Sub-minute phosphoregulation of cell-cycle systems during *Plasmodium* gamete formation revealed by a high-resolution time course

Brandon M. Invergo[1,2*], Mathieu Brochet[2,3*], Lu Yu[2], Jyoti Choudhary[2+], Pedro Beltrao[1+], and Oliver Billker[2+]

6th April 2017

Abstract

As a first step in the transmission of malaria, *Plasmodium* parasites ingested by a mosquito must rapidly produce gametes prior to sexual reproduction. The phosphorylation-based signalling mechanisms underlying the initiation of gametogenesis remain largely unmapped. We have measured a high-resolution time course of phosphorylation in *P. berghei* gametocytes during the first minute of gametogenesis to elucidate the scope and dynamics of the response. The data reveals rapid phosphoregulation of microtubule motor proteins and DNA replication initiation proteins within the first 18 s and of DNA replication machinery by 60 s. The data also implicate several protein kinases and phosphatases in the gametogenesis signalling pathway. Through gene knock-out experiments, we verify that the protein kinases CDPK4 and SRPK1 have distinct influences over the phosphorylation of similar downstream targets. Together, the results show that key cell-cycle systems related to both replication and mitosis undergo simultaneous, rapid phosphoregulation.

1: European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Hinxton, United Kingdom
2: Wellcome Trust Sanger Institute, Hinxton, United Kingdom
3: Department of Microbiology & Molecular Medicine, University of Geneva, Switzerland
*: These authors contributed equally to this work.
+: Corresponding authors — JC: jc4@sanger.ac.uk, PB: pbeltrao@ebi.ac.uk, OB: ob4@sanger.ac.uk

Introduction

Malaria represents a major global health concern, causing an estimated 214 million cases and resulting in approximately 438 000 deaths in the last year [1]. However, much of the fundamental biology of the causative agents, intracellular parasites of the eukaryotic genus *Plasmodium*, remains unknown. Nearly half of the proteome for any given *Plasmodium* species remains unannotated for function and the molecular systems that are key to the parasite's survival are largely unmapped due to significant divergence from well-studied model organisms and to experimental challenges posed by the parasite.

Plasmodium species have a complex life cycle between a vertebrate host and a mosquito vector of the genus *Anopheles.* Parasite transmission to mosquitoes is initiated when developmentally arrested sexual-precursor stages, the gametocytes, are ingested by a susceptible mosquito: micro- and macrogametocytes respond to a small mosquito molecule, xanthurenic acid (XA), which must coincide with a drop in temperature to trigger their differentiation into male microgametes and female macrogametes [2]. Cytosolic Ca^{2+} levels rise steeply after a lag phase of 6–8 s and peak within the first 20 s of activation [3]. Within 15 s, a microtubule organising centre gives rise to eight kinetosomes [4]. As soon as 60 s after activation, microgametocytes have assembled monomeric tubulin from the nucleoplasm into the first mitotic spindle, and four axonemes start to grow on the templates of kinetosomes at each spindle pole [5]. Since the DNA content of non-activated gametocytes is only slightly in excess of the haploid value, it is likely that replication in the microgametocyte is also initiated during the first minute, *i.e.* before the first mitosis. The same is probably true for events leading to egress of both micro- and macrogametocytes from the host erythrocyte, since these are also dependent on the early Ca^{2+} signal [3].Within 8–10 min, each microgametocyte must replicate its genome three times, undergo three rounds of endomitosis, and assemble eight axonemes in order to produce eight microgametes. The microgametes then extrude from the gametocyte in a flagellar manner, through a process called exflagellation. Fertilisation of the macrogametes by microgametes results in the development of motile ookinetes, which may then escape the mosquito midgut.

The signal transduction pathway initiated through gametocyte activation by XA is an experimentally validated target for interventions to block parasite transmission to the mosquito [6], and a more detailed understanding of the mechanisms that link extracellular triggers to the cell and developmental cycle of the parasite is therefore of significant interest. While a parasite receptor for XA has remained elusive, a very early response of the gametocyte is thought to be a rise in intracellular cyclic guanosine monophosphate (cGMP) concentration through regulation of guanylyl cyclase and phosphodiesterase activity [7, 8], causing the activation of the cGMP-dependent protein kinase PKG [9]. PKG activity

stimulates the rapid release of intracellular Ca^{2+} stores via phosphoinositide metabolism [7, 9–11]. This Ca^{2+} signal, in turn, activates the Ca^{2+}-regulated kinase CDPK4 and the Ca^{2+}-regulated protein phosphatase calcineurin A (CnA), both of which are required to initiate DNA replication during microgamete formation [3, 12]. Another Ca^{2+}-dependent kinase, CDPK1, may well be activated by the same early Ca^{2+} signal but is only known to be required for the later event of gametocyte egress from the host cell and in female gametes is also essential for the subsequent expression of translationally repressed mRNAs [13]. A mitogen-activated protein kinase (MAPK2) has been implicated in the late stages of gametogenesis, playing a role in mitosis, cytokinesis, and axoneme motility [14–16]. Two other enzymes, an SR-protein phosphorylating kinase SRPK1 [17, 18] and a protein phosphatase PPM1 [19], have been found to be required to complete exflagellation, however more precise characterisations of their function have not yet been made. Any regulatory relationships between these enzymes are unknown. Nevertheless, it is clear that protein phosphorylation plays a central role in the regulation of gametogenesis.

Phosphorylation patterns have been investigated at a proteomic scale in *Plasmodium* parasites at various life cycle stages of different species [11, 20–27]. However, lack of large-scale functional knowledge in these species has made it difficult to assign observed phosphosites to functional modules. Also, the dynamic nature of phosphorylation implies that temporally relevant events are indistinguishable from established phosphosites. The construction of phosphoproteomic time courses, in which the phosphoproteome is measured at several time points following a perturbation, has been shown to be a powerful approach to study cell signalling responses [28–33]. The highly dynamic and rapid gametocyte activation process is thus an interesting candidate for time-course analysis. It has the advantage of being highly synchronised: activation of developmentally arrested gametocytes can be reliably and rapidly effected *in vitro* through a decrease in temperature and a concomitant introduction of XA [2]. In this manner, the early signalling events underlying gametogenesis can be readily observed with a low amount of cross-talk from other processes.

In an attempt to unravel the early signalling processes during gamete formation, we have constructed a high-resolution phosphorylation time course of the process in *P. berghei*. Our results implicate several kinases as members of a pathway comprising the early gametogenesis signalling response. Additionally, they show that systems that are fundamental to microgamete formation, particularly the assembly of axonemes, proteins required for mitotic spindle formation and for DNA-replication processes, undergo phosphoregulation almost instantaneously and simultaneously during the first minute of cell activation. We also present a resource of hundreds of regulated phosphosites of likely importance for controlling basic biological processes in malaria parasites.

Results

Robust quantification of phosphosites

In order to observe the global patterns of protein phosphorylation following gametocyte activation, we measured a high-resolution time course of the phosphoproteome of *P. berghei*. To decide which time points would provide the best picture of relevant regulated phosphosites, we considered that the activation of gametocytes by XA at a permissive temperature is very rapid (Figure 1a). Our time course therefore focuses on the first 18 s, at 6 s intervals, to capture the first cellular response to the initial Ca^{2+} signal. It finishes after the first minute, since this period should be sufficient to capture the key regulatory events associated with the induction of major cell cycle events in the microgametocyte. For each time point, we used four label-free biological replicates and produced two technical replicates from the each phosphopeptide-enriched sample (Figure 1b). A total of 17 238 unique peptides were identified, of which 8982 were phosphopeptides. 12 672 phosphosites were identified within these peptides, with 10 237 localised at site-level with high-confidence (confidence > 0.75); 8299 of these high-confidence sites were found on *P. berghei* proteins and 1938 were from *M. musculus* proteins (Supplementary Table S1). Phosphosite intensities, estimated from the intensities of the different phosphopeptides on which the sites are identified, were strongly correlated between replicates (Figure 1c). Furthermore, we detected several of the proteins and phosphosites observed in a previous, two-dimensional gel electrophoretic screen of phosphorylation in gametocytes and gametes [34]. Thus, the data represents a comprehensive and reproducible view of early phosphorylation during gametogenesis, from which we could reliably reconstruct time courses.

Phosphorylation time courses cluster into distinct response groups

We quantified the change in phosphorylation states over time by calculating log_2-ratios against unstimulated gametocytes (0 s), producing time courses of changes in abundance from 0–60 s. In order to account for sites which were not detectable as phosphorylated at the beginning or the end of the experiment, we also constructed two "truncated" time courses: 6–60 s time courses were constructed for sites that were not observed at 0 s and 0–18 s time courses for those that were not observed at 60 s.

We then devised a method to robustly distinguish sites that significantly changed during the time course, accounting for technical variation in the observed ratios, by exploiting the fact that the phosphopeptide enrichment was not 100 % specific. To this end, we constructed time courses, as described above, for the

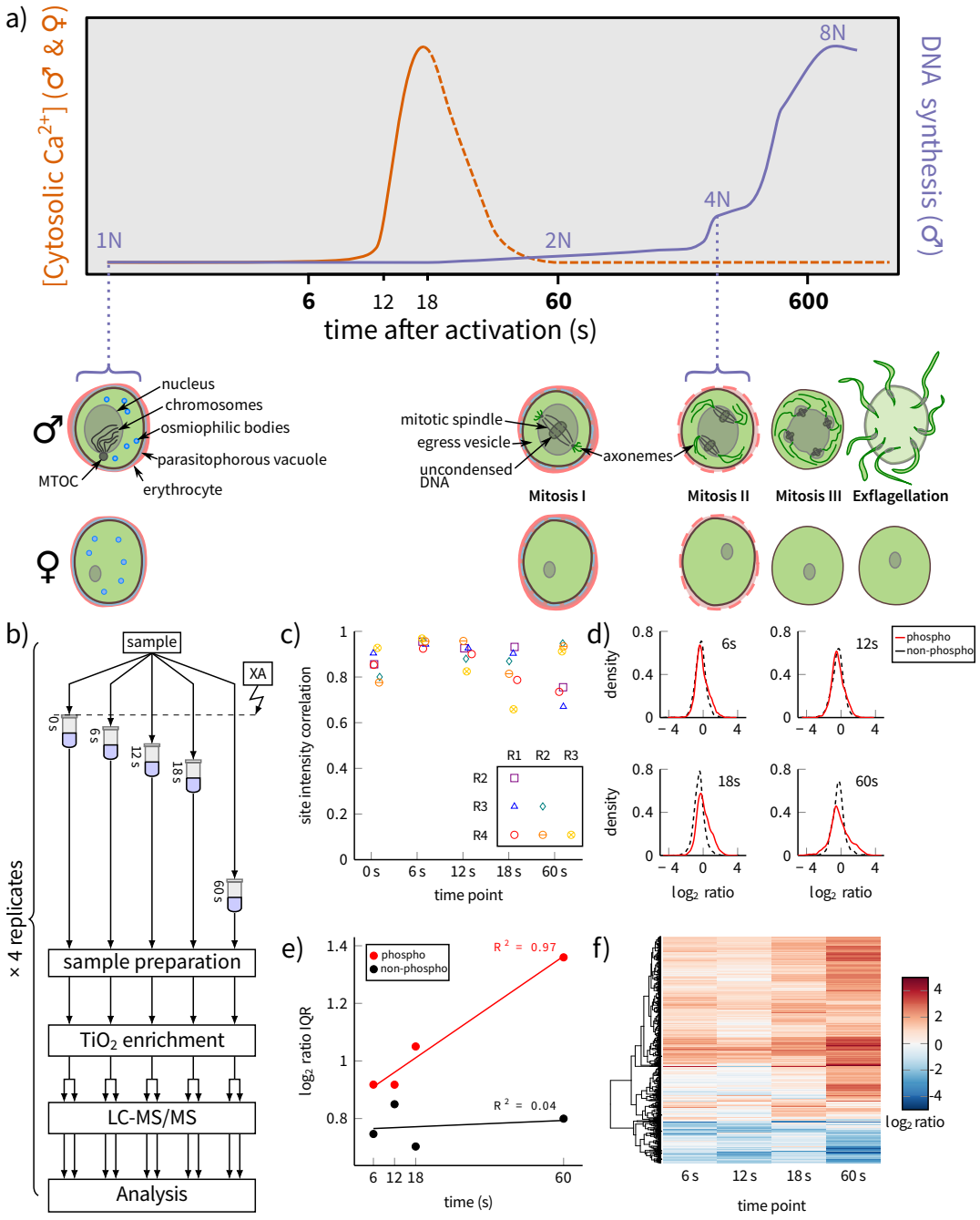

Figure 1: a) Timing of key events during *P. berghei* gametogenesis based on published data. Rapid mobilisation within 20 s of intracellular Ca^{2+} stores leads to egress from the parasitophorous vacuole and the erythrocyte in both micro- (♂) and macrogametocytes (♀). Microgametocytes also produce eight axonemes and undergo three rounds of DNA replication and endomitosis within 10 min, followed by exflagellation. (MTOC: microtubule organising centre) b) A schematic of the experimental design. Ca^{2+} and DNA synthesis dynamics were adapted from refs. [3] and [35], respectively. c) Correlation coefficients for phosphosite intensities between biological replicates for each time point (R1: replicate 1, R2: replicate 2, *etc.*). d) Kernel density distributions of log_2-ratios for phosphosites and non-phosphopeptides at the 6, 12, 18 and 60 s time points relative to unstimulated gametocytes (0 s). The non-phosphopeptideratio distributions were used as null-distributions to test significance of the phosphosite ratios. e) The interquartile ranges (IQRs) of log_2-ratio distributions from (d), with least-square regression lines and Pearson correlation coefficients shown. The increase in IQR for phosphosites represents the phosphorylation signal spreading out to a wider range of targets over time. f) Hierarchical clustering of log_2-ratios of the significantly regulated phosphosites.

non-phosphorylated peptides that we detected. Given the short time scale of the experiment, we assumed that any variation in the intensities of these non-phosphopeptides could be attributed to random, technical noise. We could then use this to construct null distributions of log_2-ratios, against which we tested for significant variation in the phosphosite ratios (Figure 1d). The interquartile ranges (IQRs) of the phosphosite ratio distributions were found to increase with time while those of the non-phosphopeptides remained constant (Figure 1e), supporting our assumptions. This growth in the IQR over time may be interpreted as the signalling response spreading out to a wider range of targets over the course of the experiment [33].

In this way, 1089 phosphosites on 549 *P. berghei* proteins were found to show evidence of significant change during one of the three time courses (referred to herein as "regulated sites", however it is important to emphasise that this is in reference to regulation of phosphorylation state and may not in all cases correspond to regulation of protein activity). For the majority of regulated sites, full, 0–60 s time courses were measurable (0–60 s: 926 sites; 6–60 s: 64 sites; 0–18 s: 99 sites; Supplementary Table S2). Hierarchical clustering of all significant time courses revealed that the majority of the sites show positive trends ("up-regulation") and only about a fifth of the sites show consistently negative trends ("down-regulation") (Figure 1f). To gain finer detail, we employed a Gaussian mixture-model clustering algorithm to re-partition the data (Supplementary Table S2). Using this method, the 0–60 s time courses were partitioned into eight clusters: seven clusters that show general up-regulation and one showing down-regulation (Figure 2a). Clusters 1, 2, and 3 consist of sites that are already up-regulated within 6–12 s, while clusters 4, 5, and 6 show predominantly later up-regulation. Cluster 7 contains up-regulated sites that did not fit well in the other clusters. For the truncated time courses, the algorithm also produced two clusters each (one showing up-regulation and one showing down-regulation) (Figure 2a; 6–18 s: clusters 8 and 11; 0–18 s: clusters 9 and 12). Through this clustering of the time course results, we see that the phosphosites in the parasite are characterised by several classes of dynamic responses to gametocyte activation, with many displaying activity within the first seconds after the initial stimulus.

DNA-replication and microtubule-related proteins are primary targets of phosphoregulation

In order to detect biological trends in the observed phosphoregulation, we next tested for enrichment in particular functional annotations in the regulated phosphoproteins. Of the 549 significantly regulated proteins, 436 were annotated with at least one GO term. We tested for enrichment in the sets of proteins carrying phosphosites in each time-course cluster (Figure 2b; see Supplementary Figure **??** and

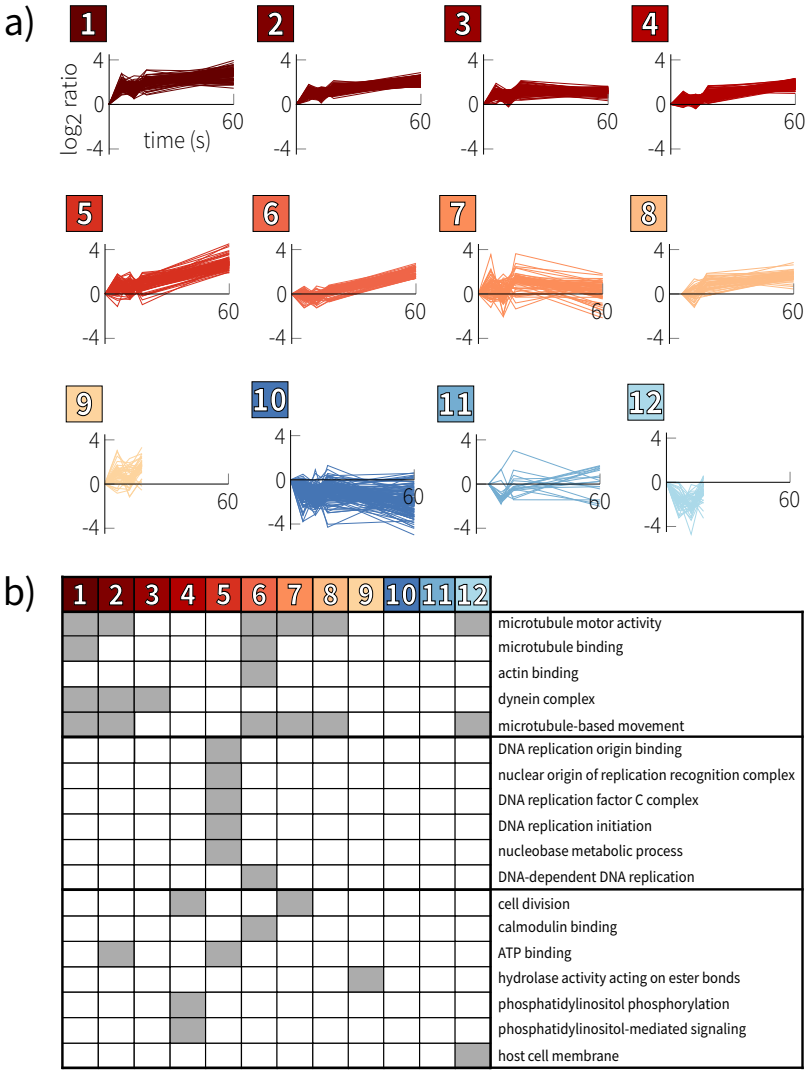

Figure 2: a) Gaussian mixture-model-based clustering of *P. berghei* phosphorylation time courses. b) Enriched Gene Ontology terms for the *P. berghei* time-course clusters.

Supplementary Table S3 for more details, including enrichment of higher-level parent terms). Most clusters individually show enrichment for microtubule-related terms (Figure 2a; clusters 1, 2, 3, 6, 7, 8, and 12, following the numbering scheme in Figure 2a); this indicates that dynamic phosphoregulation of microtubule motor proteins occurs throughout the first minute after gametocyte activation. Other clusters together suggest that proteins related to DNA replication-initiation and nucleic acid metabolism are moderately phosphorylated within the first 18 s (cluster 5), and these proteins along with DNA replication machinery are significantly phosphorylated by 60 s (clusters 5 and 6). We also observed that one early and one late up-regulation cluster (clusters 2 and 6) were enriched in male-specific proteins ($p = 0.035$ and $p = 0.023$ respectively), while another early up-regulation cluster (cluster 3) was enriched for female-specific proteins ($p < 0.001$) and the primary down-regulation cluster (cluster 10) was enriched for proteins that are exclusive to neither male nor female gametocytes ($p = 0.035$).

The broad trends of these results suggest that a large portion of the response is dominated by phosphoregulation of DNA replication and microtubule-related proteins in male gametocytes. Within the first minute of microgametogenesis, spindle plaques need to split, nucleate the spindle of mitosis I, give rise to basal bodies and start growing axonemes. In other eukaryotes, microtubules are nucleated from a ring of γ-tubulin in complex with a family of interacting proteins, many of which are phosphorylated by mitotic protein kinases [36]. In *Plasmodium*, only γ-tubulin and one component of the ring complex, PBANKA_083430, appear to be conserved. The latter is characterised in our data by an early phosphorylation event from 6 s. Similarly, phosphosites in the centrosome component centrin-2, in a number of axonemal dynein chains and in other axoneme associated proteins [37] are contained in the rapid-response clusters 1–3. This is unexpected, because it suggests that mitotic kinases may become active at the same time as DNA replication is initiated, not after the cell has completed S-phase.

Interestingly, while phosphatidylinositol (PI) signalling is known to be triggered early by PKG and is required for the rapid mobilisation of Ca^{2+} [11], we saw enrichment for PI metabolism in sites which show predominantly late up-regulation (cluster 4). Closer inspection of the data reveals a more differentiated picture, however. This cluster, indeed, includes phosphosites on a PI 3-kinase (PI3K, PBANKA_111490), a PI 4-phosphate-5-kinase (PI4P5K, PBANKA_020310) and a PI 4-kinase (PI4K, PBANKA_110940). The PI4K, a validated drug target [38], carries two regulated sites that are both up-regulated within 6 s, as would be predicted: S534 (cluster 2) and S538 (the aforementioned site in cluster 4). However, S538 has more complex dynamics, with its initial phosphorylation suppressed by 18 s and then reappearing by 60 s, hence its inclusion in the late up-regulation cluster. Later phosphorylation events in the pathway such as this may be involved in signal termination or restoration of IP3 levels. It is noteworthy that in another life-cycle stage of *P. berghei*, the motile ookinete, the same sites were identified as putative

targets for PKG and were shown to be of functional importance during ookinete gliding [11]. In the same work, cGMP-mediated phosphorylation was reported within the Sec7 domain of a putative guanine nucleotide exchange factor for the small GTPase ARF (PBANKA_130680), which in *P. falciparum* can interact with a PI4P5 kinase to regulate its activity [39]. The neighbouring serine, S1229, was found to undergo early phosphoregulation in our time course. Taken together, these data indicate that at least some phosphosites we have detected are functionally important and consistent with a previous analysis at another life-cycle stage, and that the functional enrichment analysis highlights high-level trends that are congruent with our limited existing knowledge of gametogenesis.

The phosphoregulated proteins are functionally associated

We investigated whether the regulated phosphoproteins tend to be functionally related to each other in a previously published protein-protein association network, in which associations are predicted from diverse data such as co-expression or co-evolutionary patterns [40]. We found that the subset of the network consisting of phosphoregulated proteins tended to have a higher number of associations between them than expected by random chance (Supplementary Figure **??**). We then visualised the subset of the network that consists of regulated proteins associated with the functional terms described in the previous section (Figure 3). If the phosphoregulated proteins represented a diverse set of systems, then this sub-network would be composed of several disconnected components. However, of the 124 proteins annotated with the functional terms, the largest connected component of the sub-network included 102 of them. In other words, many of the proteins that we detected as undergoing phosphoregulation during gametogenesis are functionally associated with each other. Visualising the network in this manner reveals two major clusters of proteins (Figure 3). One is composed of proteins annotated for microtubule motor activity and which tend to be phosphorylated at many different sites. The other consists of helicases, protein kinases, several other ATP-binding enzymes, and DNA-binding proteins, comprising the replisome. Other minor clusters are present, including one of ribosomal proteins and one related to phosphoinositide signalling. These clusters are interconnected via a CDK-related (cyclin-dependent kinase) protein kinase, CRK5 (PBANKA_123020), holding a clear central position between them. An orphan kinase PBANKA_062210, found within the replisome cluster, also had evidence for a very central position among these proteins. This connection between the replisome cluster and the microtubule cytoskeleton cluster suggests that the phosphorylated motor proteins are more likely to be related to the mitotic spindle than to the axonemes, and that regulation of the cell cycle following microgametocyte activation forms the core of the phosphorylation response.

Figures 3B–D show sub-networks of regulated proteins that are annotated with specific, relevant molecular function annotations: motor activity, helicase activity, and kinase and phosphatase activity, respectively. In Figure 3b, we again see a large, interconnected cluster of motor proteins with another, smaller cluster consisting of a few otherwise unannotated motor proteins. In Figure 3c, visualising the sub-network of regulated kinases and phosphatases reveals a tightly interconnected cluster, which comprises all of the enzymes that underwent phosphoregulation during the time course (none are omitted due to poor connectivity). Thus the kinases that we detected as undergoing phosphoregulation all have prior evidence of being functionally related. Finally, in Figure 3d, consisting of proteins annotated for helicase activity, we see a cluster containing several proteins of the DNA replication licensing factor (MCM) family and another cluster of helicases involved in DNA repair, as well as some RNA helicases. While these represent a broad range of helicase functionality, the connections between these clusters are limited.

Protein kinases and phosphatases associate with specific dynamic clusters

The phospho-signalling enzymes that regulate gamete formation and the regulatory relationships between them remain largely unknown. Likewise, the downstream functional pathways regulated by these enzymes are unclear. While we cannot infer direct enzyme-substrate relationships from the data set, we can use the information encoded in the protein-protein functional-association network to infer relationships between enzymes and phospho-regulated substrates (see Methods). This allows us to construct a coarse-grained view of the putative roles of the signalling enzymes. Using the protein-protein association network, we found statistical associations between enzymes and proteins from the phosphosite clusters for an aurora kinase (ARK2; PBANKA_040740), CRK5 (PBANKA_123020), the orphan kinase PBANKA_062210, NIMA-related kinase 1 (NEK1; PBANKA_144300), and a putative protein phosphatase 2C (PBANKA_142720) (Figure 4).

We found similar associations with late up-regulation and general down-regulation clusters for CRK5 (clusters 5 and 10), PBANKA_062210 (cluster 5) and the phosphatase 2C (clusters 5, 8 and 10), associating these enzymes with DNA replication via cluster 5. PBANKA_062210 was found to be strongly associated with the proteins of the replisome (Figure 3a), stressing the potential importance of this yet undescribed kinase. The phosphatase 2C has not been annotated with any of the enriched functional terms (thus it is not found in figure 3a), however it has evidence of functional association with CRK5 (Figure 3c). Accordingly, the association of these two enzymes with both phosphorylation and dephosphorylation clusters, albeit largely via a different set of proteins for each enzyme, could be explained by

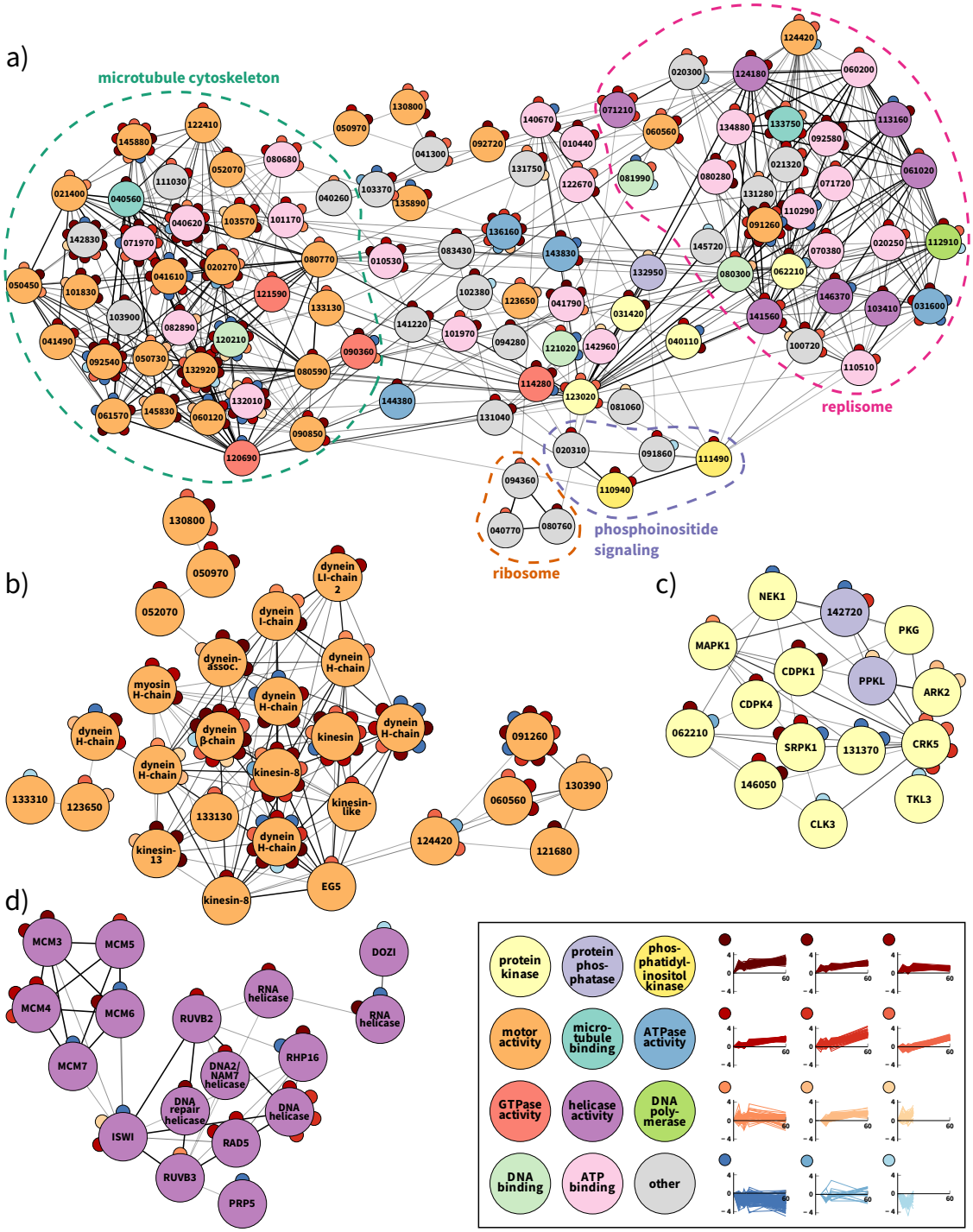

Figure 3: The proteins that are phosphoregulated following gametocyte activation have evidence of tightly interconnected functional associations. a) The subset of the functional-association network consisting of regulated phosphoproteins that were associated with enriched GO terms (minimum edge score of 300). Edges are shaded and weighted according to their edge score. Nodes are coloured according to Molecular Function annotations and contain the numeric portion of the protein's ID. Sub-nodes represent significantly regulated phosphosites, with their colour reflecting the time course cluster to which they belong. Some functional clusters are highlighted. b) The subset of the network consisting of regulated proteins annotated for motor activity. Nodes contain a description of the gene product if available, otherwise they contain the numeric portion of the protein ID. c) The sub-network of regulated proteins annotated for protein kinase (yellow) or protein phosphatase (light violet) activity (PKG, a kinase known to be active after gametocyte activation, is included, despite not having been detected as regulated). d) The sub-network of regulated proteins annotated for helicase activity.

their working in a complementary manner to regulate the same systems.

ARK2 is enriched in network interactions with proteins that carry phosphosites from two clusters characterised by rapid initial responses that remain stable (cluster 1) or increase over 60 s (cluster 6) (Figure 4). While these two sets of proteins are overall enriched in similar functional terms, the ARK2 interaction partners among them comprise two distinct sets of mainly kinesin and dynein chains. This links ARK2 to the phosphoregulation of the microtubule motor proteins, which is consistent with another member of the aurora kinase family being found at microtubule organising centres in *P. falciparum* [41]. This late up-regulation cluster (cluster 6) was also enriched in male-specific proteins, thus, we speculate that ARK2 plays a role in regulating mitotic processes during microgametogenesis. It is interesting to note that CRK5 is among the proteins from this cluster that are functionally associated with ARK2, suggesting that the regulatory functionality of ARK2 and CRK5 may be related.

The kinase NEK1 (PBANKA_144300), thought to be involved in MAPK signalling, is known to be essential and is expressed highly in males at the gametocyte stage [17, 41–44]. It was associated with a cluster characterised by initial down-regulation followed, in some cases, by eventual up-regulation (cluster 11). NEK1 was phosphoregulated at one site, characterised by rapid dephosphorylation within 6 s followed by recovery to pre-activation phosphorylation levels by 18 s. The two proteins driving its statistical association with the cluster, the kinase PBANKA_062210 and an unannotated protein PBANKA_124420, both have phosphosites that are characterised by the later up-regulation seen in the cluster, as described above. Thus, we would speculate that NEK1 is primarily active by 60 s and indeed may be functionally repressed during early signalling events. This is consistent with the role of MAPK signalling during later gametogenesis events [14].

This analysis places these enzymes alongside others that have been more directly implicated in gametogenesis through knock-out or loss-of-function experiments (Figure 4). However, beyond the initial events around the second messengers cGMP and Ca^{2+}, we cannot yet derive a hierarchical structure between the enzymes in order to reconstruct a gametogenesis signalling pathway. The second-messenger dynamics suggest that the signal originates with PKG and the calcium-dependent enzymes, but specific substrates of these and other enzymes remain largely unknown. In *P. falciparum*, CDPK1 has been identified as a substrate of PKG, so we indicate this relationship in Figure 4 [25]. Similarly, the NIMA-related kinases NEK1 and NEK3 of *P. falciparum* have been shown to be able to phosphorylate MAPK2 *in vitro*, suggesting that they could regulate the functionality of this kinase [43, 45]. This, in turn, would tentatively link one cluster (cluster 11) to mitosis and axoneme motility via NEK1 and MAPK2. Surprisingly, none of the previously implicated enzymes were otherwise found to be associated with the time-course

clusters. This is likely due to low coverage of the network, in which many of the enzymes, such as CDPK4 and SRPK1, are poorly represented. Nevertheless, the results introduce several new enzymes as likely components of the pathway, with their specific roles requiring further validation.

Conservation of phosphosites suggests functional constraint

If a phosphosite has a vital biological function, it can be expected to be conserved at the sequence level or phosphorylation at homologous sites should be observed within the protein family across species. We assessed the sequence conservation of phosphosites at the underlying nucleotide level by estimating the ratio of the number of non-synonymous substitutions to the number of synonymous substitutions (dN/dS) per site during species divergence (Supplementary Table S4). We defined conserved sites as those that show evidence of more selective constraint (lower dN/dS) than would be predicted by known determinants of evolutionary rates. The proteins with conserved phosphosites were enriched for motor activity, the myosin complex, calmodulin binding, and cytoskeletal protein binding. When limited to only proteins undergoing active phosphoregulation during gametogenesis, the set is enriched in proteins annotated for DNA helicase activity, DNA replication and replication-initiation, and the MCM complex.

We then focused on the conserved, regulated sites that fall within a predicted protein domain and which have previously been detected as phosphorylated in *P. falciparum*. Together, these would be likely indicators of important functionality. We identified thirteen such sites (Table 1). Notably, both CDPK1 and CDPK4 have homologous, conserved and regulated phosphosites that are located within their kinase domains, suggesting that these sites can be taken as indicators of the respective kinase's activity. This is supported by the fact that the orthologous CDPK1 site in *P. falciparum* has been identified as a site of autophosphorylation [46]. Several sites also fall within domains related to DNA replication and the cell cycle, for which evidence in other species suggests functionality. In particular, phosphorylation within the MCM N-terminal domain of human MCM7 has been implicated in the regulation of formation of the MCM2-7 complex and in the progression of the cell cycle [47] and phosphorylation within the MCM domain of human MCM3 regulates the replication process [48]. The phosphorylation of MAT1 (PBANKA_111190) within its eponymous domain is also potentially related to the regulation of the cell cycle is; in other eukaryotes, this protein (MNAT1 in humans) associates with CDK7 and cyclin H to form the CDK-activating kinase enzymatic complex (CAK), which then regulates cyclin-dependent kinases by phosphorylating them. However, no functionality of phosphoregulation of MNAT1 has yet been described.

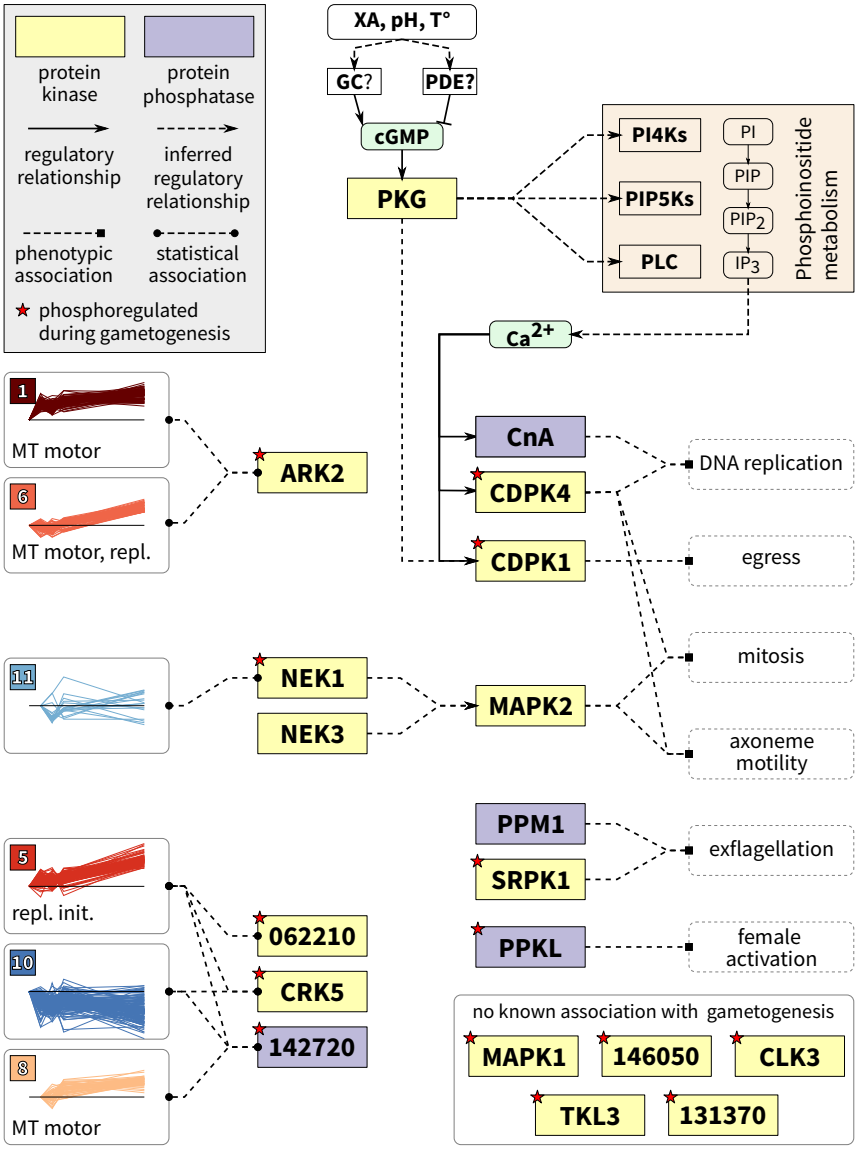

Figure 4: A schematic of early phosphorylation signalling during gamete formation. Several of the phosphoregulated kinases and phosphatases can be associated with the phosphosite clusters via statistical enrichment on a protein functional-association network; functional terms enriched in each cluster are listed (MT motor: microtubule motor activity; repl. init.: replication initiation; repl.: DNA replication). Other enzymes have been associated with specific stages of gametogenesis through knock-out or loss-of-function experiments. Inferred regulatory relationships include cases in which a kinase is known to be able to phosphorylate another kinase, but no specific functional has been described for the event.

Because we observed that many motor proteins undergo phosphoregulation at multiple sites, we next investigated whether similar phosphorylation patterns are conserved in similar proteins. If the regulated sites occur in regions of the proteins that are frequently observed to be phosphorylated within protein families, so-called phosphorylation hot-spots [49], then it may be an indication of functional importance. The dynein heavy chain and kinesin families were found both to be heavily phosphorylated along their full lengths (Supplementary Figure **??**). While no hot-spots were distinguishable in the dynein heavy chain family, several regulated sites fall within hot-spots at well-aligned positions in the kinesin family, particularly near the N- and C-termini. In a dynein light intermediate chain (PBANKA_041490), one of the two regulated phosphosites, T629 (Figure **??**a alignment position 850), falls in a major C-terminal phosphosite hot-spot. The orthologs in humans (O43237) and mice (Q6PDL0) have phosphoserines aligned with this position (S383 in both species), however no function has previously been described for this site. On the other hand, a N-terminal phosphorylation hot-spot in the dynein intermediate chain family has been found to have a regulatory role [50–55]. One site on PBANKA_050450, an ortholog of human cytoplasmic dynein 1 intermediate chain 1 (DYNC1I1), near this region (S120; Figure **??**b alignment position 232) was found to undergo regulation during gametogenesis. In human cells, phosphorylation of intermediate chain 1 within this hot-spot was found to mediate cargo binding in organelle transport [50, 51]. Such occurrences of regulated phosphosites within phosphorylation hot-spots suggest that these phosphorylation patterns have functional, regulatory roles.

Gene knock-outs link protein kinases to core signalling responses

We next sought to investigate the roles of specific protein kinases in the gametocyte activation response. We studied two knock-out lines, one lacking CDPK4 and the other with a deletion of SRPK1. Both kinase mutants make morphologically normal gametocytes but are defective in male gametogenesis [3,17]. Consistent with these genetically confirmed functions, both undergo phosphorylation during the time course, however we were not able to detect any specific functional associations for either kinase with the dynamic clusters (Figure 4). To elucidate the roles of these kinases during early gametogenesis, we compared phosphorylation patterns before and 18 s after activation in the knock-out lines and the wild type.

In the SRPK1 knock-out experiment (SRPK1-KO), 3949 peptides were quantified in the wild-type (WT) and knock-out (KO) samples, of which 2676 mapped to 1060 *P. berghei* proteins and the remaining to 787 *M. musculus* proteins (Supplementary Table S5). 236 peptides carried sites that were observed as regulated in the time-course experiment. In the CDPK4 knock-out experiment (CDPK4-KO), we

quantified 3539 unique peptides. 2491 of these peptides were mapped to 993 *P. berghei* proteins and the remaining were mapped to 663 host proteins (Supplementary Table S5) significant regulation in the time-course experiment, 190 were unambiguously quantified in CDPK4-KO. The small overlap between the different types of experiments can be accounted for largely by the inherent stochasticity of MS peptide detection in complex biological samples and potentially by the difference in time points measured.

We saw significant differences between KO and WT after activation in both experiments, however in SRPK1-KO we also saw differences between the two samples before activation (Figure 5a). Before studying the effects of deleting the kinases, we first checked that the WT samples replicated the observations of the time course experiment. The significantly up- and down-regulated WT phosphopeptides in both experiments were enriched in protein functions that generally matched those of the time course experiment, despite the shorter time-frame (18 s here versus 60 s in the time course) (Figure 5b). We therefore concluded that the WT samples reproduced the observations from the time course experiment and differences with the KO samples could be attributed to loss of the respective kinase.

Pre-activation differences between KO and WT are likely to be caused by a change in the steady-state of protein abundances or phosphorylation due to the missing kinase. If a phosphopeptide's state is not regulated (directly or indirectly) by that kinase, then its disruption in the KO versus WT 18 s after activation should be similar to that observed at steady-state prior to activation. Indeed, for both experiments, we found a positive correlation between the pre- and post-activation changes in phosphorylation, suggesting that many disruptions of phosphosites due to the absence of the kinases were already present prior to activation. Outlier peptides in this correlation were therefore interpreted to have been affected during the gametogenesis process specifically (Figure 5a). In SRPK1-KO, we identified 243 affected phosphosites, the majority of which increased in abundance in the KO versus WT after activation ("positive disruption"). In CDPK4-KO, we found 196 disrupted phosphosites, however in this case the change was largely negative. 16 of the phosphosites disrupted by loss of SRPK1 were observed in the time course, and these came from both up- and down-regulation clusters, while the 22 sites affected by deletion of CDPK4 that were also observed in the time course all came from up-regulation clusters. Thus, in the case of CDPK4 we see evidence for significant loss of phosphorylation when the kinase is deleted, while for SRPK1 we observed general disruption and a notable gain of phosphorylation events. This suggests SRPK1 may indirectly repress phosphorylation during gametogenesis via an intermediary enzyme.

We next performed an enrichment analysis on the disrupted proteins for the functional terms that were associated with the time course (Figure 5b). For both experiments, we found an enrichment in proteins related to the microtubule cytoskeleton. In CDPK4-KO, these were all negatively disrupted, while in

SRPK1 they were both positively and negatively disrupted. The phosphoregulation patterns reveal that in the case of CDPK4 the phosphosites were up-regulated in WT and unregulated in the KO after activation; while for SRPK1 the trend was reversed, with sites that were unregulated in the WT being mostly up-regulated in the KO (Figure 5c). Although some of these proteins were affected in both KO experiments, the specific phosphosites that were disrupted were unique in each case. Many of the motor proteins affected in SRPK1-KO, however, were unaffected in CDPK4-KO. Thus, CDPK4 and SRPK1 appear to both regulate different subsets of the microtubule motor system and the phosphorylation patterns support the hypothesis of a repressive function for SRPK1. We also saw a disruption in both experiments of DNA replication initiation, through components of the origin of replication initiation complex (ORC) and the MCM complex (Figure 5b). One interesting case is S60 on ORC1 (PBANKA_060200), which is up-regulated after activation in WT. When CDPK4 is not present, regulation of this site is lost. However, when SRPK1 is not present, the site significantly increases in abundance relative to WT. This suggests that the signal for phosphorylation of this site originates with CDPK4 and the stoichiometry of this event is moderated by the signal passing through an SRPK1-mediated feedback loop. This is supported by the observation that phosphorylation of SRPK1 itself is negatively disrupted when CDPK4 is deleted, placing SRPK1 downstream of CDPK4 in the early gametogenesis signalling events.

Discussion

Plasmodium gametocytes must detect the change in environment from host to vector and propagate that signal quickly and efficiently to initiate the necessary gametogenesis processes. Because this process is vital to the parasite's successful transmission, determining the underlying signalling events is a primary step towards a fundamental understanding of transmission. Our results reveal a prominent phosphorylation response within 18 s after gametocyte activation and that this signal is widespread by 60 s. In particular, despite having observed a large number of proteins undergoing active phosphoregulation during the response, we found a large fraction of the activity to occur in functionally related systems pertaining to DNA replication and mitosis, two processes known to be rapidly induced in the male gametocyte upon activation.

We observed phosphoregulation on several protein kinases and phosphatases, many of which are dispensable for the asexual erythrocytic cell cycle, but which were implicated through gene knockout studies in regulating gametocyte activation specifically, including CDPK1 and 4, MAPK2, SRPK1, CnA, PPM1 and PPKL [3, 12–15, 17–19]. In two cases, CDPK4 and CDPK1, we identified regulation of conserved

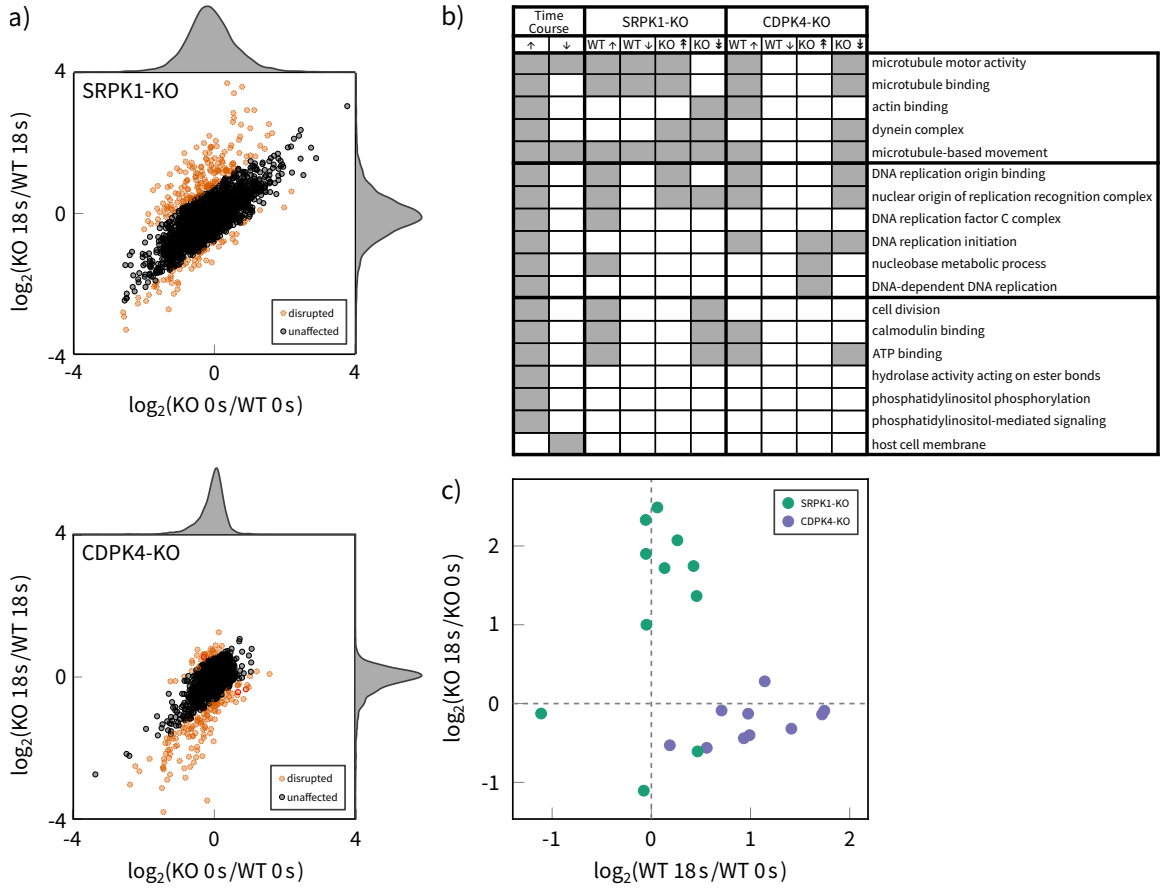

Figure 5: Individual deletions of the protein kinases CDPK4 and SRPK1 produce distinct effects on the phosphorylation response during early gametogenesis. a) Loss of CDPK4 or SRPK1 causes disruption versus wild type both before (0 s) and after (18 s) activation of the gametocytes. The degree of disruption in both states is greater in SRPK1-KO, as indicated by the distributions of fold-changes (marginal density plots). The outliers are peptides whose phosphoregulation was likely affected following gametocyte activation. CDPK4-KO appears to have caused mainly loss of phosphorylation (negative disruption) while SRPK1-KO caused mainly gain of phosphorylation (positive disruption). b) A comparison of GO-term enrichment results for the time-course experiment and the two knock-out experiments (WT: up/down-phosphoregulation; KO: positive/negative disruption). c) Disrupted phosphopeptides from motor proteins in the two KO experiments. Without SRPK1, several motor proteins become newly phosphorylated during gametogenesis, while lack of CDPK4 results in the loss of phosphorylation events. Each point represents a distinct phosphopeptide (*i.e.* each phosphosite was disrupted in only one of the two experiments).

phosphosites within their kinase domains, which may indicate regulation of their enzymatic activity during this process. For the other kinases that we observed undergoing phosphoregulation, it is more difficult to directly relate phosphorylation state to kinase activity levels. Nevertheless, through enrichment analyses in conjunction with the protein-association network, we could show that some of the kinases undergoing phosphoregulation after gametocyte activation are functionally associated with groups of proteins that are phosphoregulated with similar dynamics to each other. In particular, the evidence suggests that the phosphorylation states of ARK2, CRK5, NEK1, and the orphan kinase PBANKA_062210 change in response to gametocyte activation and that these enzymes are functionally associated with other proteins that undergo simultaneous phosphoregulation. This suggests that these enzymes may play central roles in the gametocyte activation response. While no functional data are available on the orphan kinase, the first three candidates, which all belong to more conserved protein kinase subfamilies implicated in regulating the eukaryotic cell cycle, have some evidence for expression in asexual stages, and attempts to disrupt these genes have been consistently unsuccessful in *P. berghei* [17, 20]. Roles for these protein kinases are therefore most likely not limited to gametogenesis, but may extend to regulating the *Plasmodium* cell cycle more generally. Stage-specific or inducible mutants in these genes have the potential to reveal deeper insights into the mechanisms of *Plasmodium* cell-cycle regulation. Interestingly, another cdc2-related protein kinase, CRK4, is required for both replication and mitosis in *P. falciparum* schizonts and has been implicated in phosphorylation events on the pre-replicative complex. CRK4 is also expressed in *P. falciparum* gametocytes and while it is required for *Plasmodium* transmission to the mosquito, it is not essential for exflagellation [27]. Consistent with this, in *P. berghei* gametocytes we have observed phosphopeptides belonging to CRK1, 3, 4 and 5, but only the latter shows significant evidence of marked changes in phosphorylation levels, suggesting CRK5 is the likely S-phase promoting kinase in the activated microgametocyte. However, this does not rule out potential roles for this kinase later in gametogenesis.

Unfortunately, this statistical analysis failed to reveal associations for known components of the pathway, such as CDPK4 and SRPK1. Since both protein kinases had previously been shown to be essential for exflagellation downstream of gametocyte emergence [3, 17], we used the existing mutants to elucidate further the roles that these two kinases might play in the process. Deletion mutants were characterised by the disruption of the phosphorylation processes observed in the time course, albeit in distinct ways. When CDPK4 is not present, many sites which are phosphorylated after activation in the WT are dysregulated. CDPK4 thus appears to be directly or indirectly responsible for the phosphorylation of many proteins, including those in the replisome and the microtubule cytoskeletal systems. Since CDPK4 is known to be activated in the microgametocyte soon after the XA signal is detected, this places this

kinase upstream in the pathway with a "master regulator" role in the propagation of the signal, as previously suggested [3]. SRPK1, on the other hand, plays a more indirect role, in which its presence is responsible for the repression of phosphorylation on proteins in the same systems. As a result, loss of SRPK1 leads to an increase of phosphorylation events after gametocyte activation. Because SRPK1 is itself negatively disrupted upon deletion of CDPK4, we hypothesise that SRPK1's activity depends on activation of CDPK4 and that it, in turn, negatively regulates another kinase or positively regulates a phosphatase. In at least one case that we identified, this could function as a feedback mechanism on a phosphosite that is regulated by CDPK4.

This dynamic view of phosphorylation during early gametogenesis reveals an unexpected breadth to the signalling response at short time-scales. A number of studies in malaria parasites have investigated phosphorylation of parasite proteins at proteome-scale, looking at different parasite stages and species [11, 20–27]. Even where these studies have used genetic mutations or small molecules to perturb the system in a controlled manner, identifying functional modules among the regulated proteins has been rendered difficult due to the large number of responding phosphosites in the proteome. Time-course data, in contrast, have the power to provide a more highly resolved analysis by grouping sites by the kinetics of their response [28–33]. Using this approach, we here bring into focus specific systems that are undergoing phosphoregulation as key players in the initiation of gametogenesis, especially in the rapidly developing male gametocytes. In particular, the multiple rounds of DNA replication and mitosis undergone by microgametocytes appear to be initiated via phosphorylation within the first seconds of activation. Conservation of phosphorylation sites in replication initiation proteins and motor proteins supports the hypothesis that these phosphorylation events functionally regulate the proteins. By linking individual protein kinases and phosphatases with downstream effector modules characterised by co-regulated phosphosites, we provide a rich resource from which starting points for future studies into the underlying mechanisms of cell cycle regulation during *Plasmodium* gamete formation can be derived. Some of the molecular players we identify are essential also in blood stages, and their further analysis in sexual development will therefore not only reveal new ways of blocking parasite transmission to the mosquitoes but may also serve as a model to characterise targets with important roles in asexual blood stages, which cause all the pathology associated with malaria in humans.

Cell-cycle regulation in malaria parasites is poorly understood and current evidence suggests that some of the canonical cell cycle check points are not present. During asexual schizogony in the blood stages, for example, nuclei behave as autonomous units that undergo repeated rounds of replication and mitosis in an asynchronous manner, while being retained within the same cellular envelope [56, 57]. During sexual development, the release of microgametes requires DNA synthesis [58], but on the other hand a

compound that interferes with spindle formation did not prevent DNA replication to proceed through all three rounds [5], suggesting in the male nucleus multiple rounds of replication and mitosis may progress in parallel without depending on each other. We were therefore intrigued to see that protein modules involved in origin or replication recognition and cytoskeletal reorganisation were simultaneously targeted for phosphorylation during the first few seconds of gametocyte activation. These data suggest that some mitotic processes are already initiated contemporaneously with DNA synthesis in *Plasmodium* gametocytes and raise important questions regarding the temporal control of the cell cycle in these organisms.

Materials & Methods

Parasite maintenance and preparation

All animal experiments were conducted under a license from the UK Home Office in accordance with national and European animal welfare guidelines or with the authorisation number (GE/82/15) according to the guidelines and regulations issued by the Swiss Federal Veterinary Office. *P. berghei* strain ANKA [59] derived clone 2.34 [3], CDPK4-KO [60], and SRPK1-KO [17] were maintained in CD1 outbred mice obtained from Harlan. Mice were specific pathogen-free and subjected to regular pathogen monitoring by sentinel screening. They were housed in individually ventilated cages furnished with a cardboard fun tunnel and Nestlet. Mice were maintained at $21 \pm 2\,^{\circ}\mathrm{C}$ under a 12 h light/dark cycle and given commercially prepared autoclaved dry rodent diet and water *ad libitum*. Mice were used for experimentation at 6–12 weeks of age.

For gametocyte production, mice were treated with phenyl hydrazine three days before infection. One day after infection asexually replicating parasites were eliminated by the addition of sulfadiazine (20 mg/L) in the drinking water. Parasites were harvested at day four after infection in suspended animation (SA - RPMI1640 medium containing25 mM HEPES, 5 % FCS, 4 mM sodium bicarbonate, pH 7.2) and separated from uninfected erythrocytes on a Histodenz cushion made up from 48 % of a Histodenz stock (27.6 % w/v Histodenz -Sigma- in 5.0 mM Tris-HCl [pH 7.2], 3.0 mM KCl, 0.3 mM EDTA) and 52 % SA with a final pH of 7.2. Purified gametocytes were washed twice in SA without FCS and resuspended in 400 µL of SA without FCS. Activation was induced by adding 400 µL of exflagellation medium (RPMI 1640 containing 25 mM HEPES, 4 mM sodium bicarbonate, 200 µM xanthurenic acid, pH 8). Parasites were snap frozen in liquid nitrogen at 6, 12, 18 and 60 s after activation. For the initial time point, the exflagellation

medium was replaced by SA. For each time point and parasite line, four and two independent biological replicates were produced for the time-course experiment and the KO experiments, respectively.

A lysis buffer was prepared consisting of 4 % SDS, 50 mM NaCl, 100 mM Tris buffer (pH 7.4), 5 mM EDTA, 40 mM TCEP, and Halt™ Protease & Phosphatase Inhibitor Cocktail (Thermo) (2 x). Samples were split in half (approximately 400 µL each) and 500 µL lysis buffer was added to each half. The samples were vortexed and then heated at 95 °C for 10 min. DNA was sheared via pulses of sonication for 20 s (1 s on, 1 s off) at 40 % power. The samples were centrifuged for 30 min at 14 000 rpm and the supernatant was collected. The samples were then incubated with 80 mM iodoacetamide for 1 h at room temperature. The protein digest used the FASP method [61] with Amicon Ultra-15 30K filter units (UFC903024), where Trypsin Gold (Thermo) was added at a 1:50 ratio and the digest was incubated at 37 °C overnight. The resulting peptides were desalted using Sep-Pak C18 Plus Light cartridges then dried in a SpeedVac and stored at −20 °C.

Quantitative protein mass spectrometry

Time course experiment

Phosphopeptide enrichment Phosphopeptide enrichment was performed on TiO_2 tips (Thermo), following the manufacturer's instructions. Each sample used 500 µg peptides. Phosphopeptides were eluted from the tip by 1.5 % NH_4OH followed by 5 % pyrrolidine. Both eluates were pooled, acidified and then desalted on Graphite Spin Columns (Thermo) as instructed by the manufacturer's protocol. Each sample was split into two technical replicates, dried in a SpeedVac, and stored at −20 °C.

Label-free LC-MS/MS The dried phosphopeptides were resuspended in 80 µL of 0.5 % FA / 100 % H_2O before LC-MS/MS analysis on an LTQ Orbitrap Velos coupled with an Ultimate 3000 RSLCnano System (both from Thermo Fisher). The peptides were first loaded and desalted on a PepMap C18 trap column (100 µm id x 20 mm, 5 µm) then separated on a PepMap C18 analytical column (75 µm id x 500 mm, 2 µm) (both from Thermo Fisher) over a 240 min linear gradient of 4–30 % CH_3CN/0.1 % formic acid with the total cycle time at 280 min. The Orbitrap mass spectrometer was operated in the standard "top 15" data-dependent acquisition mode while the preview mode was disabled. The MS full scan was set at m/z 380–1600 with the resolution at 30 000 at m/z 400 and a lock mass at m/z 445.120 025. The AGC was set at 1×10^6 with a maximum injection time at 200 ms. The 15 most abundant multiply-charged precursor ions, with minimal signals above 3000 counts, were dynamically

selected for CID fragmentation (MS/MS) in the ion trap, which had the AGC set at 5000 with the maximum injection time at 100 ms. The dynamic exclusion duration time was set for 60 s with ±10 ppm exclusion mass width. The isolation width was 2.0 Da and the normalised collision energy was 35 %. LC-MS/MS analyses of technical replicates were repeated sequentially.

Spectral analysis and peptide quantification Raw spectra from the time-course experiment were analysed using MaxQuant (version 1.5.2.8) [62]. Technical replicates were treated as fractions. The minimum peptide length was set at seven and protein and site identification false discovery rates (FDR) were set at 0.01. Carbamidomethyl was set as a fixed modification. Peptides were searched against the *P. berghei* annotated protein database retrieved from PlasmoDB (`http://www.plasmodb.org`; version 13.0) and the *M. musculus* protein sequence database retrieved from Uniprot (`http://www.uniprot.org`). Peptides were quantified using the iBAQ (intensity-based absolute quantification) method [63], as implemented in MaxQuant, using a logarithmic fit and spectra-matching between runs. Acetylation, deamidation, oxidation and phosphorylation were included as modifications in the protein quantification. Only phosphosites with a localisation probability of 0.75 or greater were retained for further analysis.

Data processing Time courses for the change in phosphorylation level for each phosphosite were calculated using the composite intensity scores estimated by MaxQuant from all peptide evidence for that site. For each time point (0, 6, 12, 18 and 60 s), these phosphosite intensities were quantile normalised between the four biological replicates [64]. For each site at each time point, the mean and standard error of the intensities were calculated if intensities were measured in at least two replicates, otherwise it was considered not to have been detected. The standard error of the site's intensities at a given time point was compared to the distribution of standard errors of all sites' intensities for that time point. If a site's standard error of intensities at any time point was greater than the distribution's upper quartile plus 1.5 times the interquartile range (IQR) (*i.e.* the upper "whisker" on a standard box-and-whisker plot), that site was discarded as unreliable.

Three time courses were then generated from these values: the "full" time course (0–60 s), for sites that were detected at all five time points; a "late" truncated time course (6–60 s), for sites that were not detected in unactivated parasites; and an "early" truncated time course (0–18 s), for sites that were not detected at the 60 s time point. For each site, ratios were calculated from its intensity at each time point against its intensity at the first time point and were logarithmically transformed (base 2) (*e.g.* $log_2\left(\frac{6\,\mathrm{s}}{0\,\mathrm{s}}\right)$). Only complete time courses, with no missing data, were retained for further analysis.

In order to determine which sites show significant evidence of change in phosphorylation state, we exploited the fact that enrichment of phosphopeptides is not 100 % specific and that the time-frame in question is too short to see significant variation in protein abundances. Time courses were constructed for non-phosphorylated peptides in the same manner as described above. All phosphosite ratios and non-phosphopeptide ratios were then corrected for the median non-phosphopeptide ratio for the corresponding time point. The distributions of non-phosphopeptide log_2 ratios were used as null distributions against which the significance of the phosphosite log_2 ratios could be tested. Two-tailed p-values were computed from an empirical cumulative distribution function of the non-phosphopeptide ratios. Sites showing at least one time point with a p-value less than the critical value of 0.05 were determined to have undergone significant phosphoregulation during the time course.

Significant sites were clustered according to their time courses using a normal mixture-modeling based method, using the "mclust" library (version 5.2) in R [65,66]. The library chooses the number of clusters and the specific model to fit (spherical with equal volumes or spherical with unequal volumes) via an automated method which optimises the Bayesian Information Criterion (BIC).

Knock-out experiments

Phosphopeptide enrichment The CDPK4KO and the SRPK1KO experiments consisted of 8 samples each: wild-type and knock-out samples measured at two time points, unactivated (0 s) and activated (18 s after activation), in two biological replicates.

200 µL of 2 x lysis buffer was added to the 200 µL parasite suspension, and then processed as described above. 300 µg proteins were taken and the volumes were equalised with 100 mM TEAB followed by alkylation with IAA. Proteins were precipitated by MTBE [67] and then digested with 8 µg trypsin in 150 µL of 100 mM TEAB at 37 °C for 2 h. A further 4 µg trypsin was added and digested for another 5 h. 100 µg peptides was taken from each replicate for TMT 10plex labelling, then mixed and dried in a SpeedVac. The labelled peptide mixture was fractionated at pH 10 on an XBridge BEH C18 column (4.6 mm id x 250 mm, 130 Å, 3.5 µm) (Waters) at a flow rate of 500 µL/min with a linear gradient from 5–35 % ACN/NH_3 in 30 min and total cycle time of 60 min. Fractions were collected every 30 s and between 2.5–50 min on a 96-well plate by rows and then concatenated into 12 fractions by columns and dried in a SpeedVac.

Enrichment of phosphopeptides was performed using IMAC with PHOS-Select Iron Affinity Gel (Sigma) then TiO_2 tips (Thermo Fisher) sequentially. All procedures followed the manufacturer's instruction

with some modification. The peptides were redissolved in 50 % ACN/0.1 % TFA then added to 100 µL of pre-washed PHOS-Select Iron Affinity Gel and left binding at room temperature with end-to-end rotation for 30 min. The beads were washed three times with 250 mM acetic acid/50 % ACN and once with H_2O. Phosphopeptides were eluted twice with 100 µL of 1.5 % NH_3/25 % ACN then dried in a SpeedVac. The flow-through and the first wash of IMAC beads were collected and dried in a SpeedVac, and then the phosphopeptides were enriched using TiO_2 tips as described above.

TMT-labelled LC-MS/MS analysis The enriched phosphopeptides were redissolved in 0.5 % FA before LC-MS/MS analysis on an Orbitrap Fusion Tribrid mass spectrometer coupled with an Ultimate 3000 RSLCnano system configured as above. The peptides were separated with a linear gradient of 4–36 % ACN/0.1 % FA in 120 min and total 153 min per cycle. The Orbitrap Fusion was operated using the "Top Speed" method with 2 s cycle time. The MS full scan was in the Orbitrap with the following settings: scan range at m/z 380–1500 with a lock mass at 445.120 025, resolution at 120 000 at m/z 200, and AGC at 4×10^5 with a maximum injection time at 50 ms. The multiply-charged precursor ions (2+ to 6+) at most intense, with a minimal signal above 10 000 counts, were dynamically selected for high energy collision-induced dissociation (HCD) (MS/MS) and detected in the Orbitrap with a resolution at 30 000 at m/z 200. The isolation width was 1.2 Da in quadrupole, and the collision energy was set at 40 %. The dynamic exclusion duration time was set for 60 s with ±10 ppm exclusion mass width, the AGC was set at 1×10^5 with the maximum injection time at 105 ms.

The phosphopeptides enriched via IMAC and TiO_2 were analysed separately.

Spectral analysis and peptide quantification Raw data were processed in Proteome Discoverer 2.1 (PD2.1; Thermo Fisher) using both SequestHT and Mascot search engines against a combined protein database of *P. berghei* and mouse as above. Trypsin maximum missed cleavage sites were set to 2. The dynamic modifications set in both Mascot and SequestHT were Acetyl (N-term), Deamidated (NQ), Phospho (STY) and Oxidation (M), while in SequestHT Camabidomethyl (C) was set as a fixed modification. The remaining settings were the same in both: precursor mass tolerance at 20 ppm, fragment at 0.5 Da, and TMT6plex as fixed modification. All files acquired from same set of TMT10plex experiments, including from both IMAC and TiO_2 purified samples, were set as fractions. The search result was validated by Percolate where the q-value was set at 0.01. The PSMs were identified, quantified and grouped to the peptides. Both PSMs and peptides were filtered with 1 % FDR, then further group to proteins with 1 % FDR where only peptides at high confidence were selected. Both unique and razor peptides were used for protein quantification, and protein and peptide abundance values were calculated

as summed PSM quan values (S/N values of reporter ions). The abundances were normalised on Total Peptide Amount, and then scaled with On Channels Average. The co-isolation threshold was set at 50 % to reduce the isolation interference. The phosphorylation sites were localised by phosphoRS as implemented in PD2.1 with site probability set at 75.

Data processing For each knock-out experiment, peptides were filtered to include those which were unambiguously matched to a single *P. berghei* or *M. musculus* protein with a FDR less than 0.01. Furthermore, only peptides that were quantified in all TMT channels were retained. log_2 ratios were calculated for both biological replicates using the peptides' scaled abundances provided by Proteome Discoverer for the following comparisons: KO 18 s vs. KO 0 s, WT 18 s vs. WT 0 s, KO 18 s vs. WT 18 s, and KO 0 s vs. WT 0 s. For each comparison in each biological replicate, the ratios were corrected by subtracting the median value. Finally, for each comparison, a mean ratio was calculated by first computing the mean ratio within each replicate for each of the peptide's PSMs (*i.e.* those PSMs with phosphorylation at the same site(s) but otherwise differing in other modifications), followed by computing the mean of the resulting value between the two biological replicates. For all ratios, a p-value was calculated using the empirical cumulative distribution function of the non-phosphopeptides, as described above.

In order to account for a general change in cellular state before gametocyte activation, we built a linear model to predict the disruption at 18 s from the disruption at 0 s. That is, we aimed to predict $log_2\left(\frac{\text{KO 18 s}}{\text{WT 18 s}}\right)$ from $log_2\left(\frac{\text{KO 0 s}}{\text{WT 0 s}}\right)$. We reasoned that peptides which were poorly predicted by this model could be assumed to have been affected by the knock-out during the activation process. To quantify this, we Studentised the residuals and assigned p-values from the t-distribution. Peptides that were found to have both residual p-values and $log_2\left(\frac{\text{KO 18 s}}{\text{WT 18 s}}\right)$ ratio p-values less than 0.05 were taken to be significantly affected by the knock-out. We also performed a further partitioning of the affected peptides to isolate those which suffered an apparent loss of phospho-regulation, by taking those which were: a) observed to be significantly up-regulated in the WT sample or in an up-regulation cluster in the time-course experiment; and b) either significantly down-regulated or not significantly regulated in the KO sample.

Protein function and sex enrichment

Gene Ontology (GO) association files for *P. berghei* and *P. falciparum* (version 1.3.2015) were fetched from the Wellcome Trust Sanger Institute FTP server (`ftp://ftp.sanger.ac.uk`). The *P. falciparum* GO associations were used to supplement the *P. berghei* ones. For each *P. falciparum* gene with a

one-to-one ortholog in *P. berghei*, the corresponding *P. falciparum* terms were merged with the existing terms associated with the *P. berghei* ortholog (if any).

Groups of proteins were tested for enrichment of GO terms via a binomial test [68]. Query protein groups for the time-course experiment included the list of all proteins containing significant sites and the lists of proteins containing significant sites in each of the phosphosite clusters. For the knock-out experiments, the query groups included proteins that showed significant up- or down-regulation in the wild-type sample, those whose post-activation state was significantly disrupted in the knock-out versus the wild-type, and those which we determined to be putative down-stream targets of the deleted kinase. The "background" set of proteins, against which enrichment was tested, consisted of all of the proteins detected in the time-course mass-spectrometry experiments, including those detected via non-phosphorylated peptides. GO enrichment analyses were performed using GOATOOLS (version 0.5.9) [69], modified to perform the binomial test, on the "go-basic" ontology file provided at `http://purl.obolibrary.org/obo/go/go-basic.obo`. For the knock-out experiments, we only performed targeted enrichment analyses by testing only the "leaf" terms found enriched in the time-course experiment.

Groups of proteins were further tested for enrichment for sex-specific proteins, using a previously published *P. falciparum* sex-partitioning data set by orthology [44]. Fisher's exact test was used to determine significant enrichment for female-specific, male-specific or shared proteins.

p-values for all enrichment analyses were corrected for FDR and significance was tested at a critical value of 0.05.

Network analysis

Protein interaction networks for *P. berghei* and *P. falciparum* were downloaded from the STRING database (version 10; `http://www.string-db.org`) [40]. The *P. berghei* network was extended to include additional nodes and edges from the *P. falciparum* network by orthology. New *P. berghei* nodes were added if they have a one-to-one ortholog in the *P. falciparum* proteome; edges from the *P. falciparum* STRING network were added to the *P. berghei* network only in the case that the edge was not already present; in the case that the STRING score of an edge between two orthologous pairs of proteins was different between the two networks, the original score from the *P. berghei* network was retained. Network statistics were computed using the NetworkX library for Python (version 1.8.1) [70].

Significant association on the network between protein kinases or phosphatases and the phosphosite time

course clusters was determined through an enrichment analyses for each enzyme that was found to be phosphoregulated during the time course or that has previously been implicated in gametogenesis (see Figure 4). We compared the median of all of a kinase's edge scores with the median of its edge scores shared with the proteins carrying phosphosites in a given time course cluster. We tested the hypothesis that the median edge score for the subset of proteins with sites in the cluster was higher than that of all the edges via the Mann-Whitney test. p-values were corrected for multiple testing by adjusting for FDR.

Site conservation

Site conservation was measured by fitting nucleotide alignments of orthologous genes with site-specific codon substitution models. An estimate of the ratio of the number of non-synonymous substitutions per site to the number of synonymous substitutions per site (dN/dS) during species divergence was used as a measurement of conservation; dN/dS values less than one indicate that the site has predominantly been under purifying selection, with lower values indicating stronger selective constraint and conservation.

A nucleotide multiple-sequence alignment was produced for the genes in each orthologous group using PRANK (version .140110) [71, 72]. Regions with gaps were automatically removed using TrimAL [73]. In order to estimate a species tree for the conservation analysis, a composite alignment was built by concatenating the 3150 alignments that contained orthologs from all seven species. The tree was estimated using PhyML [74]. Site dN/dS estimates were computed using the site models implemented in the "codeml" program of the PAML package (version 4.8a) [75]. The analyses were automated using the PAML interface implemented in the Bio.Phylo module of Biopython [76]. Subsequent analyses of site conservation included only sites under purifying selection ($dN/dS < 1$). A regression model of dN/dS scores was built using beta regression, as implemented in the "betareg" library for R (version 3.0-5) [77]. The residuals of this model were used to identify sites under stronger-than-expected conservation. For more details, see the Supplementary Methods.

Motor protein phosphorylation hot-spots

We identified and grouped proteins belonging to the kinesin, dynein heavy chain, dynein light intermediate chain and dynein intermediate chain families according to their annotations in the proteome databases for *P. berghei*, *P. falciparum*, *T. gondii*, *M. musculus*, *Homo sapiens*, *Rattus norvegicus* and *Saccharomyces cerevisiae*. In addition, we considered a group of *Plasmodium* proteins for which no known

function has been described (denoted as "conserved *Plasmodium* protein, unknown function"), but which have been assigned the "motor activity" GO term (GO:0003774) in *P. falciparum* at PlasmoDB, some of which we could assign by orthology to the kinesin family. Previously published human, mouse and rat phosphosite information was retrieved from Phosphosite Plus [78] (date of data set: 2016 Jan. 29). For yeast, phosphosite data was retrieved from Phosphogrid [79] (downloaded 2014 Mar. 14). Phosphosite data was manually curated for *T. gondii* [21, 80], *P. falciparum* [20–24, 26] and *P. berghei* [11]. The *P. berghei* was further supplemented with the phosphosites observed in our time course experiment. Multiple sequence alignments were constructed for each sequence and phosphosite hot-spots were estimated via kernel density estimation on phosphosite counts at each aligned column. For more details, see the Supplementary Methods.

Acknowledgements

Work at the Wellcome Trust Sanger Institute was funded by a core grant from the Wellcome Trust (WT098051). This work was supported by the Swiss National Science Foundation grant BSSGI0_155852 to MB. MB is an INSERM investigator. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE [81] partner repository with the dataset identifier PXD006266.

References

[1] World Health Organization. *World Malaria Report 2015* (World Health Organization, 2015).

[2] Billker, O., Shaw, M. K., Margos, G. & Sinden, R. E. The roles of temperature, pH and mosquito factors as triggers of male and female gametogenesis of *Plasmodium berghei in vitro*. *Parasitology* **115–Pt–1**, 1–7 (1997). URL `http://www.ncbi.nlm.nih.gov/pubmed/9280891`.

[3] Billker, O. *et al.* Calcium and a calcium-dependent protein kinase regulate gamete formation and mosquito transmission in a malaria parasite. *Cell* **117**, 503–514 (2004). URL `http://www.ncbi.nlm.nih.gov/pubmed/15137935`.

[4] Sinden, R., Canning, E. U. & Spain, B. Gametogenesis and fertilization in *Plasmodium yoelii nigeriensis*: a transmission electron microscope study. *Proceedings of the Royal Society of London B: Biological Sciences* **193**, 55–76 (1976).

[5] Billker, O. *et al.* Azadirachtin disrupts formation of organised microtubule arrays during microgametogenesis of *Plasmodium berghei*. *J. Eukaryot. Microbiol.* **49**, 489–497 (2002). URL http://www.ncbi.nlm.nih.gov/pubmed/12503686.

[6] Ojo, K. K. *et al.* Transmission of malaria to mosquitoes blocked by bumped kinase inhibitors. *J. Clin. Invest.* **122**, 2301–2305 (2012). URL http://www.ncbi.nlm.nih.gov/pubmed/17521970.

[7] Kawamoto, F., Alejo-Blanco, R., Fleck, S. L., Kawamoto, Y. & Sinden, R. E. Possible roles of Ca^{2+} and cGMP as mediators of the exflagellation of *Plasmodium berghei* and *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* **42**, 101–108 (1990). URL http://www.ncbi.nlm.nih.gov/pubmed/2172816.

[8] Muhia, D. K., Swales, C. A., Deng, W., Kelly, J. M. & Baker, D. A. The gametocyte-activating factor xanthurenic acid stimulates an increase in membrane-associated guanylyl cyclase activity in the human malaria parasite *Plasmodium falciparum*. *Molecular Microbiology* **42**, 553–560 (2001). URL http://www.ncbi.nlm.nih.gov/pubmed/11703675.

[9] McRobert, L. *et al.* Gametogenesis in malaria parasites is mediated by the cGMP-dependent protein kinase. *PLoS Biol.* **6**, e139 (2008). URL http://www.ncbi.nlm.nih.gov/pubmed/10747978.

[10] Raabe, A. C., Wengelnik, K., Billker, O. & Vial, H. J. Multiple roles for *Plasmodium berghei* phosphoinositide-specific phospholipase C in regulating gametocyte activation and differentiation. *Cell. Microbiol.* **13**, 955–966 (2011). URL http://www.ncbi.nlm.nih.gov/pubmed/11413485.

[11] Brochet, M. *et al.* Phosphoinositide Metabolism Links cGMP-Dependent Protein Kinase G to Essential Ca^{2+} Signals at Key Decision Points in the Life Cycle of Malaria Parasites. *PLoS Biol.* **12**, e1001806 (2014). URL http://www.ncbi.nlm.nih.gov/pubmed/24284631.

[12] Philip, N. & Waters, A. P. Conditional Degradation of *Plasmodium* Calcineurin Reveals Functions in Parasite Colonization of both Host and Vector. *Cell Host Microbe* **18**, 122–131 (2015). URL http://www.ncbi.nlm.nih.gov/pubmed/19915560.

[13] Sebastian, S. *et al.* A *Plasmodium* calcium-dependent protein kinase controls zygote development and transmission by translationally activating repressed mRNAs. *Cell Host Microbe* **12**, 9–19 (2012). URL http://www.ncbi.nlm.nih.gov/pubmed/22817981.

[14] Tewari, R., Dorin, D., Moon, R., Doerig, C. & Billker, O. An atypical mitogen-activated protein kinase controls cytokinesis and flagellar motility during male gamete formation in a malaria para-

site. *Molecular Microbiology* **58**, 1253–1263 (2005). URL `http://www.ncbi.nlm.nih.gov/pubmed/16313614`.

[15] Rangarajan, R. *et al.* A mitogen-activated protein kinase regulates male gametogenesis and transmission of the malaria parasite *Plasmodium berghei*. *EMBO Rep.* **6**, 464–469 (2005). URL `http://www.ncbi.nlm.nih.gov/pubmed/10699250`.

[16] Guttery, D. S. *et al.* A putative homologue of CDC20/CDH1 in the malaria parasite is essential for male gamete development. *PLoS Pathog.* **8**, e1002554 (2012). URL `http://www.ncbi.nlm.nih.gov/pubmed/16896351`.

[17] Tewari, R. *et al.* The systematic functional analysis of *Plasmodium* protein kinases identifies essential regulators of mosquito transmission. *Cell Host Microbe* **8**, 377–387 (2010). URL `http://www.ncbi.nlm.nih.gov/pubmed/19723765`.

[18] Kern, S. *et al.* Inhibition of the SR protein-phosphorylating CLK kinases of *Plasmodium falciparum* impairs blood stage replication and malaria transmission. *PLoS ONE* **9**, e105732 (2014). URL `http://www.ncbi.nlm.nih.gov/pubmed/8743695`.

[19] Guttery, D. S. *et al.* Genome-wide functional analysis of *Plasmodium* protein phosphatases reveals key regulators of parasite development and differentiation. *Cell Host Microbe* **16**, 128–140 (2014). URL `http://www.ncbi.nlm.nih.gov/pubmed/16313614`.

[20] Solyakov, L. *et al.* Global kinomic and phospho-proteomic analyses of the human malaria parasite *Plasmodium falciparum*. *Nat Commun* **2**, 565 (2011). URL `http://www.ncbi.nlm.nih.gov/pubmed/22127061`.

[21] Treeck, M., Sanders, J. L., Elias, J. E. & Boothroyd, J. C. The phosphoproteomes of *Plasmodium falciparum* and *Toxoplasma gondii* reveal unusual adaptations within and beyond the parasites' boundaries. *Cell Host Microbe* **10**, 410–419 (2011). URL `http://www.ncbi.nlm.nih.gov/pubmed/20951971`.

[22] Lasonder, E. *et al.* The *Plasmodium falciparum* schizont phosphoproteome reveals extensive phosphatidylinositol and cAMP-protein kinase A signaling. *Journal of Proteome Research* **11**, 5323–5337 (2012). URL `http://www.ncbi.nlm.nih.gov/pubmed/23025827`.

[23] Pease, B. N. *et al.* Global analysis of protein expression and phosphorylation of three stages of *Plasmodium falciparum* intraerythrocytic development. *Journal of Proteome Research* **12**, 4028–

4045 (2013). URL http://www.ncbi.nlm.nih.gov/pubmed/23914800.

[24] Collins, M. O., Wright, J. C., Jones, M., Rayner, J. C. & Choudhary, J. S. Confident and sensitive phosphoproteomics using combinations of collision induced dissociation and electron transfer dissociation. *J Proteomics* (2014). URL http://www.ncbi.nlm.nih.gov/pubmed/24657495.

[25] Alam, M. M. *et al.* Phosphoproteomics reveals malaria parasite Protein Kinase G as a signalling hub regulating egress and invasion. *Nat Commun* **6**, 7285 (2015). URL http://www.ncbi.nlm.nih.gov/pubmed/22569589.

[26] Lasonder, E., Green, J. L., Grainger, M., Langsley, G. & Holder, A. A. Extensive differential protein phosphorylation as intraerythrocytic *Plasmodium falciparum* schizonts develop into extracellular invasive merozoites. *Proteomics* **15**, 2716–2729 (2015). URL http://www.ncbi.nlm.nih.gov/pubmed/25886026.

[27] Ganter, M. *et al.* *Plasmodium falciparum* CRK4 directs continuous rounds of DNA replication during schizogony. *Nature Microbiology* **2**, 17017 (2017).

[28] Blagoev, B., Ong, S.-E., Kratchmarova, I. & Mann, M. Temporal analysis of phosphotyrosine-dependent signaling networks by quantitative proteomics. *Nature Biotechnology* **22**, 1139–1145 (2004). URL http://www.ncbi.nlm.nih.gov/pubmed/15340474.

[29] Olsen, J. V. *et al.* Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell* **127**, 635–648 (2006). URL http://www.ncbi.nlm.nih.gov/pubmed/17081983.

[30] Oyama, M. *et al.* Temporal perturbation of tyrosine phosphoproteome dynamics reveals the system-wide regulatory networks. *Mol. Cell Proteomics* **8**, 226–231 (2009). URL http://www.ncbi.nlm.nih.gov/pubmed/18815124.

[31] de Graaf, E. L., Giansanti, P., Altelaar, A. F. M. & Heck, A. J. R. Single-step enrichment by Ti4+-IMAC and label-free quantitation enables in-depth monitoring of phosphorylation dynamics with high reproducibility and temporal resolution. *Mol. Cell Proteomics* **13**, 2426–2434 (2014). URL http://www.ncbi.nlm.nih.gov/pubmed/24850871.

[32] Kanshin, E., Kubiniok, P., Thattikota, Y., D'Amours, D. & Thibault, P. Phosphoproteome dynamics of Saccharomyces cerevisiae under heat shock and cold stress. *Mol. Syst. Biol.* **11**, 813 (2015). URL http://www.ncbi.nlm.nih.gov/pubmed/21953191.

[33] Kanshin, E., Bergeron-Sandoval, L.-P., Isik, S. S., Thibault, P. & Michnick, S. W. A cell-signaling network temporally resolves specific versus promiscuous phosphorylation. *Cell Rep* **10**, 1202–1214 (2015). URL `http://www.ncbi.nlm.nih.gov/pubmed/25704821`.

[34] Alonso-Morales, A. *et al.* Protein phosphorylation during *Plasmodium berghei* gametogenesis. *Exp. Parasitol.* **156**, 49–60 (2015). URL `http://www.ncbi.nlm.nih.gov/pubmed/18567578`.

[35] Raabe, A. C., Billker, O., Vial, H. J. & Wengelnik, K. Quantitative assessment of DNA replication to monitor microgametogenesis in *Plasmodium berghei*. *Mol. Biochem. Parasitol.* **168**, 172–176 (2009). URL `http://www.ncbi.nlm.nih.gov/pubmed/11655723`.

[36] Teixidó-Travesa, N., Roig, J. & Lüders, J. The where, when and how of microtubule nucleation - one ring to rule them all. *Journal of Cell Science* **125**, 4445–4456 (2012). URL `http://www.ncbi.nlm.nih.gov/pubmed/23132930`.

[37] Talman, A. M. *et al.* Proteomic analysis of the *Plasmodium* male gamete reveals the key role for glycolysis in flagellar motility. *Malaria Journal* **13**, 315 (2014).

[38] McNamara, C. W. *et al.* Targeting *Plasmodium* PI (4) K to eliminate malaria. *Nature* **504**, 248–253 (2013).

[39] Leber, W. *et al.* A unique phosphatidylinositol 4-phosphate 5-kinase is activated by ADP-ribosylation factor in *Plasmodium falciparum*. *Int. J. Parasitol.* **39**, 645–653 (2009). URL `http://www.ncbi.nlm.nih.gov/pubmed/19171150`.

[40] Szklarczyk, D. *et al.* STRING v10: Protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Research* **43**, D447–52 (2015). URL `http://www.ncbi.nlm.nih.gov/pubmed/15461798`.

[41] Carvalho, T. G., Doerig, C. & Reininger, L. Nima- and Aurora-related kinases of malaria parasites. *Biochimica et Biophysica Acta* **1834**, 1336–1345 (2013). URL `http://www.ncbi.nlm.nih.gov/pubmed/23462523`.

[42] Khan, S. M. *et al.* Proteome analysis of separated male and female gametocytes reveals novel sex-specific *Plasmodium* biology. *Cell* **121**, 675–687 (2005). URL `http://www.ncbi.nlm.nih.gov/pubmed/15935749`.

[43] Dorin-Semblat, D. *et al.* *Plasmodium falciparum* NIMA-related kinase Pfnek-1: Sex specificity and assessment of essentiality for the erythrocytic asexual cycle. *Microbiology* **157**, 2785–2794 (2011). URL `http://www.ncbi.nlm.nih.gov/pubmed/15612927`.

[44] Tao, D. *et al.* Sex-partitioning of the *Plasmodium falciparum* stage V gametocyte proteome provides insight into falciparum-specific cell biology. *Mol. Cell Proteomics* **13**, 2705–2724 (2014). URL `http://www.ncbi.nlm.nih.gov/pubmed/17269722`.

[45] Lye, Y. M., Chan, M. & Sim, T.-S. Pfnek3: An atypical activator of a MAP kinase in *Plasmodium falciparum*. *FEBS Letters* **580**, 6083–6092 (2006). URL `http://www.ncbi.nlm.nih.gov/pubmed/17064692`.

[46] Ahmed, A. *et al.* Novel insights into the regulation of malarial calcium-dependent protein kinase 1. *FASEB J.* **26**, 3212–3221 (2012). URL `http://www.ncbi.nlm.nih.gov/pubmed/21287613`.

[47] Wei, Q., Li, J., Liu, T., Tong, X. & Ye, X. Phosphorylation of minichromosome maintenance protein 7 (MCM7) by cyclin/cyclin-dependent kinase affects its function in cell cycle regulation. *Journal of Biological Chemistry* **288**, 19715–19725 (2013). URL `http://www.ncbi.nlm.nih.gov/pubmed/10872463`.

[48] Han, X. *et al.* Phosphorylation of Minichromosome Maintenance 3 (MCM3) by Checkpoint Kinase 1 (Chk1) Negatively Regulates DNA Replication and Checkpoint Activation. *Journal of Biological Chemistry* **290**, 12370–12378 (2015). URL `http://www.ncbi.nlm.nih.gov/pubmed/23613359`.

[49] Beltrao, P. *et al.* Systematic functional prioritization of protein posttranslational modifications. *Cell* **150**, 413–425 (2012). URL `http://www.ncbi.nlm.nih.gov/pubmed/20461098`.

[50] Vaughan, P. S., Leszyk, J. D. & Vaughan, K. T. Cytoplasmic dynein intermediate chain phosphorylation regulates binding to dynactin. *Journal of Biological Chemistry* **276**, 26171–26179 (2001). URL `http://www.ncbi.nlm.nih.gov/pubmed/11340075`.

[51] Towns, W. L., Tauhata, S. B. F., Vaughan, P. S. & Vaughan, K. T. Transfection-induced defects in dynein-driven transport: Evidence that ICs mediate cargo-binding. *Cell Motil. Cytoskeleton* **66**, 80–89 (2009). URL `http://www.ncbi.nlm.nih.gov/pubmed/10545494`.

[52] Bader, J. R. *et al.* Polo-like kinase1 is required for recruitment of dynein to kinetochores during mitosis. *Journal of Biological Chemistry* **286**, 20769–20777 (2011). URL `http://www.ncbi.nlm.nih.gov/pubmed/11278950`.

[53] Rosse, C. *et al.* Binding of dynein intermediate chain 2 to paxillin controls focal adhesion dynamics and migration. *Journal of Cell Science* **125**, 3733–3738 (2012). URL `http://www.ncbi.nlm.nih.gov/pubmed/22553211`.

[54] Blasier, K. R. *et al.* Live cell imaging reveals differential modifications to cytoplasmic dynein properties by phospho- and dephosphomimic mutations of the intermediate chain 2C S84. *Journal of Neuroscience Research* **92**, 1143–1154 (2014). URL `http://www.ncbi.nlm.nih.gov/pubmed/20657784`.

[55] Young, S. A. M. *et al.* CRISPR/Cas9-Mediated Rapid Generation of Multiple Mouse Lines Identified Ccdc63 as Essential for Spermiogenesis. *Int J Mol Sci* **16**, 24732–24750 (2015). URL `http://www.ncbi.nlm.nih.gov/pubmed/19410201`.

[56] Read, M., Sherwin, T., Holloway, S., Gull, K. & Hyde, J. Microtubular organization visualized by immunofluorescence microscopy during erythrocytic schizogony in *Plasmodium falciparum* and investigation of post-translational modifications of parasite tubulin. *Parasitology* **106**, 223–232 (1993).

[57] Arnot, D. E., Ronander, E. & Bengtsson, D. C. The progression of the intra-erythrocytic cell cycle of *Plasmodium falciparum* and the role of the centriolar plaques in asynchronous mitotic division during schizogony. *International journal for parasitology* **41**, 71–80 (2011).

[58] Janse, C. J., van der Klooster, P. F., van der Kaay, H. J., van der Ploeg, M. & Overdulve, J. P. DNA synthesis in *Plasmodium berghei* during asexual and sexual development. *Molecular and biochemical parasitology* **20**, 173–182 (1986).

[59] Vincke, I. H., Bafort, J. & Scheepers-Biva, M. Recent observations on the cyclic transmission of *Plasmodium berghei*. *Ann Soc Belges Med Trop Parasitol Mycol* **46**, 327–336 (1966). URL `http://www.ncbi.nlm.nih.gov/pubmed/5955269`.

[60] Fang, H. *et al.* Multiple short windows of CDPK4 activity regulate distinct cell cycle events during *Plasmodium* gametogenesis. *Elife* (in press).

[61] Wiśniewski, J. R., Zougman, A., Nagaraj, N. & Mann, M. Universal sample preparation method for proteome analysis. *Nat. Methods* **6**, 359–362 (2009). URL `http://www.ncbi.nlm.nih.gov/pubmed/19876013`.

[62] Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.P.B.-Range mass accuracies and proteome-wide protein quantification. *Nature Biotechnology* **26**, 1367–1372 (2008). URL `http://www.ncbi.nlm.nih.gov/pubmed/19029910`.

[63] Schwanhäusser, B. *et al.* Global quantification of mammalian gene expression control. *Nature* **473**, 337–342 (2011). URL `http://www.ncbi.nlm.nih.gov/pubmed/23407496`.

[64] Bolstad, B. M., Irizarry, R. A., Astrand, M. & Speed, T. P. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* **19**, 185–193 (2003). URL `http://www.ncbi.nlm.nih.gov/pubmed/12538238`.

[65] Fraley, C. & Raftery, A. E. Model-based clustering, discriminant analysis, and density estimation. *Journal of the American statistical Association* **97**, 611–631 (2002).

[66] Fraley, C., Raftery, A. E. & Scrucca, L. mclust version 4 for r: Normal mixture modeling for model-based clustering, classification, and density estimation. *Technical Report No. 597, Department of Statistics, University of Washington.* (2012).

[67] Matyash, V., Liebisch, G., Kurzchalia, T. V., Shevchenko, A. & Schwudke, D. Lipid extraction by methyl-tert-butyl ether for high-throughput lipidomics. *Journal of Lipid Research* **49**, 1137–1146 (2008). URL `http://www.ncbi.nlm.nih.gov/pubmed/11518606`.

[68] Mi, H., Muruganujan, A., Casagrande, J. T. & Thomas, P. D. Large-scale gene function analysis with the PANTHER classification system. *Nature Protocols* **8**, 1551–1566 (2013). URL `http://www.pantherdb.org/)`.

[69] Tang, H. *et al.* Goatools: Tools for gene ontology (2015). URL `http://dx.doi.org/10.5281/zenodo.31628`.

[70] Hagberg, A. A., Schult, D. A. & Swart, P. J. Exploring network structure, dynamics, and function using NetworkX. In *Proceedings of the 7th Python in Science Conference*, SciPy, 11–15 (Los Alamos National Laboratory (LANL), 2008). URL `http://www.osti.gov/energycitations/product.biblio.jsp?osti\_id=960616`.

[71] Löytynoja, A. & Goldman, N. An algorithm for progressive multiple alignment of sequences with insertions. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 10557–10562 (2005). URL `http://www.ncbi.nlm.nih.gov/pubmed/16027352`.

[72] Löytynoja, A. & Goldman, N. Phylogeny-aware gap placement prevents errors in sequence alignment and evolutionary analysis. *Science* **320**, 1632–1635 (2008). URL `http://www.ncbi.nlm.nih.gov/pubmed/18566285`.

[73] Capella-Gutiérrez, S., Silla-Martínez, J. M. & Gabaldón, T. TrimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972–1973 (2009). URL `http://www.ncbi.nlm.nih.gov/pubmed/17452346`.

[74] Guindon, S. *et al.* New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst. Biol.* **59**, 307–321 (2010). URL `http://www.ncbi.nlm.nih.gov/pubmed/20525638`.

[75] Yang, Z. PAML 4: phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution* **24**, 1586–1591 (2007). URL `http://www.ncbi.nlm.nih.gov/pubmed/17483113`.

[76] Talevich, E., Invergo, B. M., Cock, P. J. & Chapman, B. A. Bio.Phylo: A unified toolkit for processing, analyzing and visualizing phylogenetic trees in Biopython. *BMC Bioinformatics* **13**, 209 (2012). URL `http://www.ncbi.nlm.nih.gov/pubmed/22909249`.

[77] Cribari-Neto, F. & Zeileis, A. Beta regression in R. *Journal of Statistical Software* **34**, 1–24 (2010). URL `http://www.jstatsoft.org/v34/i02/`.

[78] Hornbeck, P. V. *et al.* PhosphoSitePlus, 2014: Mutations, PTMs and recalibrations. *Nucleic Acids Research* **43**, D512–20 (2015). URL `http://www.ncbi.nlm.nih.gov/pubmed/24833420`.

[79] Sadowski, I. *et al.* The PhosphoGRID Saccharomyces cerevisiae protein phosphorylation site database: Version 2.0 Update. *Database (Oxford)* **2013**, bat026 (2013). URL `http://www.ncbi.nlm.nih.gov/pubmed/21460632`.

[80] Treeck, M. *et al.* The calcium-dependent protein kinase 3 of toxoplasma influences basal calcium levels and functions beyond egress as revealed by quantitative phosphoproteome analysis. *PLoS Pathog.* **10**, e1004197 (2014). URL `http://www.ncbi.nlm.nih.gov/pubmed/23209419`.

[81] Vizcaíno, J. A. *et al.* 2016 update of the PRIDE database and its related tools. *Nucleic acids research* **44**, D447–D456 (2016).

Tables

Protein ID	**Description**	**Site**	**Domain**	dN/dS
PBANKA_020300	chromatin assembly factor 1 protein WD40 domain	S50	CAF1C H4-bd (PF12265.5)	0.011
PBANKA_031420	CDPK1	S65	Protein kinase domain (PF00069.20)	0.05
PBANKA_041680	ubiquitin specific protease	S1076 / S1077	Ubiquitin carboxyl-terminal hydrolase (PF00443.24)	0.012 / 0.049
PBANKA_051190	RPL3	S13	Ribosomal L3 (PF00297.19)	0.013
PBANKA_061090	HSP40, subfamily A	S181	DnaJ central domain (PF00684.14)	0.008
PBANKA_061520	CDPK4	S80	Protein kinase domain (PF00069.20)	0.003
PBANKA_071590	USP13	Y410	UCH (PF00443.26)	0.02
PBANKA_080310	MCM7	T154	MCM N-terminal domain (PF14551.1)	0.006
PBANKA_081570	transporter	S251	Major Facilitator Superfamily (PF07690.11)	0.024
PBANKA_111190	MAT1	T94	MAT1 (PF06391.10)	0.015
PBANKA_112250	conserved *Plasmodium* protein, unknown function	Y108	SF1-HH (PF16275.2)	0.018
PBANKA_124180	MCM3	S626	MCM (PF00493.20)	0.007

Table 1: Significantly regulated sites which show stronger conservation than expected via a regression model, lie within a predicted protein domain, and have been previously observed in *P. falciparum*.