Published in final edited form as:

Nat Ecol Evol. 2018 February 01; 2(2): 377–387. doi:10.1038/s41559-017-0419-9.

Adaptation of *Plasmodium falciparum* to its transmission environment

Martin K. Rono^{1,2,3}, Mary A. Nyonda⁴, Joan J. Simam, Joyce M. Ngoi¹, Sachel Mok⁵, Moses M. Kortok¹, Abdullah S. Abdullah⁶, Mohammed M. Elfaki⁷, John N. Waitumbi⁸, Ibrahim M. El-Hassan⁹, Kevin Marsh^{1,3}, Zbynek Bozdech¹⁰, Margaret J. Mackinnon^{*}

¹Kenya Medical Research Institute–Wellcome Trust Research Programme, Kilifi, Kenya ²Pwani University Bioscience Research Centre, Pwani University, Kilifi, Kenya ³Centre for Tropical Medicine and Global Health, Nuffield Department of Clinical Medicine, University of Oxford, Oxford, UK ⁴Department of Microbiology and Molecular Medicine, Medical Faculty, University of Geneva, Geneva, Switzerland ⁵Columbia University Medical Center, New York, NY, USA ⁶Rochester Regional Health–Unity Hospital, Rochester, NY, USA ⁷Department of Microbiology and Parasitology, Faculty of Medicine, Jazan University, Gizan, Jazan, Saudi Arabia ⁸Walter Reed Army Institute of Research/Kenya Medical Research Institute, Kisumu, Kenya ⁹Faculty of Public Health and Tropical Medicine, Jazan University, Gizan, Jazan, Saudi Arabia ¹⁰School of Biological Sciences, Nanyang Technological University, Singapore. Margaret J. Mackinnon is unaffiliated

Abstract

Success in eliminating malaria will depend on whether parasite evolution outpaces control efforts. Here, we show that *Plasmodium falciparum* parasites (the deadliest of the species causing human malaria) found in low-transmission-intensity areas have evolved to invest more in transmission to new hosts (reproduction) and less in within-host replication (growth) than parasites found in high-transmission areas. At the cellular level, this adaptation manifests as increased production of reproductive forms (gametocytes) early in the infection at the expense of processes associated with multiplication inside red blood cells, especially membrane transport and protein trafficking. At the molecular level, this manifests as changes in the expression levels of genes encoding epigenetic and translational machinery. Specifically, expression levels of the gene encoding AP2-G—the transcription factor that initiates reproduction—increase as transmission intensity decreases. This is accompanied by downregulation and upregulation of genes encoding HDAC1 and HDA1—two histone deacetylases that epigenetically regulate the parasite’s replicative and reproductive life-

^{*}Correspondence and requests for materials should be addressed to M.J.M. mmackinnon.mackinnon@gmail.com.

Life Sciences Reporting Summary. Further information on experimental design is available in the Life Sciences Reporting Summary.

Author contributions

M.K.R., M.A.N., J.J.S., J.M.N., M.M.E., A.S.A., M.M.K. and M.J.M. collected the data. Z.B. and S.M. provided the microarray materials. I.M.E., J.N.W., K.M. and M.J.M. organized the field work. M.J.M. and M.K.R. prepared the manuscript.

competing interests

The authors declare no competing financial interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41559-017-0419-9>.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

stage programmes, respectively. Parasites in reproductive mode show increased reliance on the prokaryotic translation machinery found inside the plastid-derived organelles. Thus, our dissection of the parasite's adaptive regulatory architecture has identified new potential molecular targets for malaria control.

During their life cycle, malaria parasites alternate between asexual replication in the bloodstream of their vertebrate host and reproduction in their mosquito vector. Asexual replication leads to disease, whereas reproduction leads to transmission to new hosts, and these trade against each other. Evolutionary theory predicts that parasites will adapt their replication–reproduction schedule according to the transmission opportunities and constraints imposed by the prevailing environment^{1–3}. This has consequences for the population disease burden, as well as risks to individuals, potentially undermining the efficacy of control programmes⁴. In the light of recent widespread success in reducing malaria transmission⁵, here, we tested whether parasites are naturally pre-adapted to their transmission environment and therefore might evolve in response to human intervention. We compared transcriptomes of *Plasmodium falciparum* parasites ex vivo from non-immune children in East Africa living in areas with long-term differences in malaria transmission intensity (Fig. 1 and Supplementary Fig. 1). Three independent experiments were performed: two compared geographically isolated populations (high versus medium and medium versus low transmission) and a third compared populations from before and after a major decline in malaria within a medium-transmission area. Data on 4,615 genes measured in 96 parasite isolates at all stages of the parasite's 48 h replication cycle in the bloodstream (intra-erythrocytic development cycle (IDC); 580 samples in total; dataset 1 in Supplementary Table 1) revealed more than double the expected number of genes with significantly different expression levels between 'high' versus 'low' populations (hereafter referred to as H–L genes; Supplementary Tables 2 and 3). Many genes, especially those H–L downregulated, showed consistent effects across the three experiments and steady declines in expression levels across the full range of transmission intensities (Fig. 2c, Supplementary Table 3 and Supplementary Figs. 2 and 3). These results were robust to substantial biases in the data caused by host, gene and parasite factors (Supplementary Tables. 4 and 5 and Supplementary Figs. 2 and 4–7). Our data support the hypothesis that natural populations of *P. falciparum* are locally adapted to the level of transmission intensity in their environment.

Because parasites in high- and low-transmission populations were sourced from non-immune children (three or fewer previous clinical episodes) at the acute stage of their infections (Supplementary Table 1 and Supplementary Fig. 6), we attribute these H–L differences to general properties of the parasite population rather than facultative responses by individual parasite isolates to the immune or physiological status of their current host. To prosecute this claim, we augmented the data from one of the populations with parasite isolates from semi-immune children (dataset 2 in Supplementary Table 1) and then analysed within this population for genes showing significant responsiveness in expression levels to host immunity-related clinical factors (that is, host-responsive genes; Supplementary Table 4). H–L and host-responsive genes overlapped by only 13% (Supplementary Fig. 2). Furthermore, the set of genes significant for H–L differences did not change when adjustments for host factors were incorporated into the analyses (Supplementary Fig. 2).

Thus, we conclude that H–L differences were not due to parasite plasticity in response to cues from their immediate host, but instead reflected population-level adaptation to their wider environment.

H–L genes cluster by function

By mapping H–L genes onto a whole-genome co-expression network annotated for putative function, we revealed that H–L upregulated genes localize to regions associated with asexual stages in the blood and liver (where parasites spend ~10 days following a mosquito bite before emerging into the bloodstream), while H–L downregulated genes associated with sexual stages in the blood (gametocytes) and mosquito (Fig. 2, Supplementary Figs. 8 and 9 and Supplementary Tables. 2 and 6). H–L-regulated genes tended to fall in the periphery of the network (Fig. 2b), implying their role in non-core processes. Gene set enrichment analyses confirmed H–L associations with asexual versus sexual function: while H–L upregulated genes were significantly concentrated in functional pathways involved in biosynthesis and development of asexual stages inside the red blood cell—especially membrane transport, protein trafficking, haemoglobin digestion, glycolysis and post-translational modifications—H–L downregulated genes were enriched in pathways involved in transmission and related functions (gametocyte and ookinete production, motility and lipid metabolism) (Fig. 2d, Supplementary Fig. 9 and Supplementary Tables. 2 and 7). Thus, parasites from high-transmission areas appear to invest more in asexual replication inside their host and less in transmission between hosts than their counterparts from low-transmission areas.

Individual H–L genes and their functions

The metabolic functions most associated with H–L upregulation were transport and protein trafficking (Fig. 2d and Supplementary Figs. 10 and 11). Both of these processes are fundamental to supporting the high biosynthesis demand of ~tenfold replication every 48 h in the human bloodstream⁶. Most facets of membrane transport (nutrient import, maintenance of ion concentrations and removal of toxins) were H–L upregulated, with the exception of transport across the membranes of the mitochondrion and apicoplast, the two organelles of plastid origin found in the parasite's cytosol. Examples of H–L-upregulated membrane transport genes were nucleoside transporter 1, which imports purines, the precursors for DNA which the parasite cannot make for itself⁷, and components of V-type ATPase proton pumps, which remove from the cytosol excess H⁺ generated by glycolysis (the main source of energy during the IDC) and maintain high H⁺ in the digestive compartment to enable the breakdown of haemoglobin for supply of amino acids for protein synthesis⁸. Likewise, most protein trafficking pathways were H–L upregulated, particularly for transfer of haemoglobin to the digestive compartment, nucleus-cytosol traffic and endoplasmic reticulum–Golgi traffic via vesicles. Localization of most transport and trafficking genes to the liver or early bloodstream regions of the network (lime green, pink and maroon network modules; Supplementary Figs. 9–11) indicate their specialized roles for life inside human cells. Genes involved in housekeeping functions, such as DNA replication, cell division and the supply of transfer RNAs (tRNAs) for protein synthesis (magenta

module), and those regulating clonal antigenic variation (see below; pale green module) were generally not H–L regulated (Fig. 2d and Supplementary Tables. 6 and 7).

The most H–L downregulated functional pathway was that for early-stage gametocytes (stages I and II, turquoise module; Fig. 2b), with many genes encoding hallmark proteins of gametocytes and gametes⁹ showing significant H–L downregulation in all experiments (Fig. 2c and Supplementary Table 2). Of particular note, among these was the gene encoding AP2-G (Fig. 2e), the DNA-binding transcription factor that is the ‘master key’ to unlock the developmental programme of gametocytes¹⁰ and a member of a family of DNA-binding transcription factors in *Plasmodium* credited with regulating life-stage-specific transcriptional programmes¹¹. During the IDC, *AP2-g* is epigenetically silenced by heterochromatin: this is mediated by the histone modifier class II histone deacetylase 2 (HDA2)¹². HDA2 also causes general silencing of *var* genes¹², the large gene family that encode proteins (PfEMP1) exported by the parasite to the red blood cell surface where they bind other host cells, causing life-threatening disease¹³. By virtue of PfEMP1’s high antigenicity and variability, coupled with sequential and mutually exclusive expression afforded by HDA2-mediated background silencing, the *var* gene family is believed to be responsible for prolonging infections in the face of the adaptive immune system¹⁴. HDA2, by silencing both *AP2-g* and *var* genes during the IDC, has thus been proposed as the molecular nexus governing the replication–reproduction divide¹².

Epigenetic control of H–L adaptation

In our data, the gene encoding HDA2 (PF10_0078) was not significantly H–L upregulated ($P = 0.73$). Instead, two other members of the five genes encoding class II histone deacetylases possessed by *P. falciparum* were: one upregulated (*HDAC1*, PF11260; $P = 0.04$) and one downregulated (*HDA1*, PF14_0690; $P = 0.004$). To further understand the role of class II histone deacetylases and other histone modifiers in the asexual–sexual divide, we built a subnetwork consisting of the core epigenetic machinery and H–L genes. We further included central components of the translational machinery (Supplementary Table 8), which is known to act as a major post-transcriptional regulator of life-stage differentiation in *Plasmodium*¹⁵. This sub-network structured into distinct regulatory modules that broadly corresponded to life stages (Fig. 3). At the heart of each was a single deacetylase or demethylase surrounded by a unique set of acetylases, methylases and AP2 transcription factors. This structure is consistent with a model in which the deacetylase or demethylase acts as a general silencer that ‘wipes’ the marks from the previous life stage ready for the acetylases, methylases and AP2 transcription factors to ‘write’ new sets of marks that coordinate specific within-stage transcriptional programmes. Based on their locations in the network, we deduce that HDAC1 and HDA1 provide the background silencing that underpins the general IDC and gametocyte transcriptional programmes, respectively, and we confirm the role of HDA2 in underpinning *var* gene regulation¹² (Fig. 3c and Supplementary Table 6). This, and their significant differential regulation, leads us to conclude that downregulation of *HDAC1* and upregulation of *HDA1* are the baseline events that entrain sexual-stage development in *P. falciparum*.

Translational control of H–L adaptation

Overlaying the epigenetic control of asexual–sexual differentiation was a corresponding compartmentalization of translational machinery into eukaryotic versus prokaryotic types (Fig. 3d). The parasite's mitochondrion and apicoplast both derive from prokaryotes and perform essential metabolic functions on behalf of the parasite¹⁶. However, the genomes of these organelles house just a few genes that encode these functions, the rest having transferred to the nuclear genome over evolutionary time¹⁷. These organelle genes are translated in situ using prokaryotic translation machinery¹⁸, which, as for the proteins that perform functions inside the organelles, is partly organelle-encoded and partly nucleus-encoded. We found that almost all prokaryotic translation machinery genes fell within the sexual-stage regions of the network (Figs. 3 and 4a) and that an unusually high number were H–L downregulated (Fig. 4b). This was especially so for those encoded by the apicoplast genome (6 of 16 versus 3 of 36 nuclear-encoded apicoplast-targeted translation machinery genes ($P = 0.03$) and versus 15 of all 280 translation machinery genes ($P < 0.001$) by hypergeometric tests), and for components not shared between the eukaryotic and prokaryotic machinery (organellespecific ribosomal proteins, modifiers of the initial methionine, initiation and elongation factors, and tRNA synthetases), but not for shared components such as the tRNAs (Fig. 3d). A clear example of this pattern is the gene encoding peptide deformylase (*PDF*, PFI0380c; $P = 0.002$; Fig. 3d), which, in prokaryotes but not eukaryotes, removes formylated methionine from the start position of the newly synthesized protein. (In *Plasmodium*, this requirement seems to apply to translation in the apicoplast but not the mitochondrion¹⁹). Overall, our data indicate that translation inside the organelles is at a premium during the sexual stages. While both organelles are active in asexual blood stages, and indeed are rendered indispensable during this phase by some of their metabolic functions^{20,21}, this conclusion is supported by the following previous evidence: (1) of the four metabolic functions performed by the apicoplast, three are essential during the sexual stages^{22–24} while the fourth maintains the apicoplast and hence the other three²⁵; (2) most mutations in genes of mitochondrial pathways are lethal only in the sexual stages²⁶; (3) upon conversion from asexual blood stages to gametocytes, the parasite transfers its burden of energy production from glycolysis in the cytosol to aerobic respiration in the mitochondrion^{27,28}; and (4) drugs that target prokaryotic translation machinery show high efficacy against sexual as well as asexual stages²⁹. This conclusion has vital practical significance: resistance mutations arising in blood-stage infections treated with drugs that inhibit prokaryotic translation, several of which are already in use, will not spread to the general population if these mutations are lethal during the sexual stages³⁰.

What controls the shift towards prokaryotic translation at the asexual–sexual transition? The switch is likely to fall high in the regulatory hierarchy because many organelle- and nucleus-encoded prokaryotic translation components need to be induced. A second shift in translation machinery upon switching to the reproductive mode, namely, the change in species of eukaryotic ribosomal RNA—the RNA partner to ribosomal proteins that together form the ribosomes—when the parasite differentiates to mosquito stages, as reported previously³¹, would likewise require coordination of the parasite's many ribosomal RNA-encoding genes. Both of these shifts may represent a mechanism to selectively translate

subsets of messenger RNAs and thereby gain control over synthesis of the highly specialized sexual-stage proteins³¹. Our network suggests that the forward switch (that is, transfer from human to mosquito) involves epigenetic silencing by histone lysine demethylase, JmjC2, and the reverse switch upon transfer back involves activation of the liver-stage AP2 transcription factor, AP2-L, accompanied by general silencing orchestrated by histone lysine demethylase, JmjC1 (Fig. 3c and Supplementary Table 6). The latter proposition is supported by evidence that dormant liver stages reactivate upon inhibition of histone lysine methylases³². We speculate that the remaining two of the five class II histone deacetylases, Sir2A and Sir2B, might be additional partners in the eukaryotic–prokaryotic translation shift given that Sir2A activity represses the production of ribosomal RNA³³ and that their encoding genes locate to network modules associated with general IDC function (maroon) and mosquito stages (khaki), respectively (Fig. 3).

In addition to the eukaryotic–prokaryotic shift, there was evidence of H–L-associated transcriptional regulation within the eukaryotic arm of translation. Connectivity between translational and epigenetic machinery genes was higher than expected (Fig. 3e) with many that encode eukaryotic translation initiation and elongation factors among the most connected of all the genes (Fig. 3e). These included eukaryotic elongation factor 2 (eEF2) and eukaryotic initiation factors (EIF) 2, 3, 4, 5 and 6. eEF2 and all subunits of the alpha and beta subunits of EIF2 were associated with IDC function and connected to H–L-upregulated genes more often than expected, although they were not themselves significantly upregulated. In contrast, genes for three of the nine subunits of EIF3 were significantly regulated (Fig. 3e), six of which belonged to liver-stage modules, with subunit K (EIF3K) being particularly pivotal between mosquito and liver stages (Fig. 3d). Different EIF3 subunits may thus play distinct roles in translational regulation at different life stages of *Plasmodium*. The alpha subunit of EIF2, EIF2 α , which prepares the ribosome for translation by loading methionine-bound tRNA to the start codon of messenger RNAs, is well known as a mediator of global translational repression in *Plasmodium*. This manifests as temporary stalling of development (latency) during life-stage transitions while the parasite awaits the opportunity to transfer to a new host (that is, in gametocytes³⁴ and sporozoites³⁵), as well as within the IDC while the parasite undergoes the energy-costly process of DNA replication for the production of daughter parasites³⁴. Within the IDC, amino acid starvation leads to translational stalling via EIF2 α and this slows down the entire transcriptional programme³⁶, implying feedback between the transcriptional and translational machineries. Tight co-regulation of the translational and transcriptional programmes is a feature of the IDC under normal growth and development conditions too^{37,38}. Ribosome accumulation on 5' leaders of many messenger RNAs indicates there is widespread regulation of start codon recognition³⁸. Combined, the evidence suggests that the translation initiation complex is central to coordination of the transcriptional and translation programmes throughout the parasite's life cycle. We have revealed potential upstream epigenetic regulators of this alliance (Supplementary Table 6) and propose that downstream regulators may be found among post-translational modifiers of EIFs, as in yeast³⁹.

Discussion

We interpret the increase in replicative versus reproductive effort in *P. falciparum* parasites from high- versus low-transmission intensity areas as the outcome of differential selection on parasite life-history trade-offs under two different evolutionary models (Fig. 5). Both of these models predict higher early-life investment in replication in environments where immunity and within-host competition are common, as in high-transmission areas, but they differ fundamentally in the nature of the trade-off assumed to drive parasite populations to different optima. Under model 1, increased replication is assumed to cost early-life reproduction via decreased gametocyte conversion rates pre-peak. Under model 2, the cost is the higher risk of host death: selection is thus not necessarily on conversion rates per se since higher replication may be achieved through other means. Since our blood samples contained mixed populations of asexuals and gametocytes, and because microarrays measure relative, not absolute amounts, we cannot invoke one model over the other. We note, nonetheless, that the H–L upregulation of within-IDC functions but not of between-IDC functions expected to affect replication rates per se (cell division, DNA replication), render our results more parsimonious with model 1, although they do not preclude model 2.

Since patients were recruited upon clinical presentation and excluded if they had progressed to severe malaria, our measurements took place during the infection period most relevant to models 1 and 2. We did not assay parasites in semi-immune hosts, the environment in which the selection is proposed to have occurred, nor did we measure lifetime fitness in both non-immune and semi-immune hosts. We cannot definitively conclude, therefore, that these models explain the H–L differences observed here. We note that our findings are consistent with those from our earlier experimental evolution study in laboratory mouse malaria in which we demonstrated that pre-immunity is a highly potent selection force for increased rates of pre-peak asexual replication, leading to greater lifetime reproductive output in both non-immune and semi-immune hosts⁴², as predicted by models 1 and 2. Nonetheless, alternative models—albeit less validated by theory, although well-substantiated by lifetime fitness data on *P. falciparum* in its natural setting—also predict H–L differences in response to selection later in the infection and from transmission-related factors other than immunity and in-host competition, such as mosquito population density and seasonality⁴³.

We considered whether H–L differences were due to ascertainment bias. Since gametocytes increase as the infection progresses (even if conversion rates do not⁴⁴), lower expression of gameto-cyte genes in H populations may have resulted from capturing the infection earlier than in low-transmission populations. We think this is unlikely since any previous immunity in high-transmission hosts would delay the time to clinical presentation. In contrast, the constant arrival of new infections in high-transmission populations would be expected to reduce the average age of the infection at presentation and thus dilute the gametocyte signal. The gene encoding PUF1, which is dominantly expressed in sporozoites, was found at higher levels than the reference transcriptome in Kisumu and Kilifi-pre parasites, but was almost undetectable in Kilifi-post and Sudan parasites (Fig. 3e), perhaps supporting such an explanation. However, this argument is countered by the fact that H–L differences were robust to adjustment for multiplicity of infection (Supplementary Fig. 2).

Our conclusion that epigenetic and translationally mediated ‘soft-wired’ regulation of the transcriptional programme underlies the ‘hard-wired’ H–L differences discovered here does not sit well with the traditional roles assigned to these machineries, namely, to enable the parasite to adjust to the environment of its current host, which varies unpredictably with each transmission cycle. Indeed, a plastic approach to reproductive investment is expected from theory^{2,40} and is empirically supported by observations that gametocyte conversion rates respond to environmental stressors such as drugs⁴⁵ and anaemia⁴⁶. However, the fact that we found H–L differences in the ‘common garden’ of non-immune hosts leads us to propose that the soft-wired control of H–L differentiation described above simply reflects downstream effectors in a cascade controlled by a higher-level switch that is subject to fixed-level variation in propensity to activate. Previous evidence supports the existence of hard-wired differences in rates of conversion to reproduction: (1) in vivo-maintained lines of the rodent malaria species *Plasmodium chabaudi* display stable differences in gametocyte production in a uniform host environment⁴¹; (2) *P. falciparum* lines maintained in the same in vitro environment can vary in their requirements for mitochondrial function⁴⁷; (3) in vivo transcriptional programmes of parasites infecting children in west Africa show distinct patterns that are gametocyte related but, as in the study here, cannot be explained by host factors⁴⁸; and (4) two west African populations of *P. falciparum* exhibit strong differentiation in a genomic region housing a gene affecting the development of early gametocytes⁴⁹. Only point 4 specifically implicates genetic rather than epigenetic control of these differences. Based on our evidence of the robustness of H–L differences to a wide range of environmental factors, we have interpreted the H–L differences as fixed, hard-wired and therefore likely to be genetic rather than epigenetic in origin. However, without deep exploration of the DNA variability associated with these differences, this conclusion cannot be definitive.

The practical implications of our findings are that when malaria control programmes lead to a reduction in transmission intensity, gametocyte production early in life increases and asexual replication rates decrease. Such counter-adaptation is unlikely to threaten malaria elimination, however. This is, first, because there exist other constraints that maintain low absolute levels of reproductive activity^{3,40,50}; second, because the efficacy of transmission-reducing devices (for example, bednets) far exceeds the capacity of the parasite to compensate via upregulation of gametocyte production; and third, because early-life gametocyte production is small relative to total lifetime production. If death and infection length also decrease concomitantly with replication rates, as they appear to in human and mouse malaria⁴¹, evolution of lower early-life replication in response to transmission control programmes is expected to generate even greater reductions in the overall disease burden. Such a win–win situation reinforces our earlier conclusion that malaria control programmes that directly target transmission are likely to be more sustainable in the long term than those that target asexual replication⁴. With this principle in mind, and armed with knowledge of the crucial molecules that dictate the fitness of *P. falciparum* in its natural environment, as revealed here, new and sustainable control strategies to achieve malaria elimination can be devised.

Methods

Experimental design

With permission from the National Ethical Review Committees of Kenya (protocol SSC1292) and Sudan, parasite isolates were obtained from children below 13 years of age with three or fewer previous clinical episodes of malaria who were diagnosed with uncomplicated *P. falciparum* malaria at hospitals and dispensaries in western Kenya (2008, 'Kisumu'), coastal Kenya ('Kilifi') and eastern Sudan (2007, 'Sudan'). These populations were confirmed to have high, medium and low malaria transmission intensities, respectively, at the time of sampling based on malaria surveillance data from the Malaria Atlas Project (<http://www.map.ox.ac.uk/>). Within the Kilifi population, parasites were sampled before and after a sustained reduction in transmission intensity (1994–1996, 'Kilifi-pre' and 2010–2012, 'Kilifi-post'). Parasite transcriptomes from high versus low-transmission populations, as measured against a reference genome, were compared in three separate experiments; namely, Kisumu versus Kilifi-post (experiment A), Kilifi-post versus Sudan (experiment B) and Kilifi-pre versus Kilifi-post (experiment C). This yielded two geographic and one temporal high versus low population comparisons, each involving a common reference population (Kilifi-post) (dataset 1 in Supplementary Table 1). In a fourth experiment, data from parasites taken from children in the medium–low transmission population with a history of many previous clinical episodes were introduced to the Kilifi 2010–2012 data (dataset 2). These data were then analysed for within-population responsiveness to host clinical traits.

Sample preparation

Blood samples from children (3–5 ml) were washed and placed in *in vitro* cultures within 5 h of sampling in experiments A and B, or were revived from frozen ring-stage parasites in experiment C using standard procedures. To obtain data from all stages of the parasite's 48 h blood-stage replication cycle (the IDC), every 10 h, up to 9 times, samples of the culture were washed and stored in TRIzol, and their RNA was extracted using phenol-chloroform, as previously described⁵¹. Some 500 ng of total RNA from each time point was linearly amplified and competitively hybridized against equivalent reference material (prepared from pooled RNA from six equally spaced time points in a synchronized culture of a laboratory-adapted field isolate, P4; ref. ⁵¹) to a DNA microarray containing ~10,000 70-mer oligonucleotide probes representing ~5,500 genes⁵². Printing of microarray slides, complementary DNA preparation, and reference isolate preparation and hybridizations were performed independently for each experiment. Within experiments, samples were randomized across population groups (high- versus low-transmission populations) and maturation time points before processing and then processed in batches of eight.

Stage of maturity estimation

The stage of maturity of each sample at each time point was estimated to within 2 h intervals by comparison with the reference time-course transcriptome from laboratory isolate Dd2 using previous methods^{51,53}. Time point samples for which the Pearson correlation between the reference transcriptome and sample was < 0.2, and all samples from isolates for which

the estimated stages of maturity did not increase across consecutive sampling time points, were excluded from the data for further analysis.

Data pre-processing

Data from poor-quality spots (regression ratio <0.6 by GenePix Pro 6.0 software; Molecular Devices, http://mdc.custhelp.com/app/answers/detail/a_id/18691), probes containing a single nucleotide polymorphism (based on data from 129 genomes in PlasmoDB), probes matching multiple genes in version 3 of the 3D7 genome, genes annotated as pseudogenes, *var*, *rif* and *stevor* (that is, highly variable) genes, and probes with twofold or greater differences between the samples and the reference isolate when averaged across all samples (an indication of a deletion or polymorphism in the reference isolate; 336 probes, 45 of which code for genes involved in invasion) were excluded from the analyses. $\log_2$ expression ratios were normalized within and between arrays using the ‘normexp’ and ‘quantile’ methods, respectively, in the *limma* package⁵⁴ in R (ref. ⁵⁵).

Pre-adjustment for stage of maturity

Loess curves were fitted to time-course data on each gene using all probes and data from all experiments combined. Curves were also fitted to data from each experiment separately in order to allow for potential systematic differences between experiments in maturation patterns. Residuals from these curves were used for tests of high versus low population differences and host responsiveness (see below). The Loess curves were used to determine the time of maximum expression, t_{max} , and periodicity using the Lomb–Scargle method implemented in R (ref. ⁵⁶). Genes were classified as ‘aperiodic’ if their P value for cyclic behaviour was $> 10^{-15}$.

Testing for high- versus low-transmission-intensity population differences

For each gene, the difference between high- versus low-transmission-intensity populations (denoted ‘H–L contrast’) and its ratio to the residual variance (estimated across all genes), Student’s t value, and corresponding P value with and without Benjamini–Hochberg adjustment for multiple testing (‘nominal’ and ‘false discovery rate’ P values, respectively) were estimated by fitting linear models in the *limma* package⁵⁴ to the expression data from dataset 1. Analyses were performed on data with and without previous adjustment for stage of maturity by Loess curve fitting. The first model fitted (the ‘population-specific adaptation’ model) included fixed effects for population type (high versus low) within experiment. Pairwise contrasts between all possible high versus low population pairs, both within experiment and between different experiments, were constructed (Supplementary Fig. 1). Where these involved different experiments, experiment-specific values for the reference (Kilifi-post) population were subtracted to remove systematic differences between the experiments. Statistical tests of these pairwise population contrasts are denoted ‘test 1’. A second model (the ‘generalized adaptation’ model) was fitted with fixed main effects for population type (high versus low) and experiment: this model thus estimated the average H–L effect across experiments and assumes that these are similar across the range of transmission intensities represented in the three experiments. Statistical tests under this model are denoted ‘test 2’. A third model, likewise testing for generalized adaptation, fitted a fixed effect linear covariate for population transmission intensity, thus estimating H–L

effects as a function of gradient in transmission intensity while ignoring any systematic experiment effects. The regression slope was evaluated for significance (test 3). In all models, a random ‘block’ effect for isolates that allowed for correlations between multiple time points on the same parasite isolate was included⁵⁷.

Testing for host responsiveness

Each gene was further analysed for its responsiveness in expression levels to host traits by analysing within the Kilifi-post population after augmenting it with data from semi-immune hosts (dataset 2). The model fitted a linear covariate for one of seven traits relating to host physiological or immune status (Supplementary Table 4) and a random effect for isolate. Host responsiveness analyses were similarly performed in dataset 1 by incorporating a covariate for host trait into the models testing for the H–L effects described above (Supplementary Table 5).

Post-adjustment for gene traits

H–L contrasts were post-adjusted for systematic bias from gene traits (the sexual-to-asexual expression ratio, time of maximum expression, single nucleotide polymorphism density and non-synonymous-to-synonymous single nucleotide polymorphism ratio) by fitting a linear model with a fixed effect for each of these traits using the ‘lm’ function in the *stats* package in R (ref. ⁵⁵) and then taking the residuals for downstream analyses (Supplementary Table 4). Sexual-to-asexual expression ratios for each gene were calculated as the average expression values in stage II gametocytes, stage V gametocytes and ookinetes relative to the average expression in rings, early trophozoites, late trophozoites and schizonts based on RNA sequencing data⁵⁸. Single nucleotide polymorphism variables were based on the values from 129 isolates using data downloaded from PlasmoDB (<http://plasmodb.org/plasmo/>).

Analysis combinations

To test for the influence on H–L effects of data pre-adjustment, data subsetting and biological factors (host and gene properties), analyses were performed for all combinations of the following variables: (1) data with and without parasite densities $< 25,000 \mu\text{l}^{-1}$ (datasets 1.1 and 1.2, respectively); (2) data from three alternative lengths of interval around the maximum (6, 12 or 24 h either side of t_{max} ; that is, representing one-quarter, one-half or the full 48 h cycle); and (3) three alternative methods of stage of maturity adjustment using Loess curves (across experiment, within experiment and none). For each of these 18 combinations, H–L effects were estimated with and without adjustment for one of seven host traits (by incorporation as a covariate in the linear model) and with and without post-adjustment for one or all four gene traits. For analyses of host responsiveness effects, the same model–dataset combinations as above were applied, but to dataset 2 and dataset 1.2. This gave, for each gene, a total of 171 tests for H–L effects and 126 tests for host responsiveness effects (Supplementary Tables. 4 and 5).

A ‘reference analysis’ for H–L effects was defined as that using dataset 1.2 and all time points (that is, an interval length of 48 h), pre-adjusted by Loess curves fitted across experiments with no covariates for host traits in the model and with no post-adjustment for gene traits.

Significance criteria

For individual analysis combinations, H–L genes were declared significant if the ‘H–L’ nominal P value was < 0.05 , unless stated otherwise. For results from across-analysis combinations, H–L genes were declared significant if the geometric mean nominal P value across all 171 combinations was < 0.05 , unless stated otherwise. Robustness was defined as the proportion of individual analysis combinations with a nominal P value < 0.05 . These thresholds were applied to all H–L tests between pairs of experiments and across all experiments; that is, for tests 1–3, but, where reported, refer to test 2 (H–L effects averaged over the three experiments) unless stated otherwise. For downstream analyses of the regulatory network and functional enrichment, significant genes were defined as those reaching the threshold for across-analysis combinations for test 2.

Numbers of observed versus expected significant genes

Numbers of H–L significant genes expected under the null hypothesis of no H–L population differentiation were obtained by performing the reference analysis on data in which the H–L status of each isolate within experiment was randomly reassigned before analysis. One-tailed tests of excess of observed versus expected numbers of significant genes were calculated from the empirical distribution generated from 105 such data permutations.

Tests of overlap

Evidence of generalized adaptation was sought by testing for excess overlap between pairs of experiments in their sets of significant genes. Since the Kilifii-post population was shared by all three experiments, H–L contrasts from experiments A, B and C were correlated (negatively between experiments A and B, and between B and C; positively between A and C). This led to higher-than-expected discordance or concordance between the experiments in H–L up- and downregulated genes than if the experiments were independent. To account for this, significant genes were defined as those outside the 95% confidence limits of the bivariate distribution of H–L effects for each pair of experiments (computed from the covariance matrix between the two experiments using the *SIBER* package in R; ref. ⁵⁹). Genes were further re-classified as up- or downregulated according to whether they fell above or below the regression lines (‘adjusted axes’) obtained by fitting H–L contrasts from each experiment on the other; that is, by their residuals after accounting for the between-experiment relationships. The proportion of significant genes in common for each pair of experiments for each of the four possible combinations of up- and downregulation in the experiment pair was calculated relative to the total number of significant genes in the quadrants defined by the adjusted axes and which fell outside the 95% confidence limits. Values were compared with the empirical distribution obtained from 105 permuted datasets generated as described above.

As a further test for consistency in H–L effects across the range of transmission intensities, overlap tests were performed on individual experiment contrasts (that is, by test 1) formed not involving the same populations (for example, Kisumu–Sudan contrast versus Kilifii-pre–Kilifii-post contrast). Overlap tests were also performed on the results from an individual experiment and those from the other two experiments combined using test 2. The above tests for overlap were applied to the results from the reference analysis only, after

excluding the five isolates in experiment B that were shared with the other experiments (Supplementary Table 1).

Fisher's exact tests were used to evaluate overlap between sets of H–L genes identified by different tests (test 2 versus test 3), for results with versus without host or gene trait adjustment, and for H–L versus host-responsive genes. These were performed on the results pooled across all analysis combinations.

Gene set enrichment analyses

Gene set enrichment tests were performed on sets of genes categorized into 255 'metabolic pathways' using information from the Metabolic Pathways for Malaria database (<http://mpmp.huji.ac.il/home>) and other sources (Supplementary Table 7). These were applied to sets of H–L up- and downregulated genes separately using a categorical 'membership test' based on the hypergeometric distribution. These were applied to H–L genes based on results pooled across analysis combinations.

Full-genome network analysis

To identify sets of co-regulated genes and their response to H–L population status, gene correlation network analyses were conducted using WGCNA⁶⁰ in R following the method in ref.⁶¹. The input adjacency matrix was constructed from across-isolate gene-pair correlations computed from residuals after curve-fitting to data from datasets 1 and 2 combined. After checking for outlier isolates by hierarchical clustering, the adjacency matrix was raised to the power of 6 to give a scale-free topology. This was used to compute the topological overlap similarity matrix from which gene clusters ('modules') were defined using hierarchical clustering. Pearson correlations between eigenvalues of these clusters and experimental factors (H–L status and host clinical variables) were calculated. For each gene, the number of connections to other genes among the strongest 5% of network links was computed within and between modules, and across the entire network. Networks were visualized using the *igraph* package⁶² and the Fruchterman–Reingold algorithm after pruning to the strongest 5% of linkages then excluding genes with no connections.

Epigenetic and translation machinery sub-network

A sub-network was constructed that contained genes encoding central components of the epigenetic machinery (histone modifiers and deacetylases, histone acetylases, histone methylases, histone demethylases and AP2 transcription factors; $n = 53$), translation machinery (ribosomal proteins, tRNAs, tRNA synthetases, translation initiation and elongation factors, and processors of the initial methionine involved in translation initiation; $n = 280$), RNA-binding gene families implicated as regulators of gene subsets at life-stage transitions (the DNA/RNA-binding family of ALBA proteins, believed to be master regulators of the life-stage-specific translational programme in *Plasmodium*, and thus the translation equivalent of the ApiAP2 family; $n = 6$)⁶³ and two other RNA-binding proteins from the Puf family, which have been shown to repress the translation of subsets of messenger RNAs at the asexual–gametocyte and sporozoite–liver-stage transitions^{64,65}. In addition, H–L genes with $P < 0.01$ ($n = 136$) were included.

Analysis of network hubs

To identify potential key transcriptional regulators ('hubs') and their relationship to metabolic function, the top 100 genes for the number of connections to genes that were H–L significantly upregulated, significantly downregulated, significantly up- or downregulated and not significantly H–L regulated were identified. For each of these 400 hub genes, membership functional enrichment tests were performed on their connecting genes ('hub communities'). Each hub gene community was tested for enrichment of H–L genes, as above, and the results were pooled across hubs belonging to the same network module.

Figures in circular format showing enrichment patterns and gene properties were constructed using the *circlize* package in R (ref. ⁶⁶).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This paper is published with the permission of the director of the Kenya Medical Research Institute (KEMRI). The authors are grateful to the study participants and the parasite culture laboratory at the KEMRI–Wellcome Trust Research Programme, Kilifi, Kenya. We also thank G. McFadden, M. Greischar and A. Read for helpful comments and H. Ginsburg for assistance with the gene sets for the enrichment tests. This work was supported by the Wellcome Trust (grant numbers 088634 to M.J.M. and 092741 and 077176 to K.M.).

Code availability

Other than the code used for data manipulation before analysis, no custom code outside the R packages cited was used.

Data availability

The datasets generated and/or analysed during the current study are available in the Gene Expression Omnibus repository, <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE108034>.

References

1. Stearns, SC. The Evolution of Life Histories. Oxford Univ. Press; New York: 1992.
2. Reece SE, Ramiro RS, Nussey DH. Plastic parasites: sophisticated strategies for survival and reproduction? *Evol Appl*. 2009; 2:11–23. [PubMed: 20305703]
3. Mideo, N; Day, T. On the evolution of reproductive restraint in malaria. *Proc R Soc B*; 2008. 1217–1224.
4. Gandon S, Mackinnon MJ, Nee S, Read AF. Imperfect vaccines and the evolution of parasite virulence. *Nature*. 2001; 414:751–755. [PubMed: 11742400]
5. Noor AM, et al. The changing risk of *Plasmodium falciparum* malaria infection in Africa: 2000–10: a spatial and temporal analysis of transmission intensity. *Lancet*. 2014; 383:1739–1747. [PubMed: 24559537]
6. Kirk K, Lehane AM. Membrane transport in the malaria parasite and its host erythrocyte. *Biochem J*. 2014; 457:1–18. [PubMed: 24325549]

7. El Bissati K, et al. The plasma membrane permease PfNT1 is essential for purine salvage in the human malaria parasite *Plasmodium falciparum*. *Proc Natl Acad Sci USA*. 2006; 103:9286–9291. [PubMed: 16751273]
8. Saliba KJ, Kirk K. pH regulation in the intracellular malaria parasite, *Plasmodium falciparum*. H(+) extrusion via a V-type H(+)-ATPase. *J Biol Chem*. 1999; 274:33213–33219. [PubMed: 10559194]
9. Silvestrini F, et al. Protein export marks the early phase of gametocytogenesis of the human malaria parasite *Plasmodium falciparum*. *Mol Cell Proteom*. 2010; 9:1437–1448.
10. Kafack BF, et al. A transcriptional switch underlies commitment to sexual development in malaria parasites. *Nature*. 2014; 507:248–252. [PubMed: 24572369]
11. Painter HJ, Campbell TL, Llinas M. The Apicomplexan AP2 family: integral factors regulating *Plasmodium* development. *Mol Biochem Parasitol*. 2011; 176:1–7. [PubMed: 21126543]
12. Coleman BI, et al. A *Plasmodium falciparum* histone deacetylase regulates antigenic variation and gametocyte conversion. *Cell Host Microbe*. 2014; 16:177–186. [PubMed: 25121747]
13. Smith JD, et al. Switches in expression of *Plasmodium falciparum* var genes correlate with changes in antigenic and cytoadherent phenotypes of infected erythrocytes. *Cell*. 1995; 82:101–110. [PubMed: 7606775]
14. Recker M, et al. Transient cross-reactive immune responses can orchestrate antigenic variation in malaria. *Nature*. 2004; 429:555–558. [PubMed: 15175751]
15. Zhang M, Joyce BR, Sullivan WJ Jr, Nussenzweig V. Translational control in *Plasmodium* and *Toxoplasma* parasites. *Eukaryot Cell*. 2013; 12:161–167. [PubMed: 23243065]
16. Sheiner L, Vaidya AB, McFadden GI. The metabolic roles of the endosymbiotic organelles of *Toxoplasma* and *Plasmodium* spp. *Curr Opin Microbiol*. 2013; 16:452–458. [PubMed: 23927894]
17. Wilson RJM, Gardner MJ, Feagin JE, Williamson DH. Have malaria parasites three genomes? *Parasitol Today*. 1991; 7:134–136. [PubMed: 15463469]
18. Chaubey S, Kumar A, Singh D, Habib S. The apicoplast of *Plasmodium falciparum* is translationally active. *Mol Microbiol*. 2005; 56:81–89. [PubMed: 15773980]
19. Pino P, et al. Mitochondrial translation in absence of local tRNA aminoacylation and methionyl tRNA Met formylation in Apicomplexa. *Mol Microbiol*. 2010; 76:706–718. [PubMed: 20374492]
20. Yeh E, DeRisi JL. Chemical rescue of malaria parasites lacking an apicoplast defines organelle function in blood-stage *Plasmodium falciparum*. *PLoS Biol*. 2011; 9:e1001138. [PubMed: 21912516]
21. Painter HJ, Morrissey JM, Mather MW, Vaidya AB. Specific role of mitochondrial electron transport in blood-stage *Plasmodium falciparum*. *Nature*. 2007; 446:88–91. [PubMed: 17330044]
22. Ke H, et al. The heme biosynthesis pathway is essential for *Plasmodium falciparum* development in mosquito stage but not in blood stages. *J Biol Chem*. 2014; 289:34827–34837. [PubMed: 25352601]
23. Van Schaijk BC, et al. Type II fatty acid biosynthesis is essential for *Plasmodium falciparum* sporozoite development in the midgut of *Anopheles* mosquitoes. *Eukaryot Cell*. 2014; 13:550–559. [PubMed: 24297444]
24. Wiley JD, et al. Isoprenoid precursor biosynthesis is the essential metabolic role of the apicoplast during gametocytogenesis in *Plasmodium falciparum*. *Eukaryot Cell*. 2015; 14:128–139. [PubMed: 25446055]
25. Gisselberg JE, Dellibovi-Ragheb TA, Matthews KA, Bosch G, Prigge ST. The suf iron-sulfur cluster synthesis pathway is required for apicoplast maintenance in malaria parasites. *PLoS Pathog*. 2013; 9:e1003655. [PubMed: 24086138]
26. Jacot D, Waller RF, Soldati-Favre D, MacPherson DA, MacRae JJ. Apicomplexan energy metabolism: carbon source promiscuity and the quiescence hyperbole. *Trends Parasitol*. 2016; 32:56–70. [PubMed: 26472327]
27. Lang-Unnasch N, Murphy AD. Metabolic changes of the malaria parasite during the transition from the human to the mosquito host. *Annu Rev Microbiol*. 1998; 52:561–590. [PubMed: 9891808]
28. MacRae JJ, et al. Mitochondrial metabolism of sexual and asexual blood stages of the malaria parasite *Plasmodium falciparum*. *BMC Biol*. 2013; 11:67. [PubMed: 23763941]

29. Delves M, et al. The activities of current antimalarial drugs on the life cycle stages of *Plasmodium*: a comparative study with human and rodent parasites. *PLoS Med.* 2012; 9:e1001169. [PubMed: 22363211]
30. Goodman CD, et al. Parasites resistant to the antimalarial atovaquone fail to transmit by mosquitoes. *Science.* 2016; 352:349–353. [PubMed: 27081071]
31. Gunderson JH, et al. Structurally distinct, stage-specific ribosomes occur in *Plasmodium*. *Science.* 1987; 238:933–937. [PubMed: 3672135]
32. Démbélé L, et al. Persistence and activation of malaria hypnozoites in long-term primary hepatocyte cultures. *Nat Med.* 2014; 20:307–312. [PubMed: 24509527]
33. Mancio-Silva L, Lopez-Rubio JJ, Claes A, Scherf A. Sir2a regulates rDNA transcription and multiplication rate in the human malaria parasite *Plasmodium falciparum*. *Nat Commun.* 2013; 4:1530. [PubMed: 23443558]
34. Zhang M, et al. PK4, a eukaryotic initiation factor 2alpha(eIF2alpha) kinase, is essential for the development of the erythrocytic cycle of *Plasmodium*. *Proc Natl Acad Sci USA.* 2012; 109:3956–3961. [PubMed: 22355110]
35. Zhang M, et al. The *Plasmodium* eukaryotic initiation factor-2alpha kinase IK2 controls the latency of sporozoites in the mosquito salivary glands. *J Exp Med.* 2010; 207:1465–1474. [PubMed: 20584882]
36. Babbitt SE, et al. *Plasmodium falciparum* responds to amino acid starvation by entering into a hibernatory state. *Proc Natl Acad Sci USA.* 2012; 109:E3278–E3287. [PubMed: 23112171]
37. Bunnik EM, et al. Polysome profiling reveals translational control of gene expression in the human malaria parasite *Plasmodium falciparum*. *Genome Biol.* 2013; 14:R128. [PubMed: 24267660]
38. Caro F, Ah Yong V, Betegon M, DeRisi JL. Genome-wide regulatory dynamics of translation in the *Plasmodium falciparum* asexual blood stages. *eLife.* 2014; 3:e04106.
39. Beilsten-Edmands V, et al. eIF2 interactions with initiator tRNA and eIF2B are regulated by post-translational modifications and conformational dynamics. *Cell Discov.* 2015; 1
40. Greischar MA, Mideo N, Read AF, Bjornstad ON. Predicting optimal transmission investment in malaria parasites. *Evolution.* 2016; 70:1542–1558. [PubMed: 27271841]
41. Mackinnon MJ, Read AF. Virulence in malaria: an evolutionary viewpoint. *Phil Trans R Soc Lond B.* 2004; 359:965–986. [PubMed: 15306410]
42. Mackinnon MJ, Read AF. Immunity promotes virulence evolution in a malaria model. *PLoS Biol.* 2004; 2:E230. [PubMed: 15221031]
43. Mackinnon MJ, Marsh K. The selection landscape of malaria parasites. *Science.* 2010; 328:866–871. [PubMed: 20466925]
44. Greischar MA, Mideo N, Read AF, Bjornstad ON. Quantifying transmission investment in malaria parasites. *PLoS Comput Biol.* 2016; 12:e1004718. [PubMed: 26890485]
45. Buckling AGL, Crooks L, Read AF. *Plasmodium chabaudi*: effect of antimalarial drugs on gametocytogenesis. *Exp Parasitol.* 1999; 93:45–54. [PubMed: 10464038]
46. Reece SE, Duncan AB, West SA, Read AF. Host cell preference and variable transmission strategies in malaria parasites. *Proc R Soc B.* 2005; 272:511–517.
47. Ke H, et al. Variation among *Plasmodium falciparum* strains in their reliance on mitochondrial electron transport chain function. *Eukaryot Cell.* 2011; 10:1053–1061. [PubMed: 21685321]
48. Daily JP, et al. In vivo transcriptional profiling of *Plasmodium falciparum*. *Malar J.* 2004; 3:30. [PubMed: 15296511]
49. Mobegi VA, et al. Genome-wide analysis of selection on the malaria parasite *Plasmodium falciparum* in West African populations of differing infection endemicity. *Mol Biol Evol.* 2014; 31:1490–1499. [PubMed: 24644299]
50. Taylor LH, Read AF. Why so few transmission stages? Reproductive restraint by malaria parasites. *Parasitol Today.* 1997; 13:135–140. [PubMed: 15275099]
51. Mackinnon MJ, et al. Comparative transcriptional and genomic analysis of *Plasmodium falciparum* field isolates. *PLoS Pathog.* 2009; 5:e1000644. [PubMed: 19898609]
52. Bozdech Z, et al. The transcriptome of the intraerythrocytic developmental cycle of *Plasmodium falciparum*. *PLoS Biol.* 2003; 1:85–100.

53. Mok S, et al. Artemisinin resistance in *Plasmodium falciparum* is associated with an altered temporal pattern of transcription. *BMC Genom.* 2011; 12:391.
54. Smyth, GK. Bioinformatics and Computational Biology Solutions Using R and Bioconductor. Gentleman, R, , et al., editors. Springer; New York: 2005. 397–420.
55. R Development Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing; Vienna: 2015.
56. Glynn EF, Chen J, Mushegian AR. Detecting periodic patterns in unevenly spaced gene expression time series using Lomb–Scargle periodograms. *Bioinformatics.* 2006; 22:310–316. [PubMed: 16303799]
57. Smyth GK, Michaud J, Scott H. The use of within-array replicate spots for assessing differential expression in microarray experiments. *Bioinformatics.* 2005; 21:2067–2075. [PubMed: 15657102]
58. Lopez-Barragan MJ, et al. Directional gene expression and antisense transcripts in sexual and asexual stages of *Plasmodium falciparum* . *BMC Genom.* 2011; 12:587.
59. Jackson AL, Inger R, Parnell AC, Bearhop S. Comparing isotopic niche widths among and within communities: SIBER–Stable Isotope Bayesian Ellipses in R. *J Anim Ecol.* 2011; 80:595–602. [PubMed: 21401589]
60. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics.* 2008; 9:559. [PubMed: 19114008]
61. Langfelder, P; Horvath, S. Tutorials for the WGCNA Package. <https://labs.genetics.ucla.edu/horvath/CoexpressionNetwork/Rpackages/WGCNA/Tutorials/index.html>
62. Csardi G, Nepusz T. The igraph software package for complex network research. *InterJournal Complex Systems.* 2006; 1695:1–9.
63. Vembar SS, Macpherson CR, Sismeiro O, Coppee JY, Scherf A. The PfAlba1 RNA-binding protein is an important regulator of translational timing in *Plasmodium falciparum* blood stages. *Genome Biol.* 2015; 16:212. [PubMed: 26415947]
64. Muller K, Matuschewski K, Silvie O. The Puf-family RNA-binding protein Puf2 controls sporozoite conversion to liver stages in the malaria parasite. *PLoS ONE.* 2011; 6:e19860. [PubMed: 21673790]
65. Miao J, et al. The Puf-family RNA-binding protein PfPuf2 regulates sexual development and sex differentiation in the malaria parasite *Plasmodium falciparum* . *J Cell Sci.* 2010; 123:1039–1049. [PubMed: 20197405]
66. Gu Z, Gu L, Eils R, Schlesner M, Brors B. circlize implements and enhances circular visualization in R. *Bioinformatics.* 2014; 30:2811–2812. [PubMed: 24930139]

Fig. 1. Experimental design and analysis.

a, *P. falciparum* parasites were sampled from three pairs of high (H) versus low (L) transmission populations in East Africa separated by space and time. **b**, Transcriptional profile (log₂ expression values relative to the reference isolate) for an example gene (PF11_0513) during the 48 h asexual stage maturation cycle. Dots represent observed values (orange, high transmission; blue, low transmission), while the black line shows the Loess fitted curve. Analysis of the curve's residual values revealed highly significant upregulation of this gene in high-transmission versus low-transmission populations (right panel: mean and 95% confidence intervals; $P < 0.001$ by significance test 2; see Methods). **c**, Overview of the data analysis.

Fig. 2. Full-genome gene correlation network.

a. Network of co-regulated genes coloured by module. The network is built from the strongest 5% of linkages between genes and this visual representation includes only genes with at least 50 connections ($n = 3,401$ out of 4,615 genes). The black curved lines approximately delineate the three life stages: asexual forms in the blood (maroon, pale green, magenta and forest green), mosquito stages (khaki and goldenrod) and liver stages (lime green, dark blue, cornflower blue, purple and pink) (Supplementary Fig. 8). Sexual stages in the blood and early in the mosquito phase occupy the periphery (turquoise, yellow

and tomato). **b**, Proportion of network modules with significant H–L-regulated genes ($P < 0.05$ by significance test 2), ordered from left to right by the module's average network connectivity, an index that negatively relates to occupancy of the network's periphery. **c**, Heat map showing significance (orange–blue colour legend) for genes with consistent H–L differences in the three experiments (significance test 1; first three columns) and across experiments (by significance tests 2 and 3; last two columns). The colour bar on the left side indicates the network module in **a** to which each gene belongs. **d**, Functional enrichment of H–L significant genes ($P < 0.05$ by test 2). The bar heights on the inner two tracks indicate the proportion of 171 analysis combinations, scaled from 0 to 0.5, with significant enrichment ($P < 0.05$ by membership enrichment test) of H–L genes (upregulated, orange; downregulated, blue) in each of 255 small functional groups (Supplementary Table 7) delineated by tick marks. Large and medium functional groups are indicated by the outer coloured track (coloured segments and adjacent labels) and inner grey-scale tracks ($n = 32$; black radiating labels), respectively. The third-from-outside track indicates the network module most strongly enriched for the small functional group. More details relating the network modules to functional groups are given in Supplementary Fig. 9. **e**, Across-experiment regression of expression ratios on transmission intensity (by significance test 3) for the *AP2-g* gene (PFL1085w). Coloured dots and vertical bars indicate means and 95% confidence intervals for high- (red colours) and low-transmission (blue colours) populations in each of the three experiments.

Fig. 3. Epigenetic and translation machinery networks.

a-d, Sub-network of genes coding for epigenetic and translational machinery alongside highly significant H-L-regulated genes ($P < 0.01$ by test 2), coloured by network module (**a**), H-L significance (**b**; see legend), classes of epigenetic machinery (**c**) and translational machinery (**d**; legends in panel **e**). Key genes mentioned in the main text are labelled with their abbreviated names (Supplementary Table 8). Symbol shapes indicate cellular location (square, mitochondrion; triangle, apicoplast; circle, cytosol). **e,** Connectivity of epigenetic machinery (right sector) and translational machinery genes (left sector). Colours in the outer

three tracks show the H–L significance (legend in panel **b**), network module and class of epigenetic machinery and translational machinery, respectively. The next track (grey bars) shows the numbers of connections to all other genes as a proportion of all genes (range 0–0.25). The inner three tracks show the ratio of the observed to expected number of connections to H–L-upregulated genes (orange), H–L downregulated genes (blue), epigenetic machinery genes (gold) and translational machinery genes (dark gold) on a $\log_2$ scale, ranging from 0 to 1. Ribosomal proteins in the bottom 80% for global connectivity were excluded in **e**.

Fig. 4. I Enrichment of organellar translation genes in downregulated H-L genes and sexual-stage network modules.

a. Distribution of classes of translation machinery genes (xaxis; Supplementary Table 8) across network modules (yaxis). Horizontal lines on the y axis delineate asexual blood stages, sexual blood stages, mosquito stages and liver stages, from top to bottom (Supplementary Table 6). **b.** Proportions of H-L-upregulated, -downregulated and -non-regulated genes by translation machinery classes (shared xaxis with **a**). api, apicoplast; iMet, initiator methionine; mit, mitochondrion.

Fig. 5. Selection for environment-specific levels of replication and reproduction under two different trade-off models.

Differential selection (orange and blue vertical arrows) acting on life-history trade-offs leads to different optimal levels of replication in high- versus low-transmission environments (orange and blue lines and text). **a.** Under model 1, a direct trade-off between replication and reproduction, indicated by a single asterisk, arises from the fact that gametocytes, which are formed from asexual parasites, do not replicate. Theoretical models⁴⁰ and empirical data⁴¹ have shown that low investment in reproduction (higher investment in replication) early in the infection leads to larger asexual populations and thus more gametocytes later in the infection, thus creating a trade-off between future and current reproduction^{2,40}. When host immunity and within-host competition, as occur in high-transmission areas, impose greater demands on replication ability, lower reproductive investment early in the infection is favoured⁴⁰, thus creating different outcomes in high- versus low-transmission environments (crossed blue and orange lines). **b.** Under model 2, a higher asexual replication rate is assumed to lead to more gametocytes throughout the infection (parallel orange and blue lines) and/or longer infections (persistence), but the trade-off (double asterisk) is the higher risk of host death (virulence), which causes zero transmission. In high-transmission areas, where immunity protects hosts from dying, the cost of virulence is reduced, thus allowing higher replication rates to evolve⁴. Model 2, like model 1, predicts that in-host competition will further drive replication upwards in high-transmission areas⁴. Note that the model 2 trade-off may also operate within model 1 (dotted arrows) and vice versa without altering their individual predictions of higher early asexual replication in high-transmission environments. Traits that impact fitness through early-life decisions are shown in green.

Fitness–trait relationships considered to be most influential under the model in question are shown with larger ‘+’ or ‘-’ symbols.