providing all known relevant genomic information related to *Plasmodium* species, PlasmoDB provides an elaborate search system that allows users to perform complex queries concerning the available data.

The current version of the database (PlasmoDB 31 released 8 March 2017) includes 20 fully sequenced and annotated genomes belonging to 15 species of the *Plasmodium* genus, including 5 species infecting humans (*P. falciparum*, *Plasmodium knowlesi*, *Plasmodium malariae*, *Plasmodium vivax* and *Plasmodium ovale*).

PlasmoDB is still expanding the content and type of data it contains as well as its functionality. PlasmoDB is undoubtedly the most comprehensive publicly accessible *Plasmodium* database developed to date.

In the next section, we will illustrate how we have used these data in our studies of malaria parasites in combination with population genetics.

Reverse ecology of malaria parasites

The 'reverse ecology' term was introduced to science by Matthew Rockman [68]. This term describes an approach in which ecology is studied using bioinformatics and genomics.

The main source of knowledge under this approach is the comparison of genomes of different individuals of a given species and closely related species. This approach can be used to detect rapidly and slowly evolving genes.

Different species of malaria parasites (e.g. the rodent malaria parasite *Plasmodium berghei*) and different isolates of *P. falciparum* have been sequenced (e.g. the Asian DD2 and African Ghana isolates).

In our studies, we have used these data to analyse the evolution of *P. falciparum*. Using these data, we found that rapid evolution of *Plasmodium* genes resulted from antigenic escape (i.e. selection of mutations in antigens not recognized by the host immune system) [15]. To this end, sets of proteins harbouring experimentally evident epitopes were obtained from PlasmoDB, and the evolution of these proteins was compared with the evolution of proteins without epitopes. More protein polymorphisms were found in putative antigenic proteins [25, 26]. This finding indicated that the host immune response led to the selection of mutations in antigens not recognized by antigens (i.e. diversifying selection in natural populations).

We have also examined proteomes expressed in different stages of the parasitic life cycle, such as the proteomes of male and female gametocytes described by our co-workers [25, 26]. Additionally, different proteomes are available in the PlasmoDB database. Using these data, we showed that the 'male' proteome contained more antigens than the female proteome, probably because of stronger host immune pressure. Proteins expressed in male gametocytes evolve more rapidly than proteins expressed in female gametocytes. Subsequently, we compared proteomes from other stages of the life cycle and found that the greatest fraction of putative antigens was present in the sporozoite stage. In this stage, we observed more rapid accumulation of protein mutations in natural populations [25]. In conclusion, our results indicate that host immune pressure has a particularly strong impact on parasite evolution.

The application of reverse ecology is not restricted to analysis of the evolution of whole genes. For example, we identified rapidly evolving fragments of the vaccine candidates Pf47 and p230 [17].

Genome-wide association studies

Genome-wide association study (GWAS) is relatively recent field of study. The first article in which the authors applied this method was published in 2005 [69]. GWAS are high-throughput genotyping assays in which variants and/or polymorphisms are compared in case-control groups [70]. Comparing large number of individuals allows phenotypes to be linked with genomic mutations. GWAS can be applied to *Plasmodium* or human populations. The former can be used to identify drug resistance determinants [71], and the latter can be applied to discover the underlying molecular mechanism of infections and/or immunological responses. For instance, the connections of Th1 and Th2 cytokines with the development of severe malarial anaemia and erythropoiesis perturbation were demonstrated [4]. Another recent study discovered the mechanism underlying the increased immunity to malaria found in the Eastern African population, which was based on copy number variations of glycophorin genes of the MNS blood group system. The products of these genes are receptors for *Plasmodium* ligands during the erythrocytic schizogony cycle [72].

Metabolic databases

The metabolic network for the malaria parasite has been described using a number of different databases. Some of these databases are accessible from PlasmoDB and facilitate analysis, e.g. KEGG [73] or MPMP [74]. Combining metabolomics data allowed the construction of a metabolic model of *Plasmodium*, which in turn enabled the identification of 307 essential metabolic reactions [75]. This information is extremely useful for the design of new antimalarial drugs. A recent article explains how the host cell tropism of *P. falciparum* and *P. berghei* differs in the context of *Plasmodium* metabolism and connects the host metabolism with the *Plasmodium* parasite by indicating which of the reticulocyte's metabolic products can be used by the parasite [76]. This information allows further elimination of potential drug targets before running any actual clinical tests.

Conclusion

Obtaining a malaria vaccine requires a deep understanding of the mechanisms underlying *Plasmodium* adaptations. This goal can be achieved only by proper analysis of experimental data. A number of bioinformatics tools can be useful for predicting which genes undergo the most rapid evolutionary changes and the underlying mechanisms.

Key Points

- Malaria remains one of the world's most important health problems.
- The host–parasite interactions of malaria parasites are not well understood.
- The development of genomics and bioinformatics is playing an important role in the progress of current malaria studies.
- Sequence analysis has been successfully applied many times to predict protein functions and surface localization.
- The application of 'reverse ecology' allows better understanding of the evolutionary change rate of malaria genes and thus more efficient vaccine/drug selection.

Funding

This work conducted at the Institute of Biochemistry and Biophysics was supported by a grant from the National Science Centre of Poland (grant number 2014/13/B/NZ8/04719) to S.K.

References

1. Murray CJ, Rosenfeld LC, Lim SS, *et al*. Global malaria mortality between 1980 and 2010: a systematic analysis. *Lancet* 2012; **379**(9814):413–31.
2. Kleinschmidt I, Schwabe C, Benavente L, *et al*. Marked increase in child survival after four years of intensive malaria control. *Am J Trop Med Hyg* 2009;**80**:882–8.
3. Dourado HV. Malaria parasites, vectors and biologic cycle. *Rev Inst Med Trop Sao Paulo* 1992;**34(Suppl 9)**:S6–9.
4. Haldar K, Mohandas N. Malaria, erythrocytic infection, and anemia. *Hematology Am Soc Hematol Educ Program* 2009; **2009**(1):87–93.
5. Krogstad DJ, Gluzman IY, Kyle DE, *et al*. Efflux of chloroquine from *Plasmodium falciparum*: mechanism of chloroquine resistance. *Science* 1987;**238**(4831):1283–5.
6. Dondorp AM, Yeung S, White L, *et al*. Artemisinin resistance: current status and scenarios for containment. *Nat Rev Microbiol* 2010;**8**(4):272–80.
7. Schneider P, Bell AS, Sim DG, *et al*. Virulence, drug sensitivity and transmission success in the rodent malaria, Plasmodium chabaudi. *Proc Biol Sci* 2012;**279**(1747):4677–85.
8. Hargreaves K, Koekemoer LL, Brooke BD, *et al*. Anopheles funestus resistant to pyrethroid insecticides in South Africa. *Med Vet Entomol* 2000;**14**(2):181–9.
9. Nussenzweig RS, Vanderberg J, Most H, *et al*. Protective immunity produced by the injection of x-irradiated sporozoites of *Plasmodium berghei*. *Nature* 1967;**216**(5111):160–2.
10. Miller LH, Howard RJ, Carter R, *et al*. Research toward malaria vaccines. *Science* 1986;**234**(4782):1349–56.
11. Carter R, Mendis K. Transmission immunity in malaria: reflections on the underlying immune mechanisms during natural infections and following artificial immunization. *Mem Inst Oswaldo Cruz* 1992;**87(Suppl 3)**:169–73.
12. Luke T, Hoffman S. Rationale and plans for developing a nonreplicating, metabolically active, radiation-attenuated *Plasmodium falciparum* sporozoite vaccine. *J Exp Biol* 2003;**206 (Pt 21)**:3803–8.
13. Scarselli E, Tolle R, Koita O, *et al*. Analysis of the human antibody response to thrombospondin-related anonymous protein of *Plasmodium falciparum*. *Infect Immun* 1993;**61**(8):3490–5.
14. Hodgson SH, Ewer KJ, Bliss CM, *et al*. Evaluation of the efficacy of ChAd63-MVA vectored vaccines expressing circumsporozoite protein and ME-TRAP against controlled human malaria infection in malaria-naive individuals. *J Infect Dis* 2015;**211**(7): 1076–86.
15. Good MF, Miller LH, Kumar S, *et al*. Limited immunological recognition of critical malaria vaccine candidate antigens. *Science* 1988;**242**(4878):574–7.
16. Godson GN, Ellis J, Svec P, *et al*. Identification and chemical synthesis of a tandemly repeated immunogenic region of *Plasmodium knowlesi* circumsporozoite protein. *Nature* 1983; **305**(5929):29–33.
17. van Dijk MR, van Schaijk BC, Khan SM, *et al*. Three members of the 6-cys protein family of Plasmodium play a role in gamete fertility. *PLoS Pathog* 2010;**6**(4):e1000853.
18. Outchkourov NS, Roeffen W, Kaan A, *et al*. Correctly folded Pfs48/45 protein of *Plasmodium falciparum* elicits malaria transmission-blocking immunity in mice. *Proc Natl Acad Sci USA* 2008;**105**(11):4301–5.
19. van Dijk MR, Janse CJ, Thompson J, *et al*. A central role for P48/45 in malaria parasite male gamete fertility. *Cell* 2001;**104**(1): 153–64.
20. Saxena AK, Wu Y, Garboczi DN. Plasmodium p25 and p28 surface proteins: potential transmission-blocking vaccines. *Eukaryot Cell* 2007;**6**(8):1260–5.
21. de Koning-Ward TF, Olivieri A, Bertuccini L, *et al*. The role of osmiophilic bodies and Pfg377 expression in female gametocyte emergence and mosquito infectivity in the human malaria parasite *Plasmodium falciparum*. *Mol Microbiol* 2007;**67**(2): 278–90.
22. Goodman AL, Draper SJ. Blood-stage malaria vaccines—recent progress and future challenges. *Ann Trop Med Parasitol* 2010;**104**(3):189–211.
23. Clemens J, Moorthy V. Implementation of RTS, S/AS01 Malaria vaccine–the need for further evidence. *N Engl J Med* 2016;**374**(26):2596–7.
24. Mackinnon MJ, Gandon S, Read AF. Virulence evolution in response to vaccination: the case of malaria. *Vaccine* 2008;**26**: C42–52.
25. Krzyczmonik K, Świtnicki M, Kaczanowski S. Analysis of immunogenicity of different protein groups from malaria parasite *Plasmodium falciparum*. *Infect Genet Evol* 2012;**12**(8):1911–16.
26. Shahid M, Khan SER, Waters AP, Janse CJ, Kaczanowski S. Why are male malaria parasites in such a rush? Sex-specific evolution and host–parasite interactions. *Evol Med Public Health* 2013;**2013**(1):3–13.
27. Urban BC, Ferguson DJ, Pain A, *et al*. *Plasmodium falciparum*-infected erythrocytes modulate the maturation of dendritic cells. *Nature* 1999;**400**(6739):73–7.
28. Wykes MN, Good MF. What really happens to dendritic cells during malaria? *Nat Rev Microbiol* 2008;**6**(11):864–70.
29. Farias AS, Talaisys RL, Blanco YC, *et al*. Regulatory T cell induction during Plasmodium chabaudi infection modifies the clinical course of experimental autoimmune encephalomyelitis. *PLoS One* 2011;**6**(3):e17849.
30. Gardner M, Hall N, Fung E, *et al*. Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* 2002; **419**(6906):498–511.
31. Su XZ, Heatwole VM, Wertheimer SP, *et al*. The large diverse gene family var encodes proteins involved in cytoadherence and antigenic variation of *Plasmodium falciparum*-infected erythrocytes. *Cell* 1995;**82**(1):89–100.
32. Scherf A, Hernandez-Rivas R, Buffet P, *et al*. Antigenic variation in malaria: in situ switching, relaxed and mutually exclusive transcription of var genes during intra-erythrocytic development in *Plasmodium falciparum*. *Embo J* 1998;**17**(18): 5418–26.
33. Smith JD, Chitnis CE, Craig AG, *et al*. Switches in expression of *Plasmodium falciparum* var genes correlate with changes in antigenic and cytoadherent phenotypes of infected erythrocytes. *Cell* 1995;**82**(1):101–10.
34. Kyes S, Horrocks P, Newbold C. Antigenic variation at the infected red cell surface in malaria. *Annu Rev Microbiol* 2001;**55**: 673–707.
35. Hoppe GJ, Tanny JC, Rudner AD, *et al*. Steps in assembly of silent chromatin in yeast: Sir3-independent binding of a Sir2/Sir4 complex to silencers and role for Sir2-dependent deacetylation. *Mol Cell Biol* 2002;**22**(12):4167–80.
36. Tonkin CJ, Carret CK, Duraisingh MT, *et al*. Sir2 paralogues cooperate to regulate virulence genes and antigenic variation in *Plasmodium falciparum*. *PLoS Biol* 2009;**7**(4):e1000084.

37. Freitas-Junior LH, Hernandez-Rivas R, Ralph SA, *et al*. Telomeric heterochromatin propagation and histone acetylation control mutually exclusive expression of antigenic variation genes in malaria parasites. *Cell* 2005;**121**(1):25–36.
38. Guizetti J, Scherf A. Silence, activate, poise and switch! Mechanisms of antigenic variation in *Plasmodium falciparum*. *Cell Microbiol* 2013;**15**(5):718–26.
39. Miller JL, Harupa A, Kappe SH, *et al*. Plasmodium yoelii macrophage migration inhibitory factor is necessary for efficient liver-stage development. *Infect Immun* 2012;**80**(4):1399–407.
40. Sun T, Holowka T, Song Y, *et al*. A plasmodium-encoded cytokine suppresses T-cell immunity during malaria. *Proc Natl Acad Sci USA* 2012;**109**(31):E2117–26. 21
41. Kaczanowski S, Zielenkiewicz P. A TIP on malaria (genomics). *Nat Biotechnol* 2003;**21**(7):733.
42. Fiscella M, Perry JW, Teng B, *et al*. TIP, a T-cell factor identified using high-throughput screening increases survival in a graft-versus-host disease model. *Nat Biotechnol* 2003;**21**(3): 302–7.
43. Hurd H, Grant K, Arambage S. Apoptosis-like death as a feature of malaria infection in mosquitoes. *Parasitology* 2006;**132**: S33–47.
44. Arambage S, Grant K, Pardo I, *et al*. Malaria ookinetes exhibit multiple markers for apoptosis-like programmed cell death *in vitro*. *Parasit Vectors* 2009;**2**(1):32.
45. Taylor-Brown E, Hurd H. The first suicides: a legacy inherited by parasitic protozoans from prokaryote ancestors. *Parasit Vectors* 2013;**6**(1):108.
46. Duszenko M, Figarella K, Macleod E, *et al*. Death of a trypanosome: a selfish altruism. *Trends Parasitol* 2006;**22**(11):536–42.
47. Kaczanowski S. Apoptosis: its origin, history, maintenance and the medical implications for cancer and aging. *Phys Biol* 2016;**13**(3):031001.
48. Kaczanowski S, Sajid M, Reece SE. Evolution of apoptosis-like programmed cell death in unicellular protozoan parasites. *Parasit Vectors* 2011;**4**:44.
49. Meslin B, Beavogui AH, Fasel N, *et al*. *Plasmodium falciparum* metacaspase PfMCA-1 triggers a z-VAD-fmk inhibitable protease to promote cell death. *PLoS One* 2011;**6**(8):e23867.
50. Muralidharan V, Goldberg DE. Asparagine repeats in *Plasmodium falciparum* proteins: good for nothing? *PLoS Pathog* 2013;**9**(8):e1003488.
51. Singh GP, Chandra BR, Bhattacharya A, *et al*. Hyper-expansion of asparagines correlates with an abundance of proteins with prion-like domains in *Plasmodium falciparum*. *Mol Biochem Parasitol* 2004;**137**(2):307–19.
52. Kreil DP, Ouzounis CA. Comparison of sequence masking algorithms and the detection of biased protein sequence regions. *Bioinformatics* 2003;**19**(13):1672–81.
53. Wootton JC, Federhen S. Statistics of local complexity in amino acid sequences and sequence databases. *Comput Chem* 1993;**17**(2):149–63.
54. Finn RD, Mistry J, Tate J, *et al*. The Pfam protein families database. *Nucleic Acids Res* 2010;**38**:D211–22.
55. Kaslow DC, Quakyi IA, Syin C, *et al*. A vaccine candidate from the sexual stage of human malaria that contains EGF-like domains. *Nature* 1988;**333**(6168):74–6.
56. Holder AA, Blackman MJ, Borre M, *et al*. Malaria parasites and erythrocyte invasion. *Biochem Soc Trans* 1994;**22**(2):291–5.
57. Marshall VM, Silva A, Foley M, *et al*. A second merozoite surface protein (MSP-4) of *Plasmodium falciparum* that contains an epidermal growth factor-like domain. *Infect Immun* 1997; **65**(11):4460–7.
58. Marshall VM, Tieqiao W, Coppel RL. Close linkage of three merozoite surface protein genes on chromosome 2 of *Plasmodium falciparum*. *Mol Biochem Parasitol* 1998;**94**(1):13–25.
59. Black CG, Wu T, Wang L, *et al*. Merozoite surface protein 8 of *Plasmodium falciparum* contains two epidermal growth factor-like domains. *Mol Biochem Parasitol* 2001;**114**(2):217–26.
60. Black CG, Wang L, Wu T, *et al*. Apical location of a novel EGF-like domain-containing protein of *Plasmodium falciparum*. *Mol Biochem Parasitol* 2003;**127**(1):59–68.
61. Kuehn A, Simon N, Pradel G. Family members stick together: multi-protein complexes of malaria parasites. *Med Microbiol Immunol* 2010;**199**(3):209–26.
62. Krogh A, Larsson B, von Heijne G, *et al*. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* 2001;**305**(3):567–80.
63. Bendtsen JD, Nielsen H, von Heijne G, *et al*. Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol* 2004;**340**(4): 783–95.
64. Hiller NL, Bhattacharjee S, van Ooij C, *et al*. A host-targeting signal in virulence proteins reveals a secretome in malarial infection. *Science* 2004;**306**(5703):1934–7.
65. Marti M, Baum J, Rug M, *et al*. Signal-mediated export of proteins from the malaria parasite to the host erythrocyte. *J Cell Biol* 2005;**171**(4):587–92.
66. Sargeant TJ, Marti M, Caler E, *et al*. Lineage-specific expansion of proteins exported to erythrocytes in malaria parasites. *Genome Biol* 2006;**7**:R12.
67. Aurrecoechea C, Brestelli J, Brunk BP, *et al*. PlasmoDB: a functional genomic database for malaria parasites. *Nucleic Acids Res* 2009;**37**:D539–43.
68. Li YF, Costello JC, Holloway AK, *et al*. "Reverse ecology" and the power of population genomics. *Evolution* 2008;**62**(12): 2984–94.
69. Klein RJ, Zeiss C, Chew EY, *et al*. Complement factor H polymorphism in age-related macular degeneration. *Science* 2005; **308**(5720):385–9.
70. Pearson TA, Manolio TA. How to interpret a genome-wide association study. *JAMA* 2008;**299**(11):1335–44.
71. Amato R, Lim P, Miotto O, *et al*. Genetic markers associated with dihydroartemisinin-piperaquine failure in *Plasmodium falciparum* malaria in Cambodia: a genotype-phenotype association study. *Lancet Infect Dis* 2017;**17**:164–73.
72. Leffler EM, Band G, Busby GBJ, *et al*. Resistance to malaria through structural variation of red blood cell invasion receptors. *Science* 2017;**356**. doi: 10.1126/science.aam6393.
73. Kanehisa M, Araki M, Goto S, *et al*. KEGG for linking genomes to life and the environment. *Nucleic Acids Res* 2008;**36**: D480–4.
74. Ginsburg H, Abdel-Haleem AM. Malaria Parasite Metabolic Pathways (MPMP) upgraded with targeted chemical compounds. *Trends Parasitol* 2016;**32**(1):7–9.
75. Huthmacher C, Hoppe A, Bulik S, *et al*. Antimalarial drug targets in *Plasmodium falciparum* predicted by stage-specific metabolic network analysis. *BMC Syst Biol* 2010; **4**(1):120.
76. Srivastava A, Creek DJ, Evans KJ, *et al*. Host reticulocytes provide metabolic reservoirs that can be exploited by malaria parasites. *PLoS Pathog* 2015;**11**(6): e1004882.