



**HAL**  
open science

## **PfGBP2 is a novel G-quadruplex binding protein in Plasmodium falciparum**

Pratima Gurung, Ana Rita Gomes, Rafael Martins, Stefan Juranek, Patrizia Alberti, Diane-Ethna Mbang-Benet, S Urbach, Elodie Gazanion, Vincent Guitard, Katrin Paeschke, et al.

► **To cite this version:**

Pratima Gurung, Ana Rita Gomes, Rafael Martins, Stefan Juranek, Patrizia Alberti, et al.. PfGBP2 is a novel G-quadruplex binding protein in Plasmodium falciparum. Cellular Microbiology, 2021, 23 (4), pp.e13303. 10.1111/cmi.13303 . hal-03152279

**HAL Id: hal-03152279**

**<https://hal.umontpellier.fr/hal-03152279>**

Submitted on 6 Oct 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**PfGBP2 is a novel G-quadruplex binding protein in  
*Plasmodium falciparum***

Journal:	<i>Cellular Microbiology</i>
Manuscript ID	CMI-20-0187
Manuscript Type:	Research article
Date Submitted by the Author:	10-Jul-2020
Complete List of Authors:	Gurung, Pratima; University of Montpellier, Laboratory of Pathogen-Host Interactions (LPHI) Gomes, Ana Rita; University of Montpellier, Laboratory of Pathogen-Host Interactions (LPHI) Martins, Rafael; University of Montpellier, Laboratory of Pathogen-Host Interactions (LPHI) Juranek, Stefan; University Hospital Bonn, Department of Oncology, Hematology and Rheumatology Alberti, Patrizia; Muséum National d'Histoire Naturelle, Nucleic Acid Structures, Telomeres and Evolution Mbang-Benet, Diane-Ethna ; University of Montpellier, Laboratory of Pathogen-Host Interactions (LPHI) Urbach, Serge; BioCampus Montpellier, CNRS UMR 5203, IGF,34090 Gazanion, Elodie; MIVEGEC, Laboratory of Parasitology and Mycology, University of Montpellier, CNRS UMR 5290 Guitard, Vincent; University of Montpellier, Laboratory of Pathogen-Host Interactions (LPHI) Paeschke, Katrin; University Hospital Bonn, Department of Oncology, Hematology and Rheumatology Lopez-Rubio, Jose-Juan; University of Montpellier, Laboratory of Pathogen-Host Interactions (LPHI)
Key Words:	Antigenic variations, Diseases, Genetic, Molecular genetic, Protozoa

PfGBP2 is a novel G-quadruplex binding protein in *Plasmodium falciparum*

Authors: Pratima Gurung<sup>a,b</sup>, Ana Rita Gomes<sup>a,b</sup>, Rafael M. Martins<sup>a</sup>, Stefan A. Juranek<sup>c</sup>,  
Patrizia Alberti<sup>d</sup>, Diane-Ethna Mbang-Benet<sup>a,b</sup>, Serge Urbach<sup>e</sup>, Elodie Gazanion<sup>b</sup>, Vincent  
Guitard<sup>a,b</sup>, Katrin Paeschke<sup>c</sup>, Jose-Juan Lopez-Rubio<sup>a,b\*</sup>

<sup>a</sup> *LPHI, UMR 5235, CNRS, University of Montpellier, 2 Place E Bataillon-Bat 24 cc107, 34095 Montpellier, France.*

<sup>b</sup> *Laboratory of Parasitology and Mycology, University of Montpellier, CNRS UMR 5290 MIVEGEC, 39 Avenue Charles Flahault, 34295 Montpellier, France*

<sup>c</sup> *Department of Oncology, Hematology and Rheumatology, University Hospital Bonn, 53127, Bonn, Germany*

<sup>d</sup> *Nucleic Acid Structures, Telomeres and Evolution, MNHN, CNRS UMR 7196, INSERM U1154, Sorbonne University, 43 rue Cuvier, 75231 Paris, France*

<sup>e</sup> *BioCampus Montpellier, CNRS UMR 5203, IGF, 34090 Montpellier, France.*

\* *e-mail: jose-juan.lopez-rubio@inserm.fr*

## Abstract

Guanine-quadruplexes (G4s) are non-canonical DNA structures that can regulate key biological processes such as transcription, replication and telomere maintenance in several organisms including eukaryotes, prokaryotes, and viruses. Recent reports have identified the presence of G4s within the AT-rich genome of *Plasmodium falciparum*, the protozoan parasite causing malaria. In *Plasmodium*, potential G4-forming sequence (G4FS) are enriched in the telomeric and sub-telomeric regions of the genome where they are associated with telomere maintenance and recombination events within virulence genes. However, there is a

1  
2  
3 little understanding about the biological role of G4s and G4-binding proteins. Here, we  
4 provide the first snapshot of G4-interactome in *P. falciparum* using DNA pull-down assay  
5 followed with LC-MS/MS. Interestingly, we identified ~29 potential G4-binding proteins  
6 (G4-BP) that binds to a stable G4FS (AP2\_G4). Furthermore, we characterized the role of G-  
7 strand binding protein 2 (PfGBP2), a putative telomere binding protein in *P. falciparum*. We  
8 validated the interaction of PfGBP2 with G4 *in vitro* as well as *in vivo*. PfGBP2 is expressed  
9 throughout the intra-erythrocytic developmental cycle and is essential for the parasites in the  
10 presence of G4-stabilizing ligand, Pyridostatin. Gene knockout studies showed the role of  
11 PfGBP2 in expression of *var* genes. Taken together, this study suggests that PfGBP2 is  
12 a *bona fide* G4-binding protein, which are likely to be involved in the regulation of G4-  
13 related functions in these malarial parasites. In addition, this study sheds light on this  
14 understudied G4 biology in *P. falciparum*.

## 35 Introduction

36  
37  
38  
39 G-quadruplexes (G4s) are non-canonical secondary nucleic acid structures that can  
40 form in guanine-rich nucleic acid sequences. These structures act as enhancers,  
41 repressors, and/or as blockades in multiple biological processes (Rhodes and Lipps,  
42 2015). In *Plasmodium falciparum*, an etiological agent of malaria, several putative  
43 G4-forming sequences (G4FS) are identified within the key regulatory regions of this  
44 AT-biased parasite genome. Moreover, these G4FS are highly enriched in telomeric  
45 and sub-telomeric regions including virulence gene families, particularly *var*  
46 genes (Smargiasso *et al.*, 2009; Calvo and Wasserman, 2016; Stanton *et al.*, 2016;  
47 Harris *et al.*, 2018; Marsico *et al.*, 2019; Gazanion *et al.*, 2020). The *var* gene family  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 encodes for erythrocyte membrane protein 1 (PfEMP1), which is an exported surface  
4 protein on infected erythrocytes. PfEMP1 plays a crucial role in malaria pathogenesis  
5 through cytoadherence and antigenic variation (Scherf *et al.*, 2008). The G4 are  
6 associated with telomeric maintenance and mitotic recombination events amongst *var*  
7 genes of these parasites (De Cian *et al.*, 2008; Calvo and Wasserman, 2016; Stanton *et*  
8 *al.*, 2016; Gazanion *et al.*, 2020). Several G4 ligands such as naphthalene di-imide,  
9 TMPyP4 and telomestatin affect the parasite growth, thus targeting G4-associated  
10 processes might represent new strategies for malaria intervention (Calvo and  
11 Wasserman, 2016; Anas *et al.*, 2017; Belmonte-Reche *et al.*, 2018; Harris *et al.*,  
12 2018).

13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28 Over the past decade, an increasing number of studies reported the identification of G4  
29 binding proteins (G4-BPs) and their role in modulating the functions of G4 in  
30 numerous organisms including yeast and mammalian cells (Sissi *et al.*, 2011; Wang *et*  
31 *al.*, 2012; Brázda *et al.*, 2014; Von Hacht *et al.*, 2014; Mcrae *et al.*, 2017; Brázda *et*  
32 *al.*, 2018). Most of the G4-BPs were identified by homology-based search and/or  
33 DNA pull-down based approaches (González *et al.*, 2009),(De Silanes *et al.*, 2010;  
34 Martadinata and Phan, 2013). For instance, the affinity-based approach revealed that  
35 Nucleolin stabilizes DNA G4 in the promoter of c-myc proto-oncogene and reduces  
36 the promoter activity (González *et al.*, 2009). Some RNA binding proteins including  
37 hnRNPs are shown to interact with telomeric repeat-containing RNA (TERRA) that  
38 forms G4 (De Silanes *et al.*, 2010; Martadinata and Phan, 2013). However, despite the  
39 availability of vast information on G4-BPs in different organisms, only RecQ helicase  
40 (particularly PfWRN) has been reported to influence G4-related phenotypes such as a  
41 change in *var* gene recombination patterns and high sensitivity against G4 ligands in  
42 *P. falciparum* (Stanton *et al.*, 2016; Claessens *et al.*, 2018).

1  
2  
3 Hence, to identify a putative G4-BPs in *P. falciparum*, we performed a DNA pull-down assay  
4 followed by LC-MS/MS using a validated G4 (AP2\_G4) as a bait. We next characterized  
5  
6 PfGBP2, one of the potential G4-BPs obtained in our study. A previous study showed that  
7  
8 PfGBP2 contains two RNA recognition motifs (RRM) and binds telomeric DNA *in vitro*  
9  
10 PfGBP2 contains two RNA recognition motifs (RRM) and binds telomeric DNA *in vitro*  
11  
12 (Calvo and Wasserman, 2015). We showed that PfGBP2 binds to the G4 formed by the  
13  
14 AP2\_G4 motif *in vitro* and it also interacts with the G4FS of *P. falciparum*. Thus, suggesting  
15  
16 that PfGBP2 is a *bona fide* G4-BP in *P. falciparum*. The ChIP-Seq data also confirmed that  
17  
18 PfGBP2 binds to the telomeres *in vivo*. Besides, we showed that the PfGBP2 is essential for  
19  
20 the intra-erythrocytic cycle of the malarial parasites in the presence of G4-stabilizing ligand,  
21  
22 Pyridostatin. The loss of PfGBP2 does not affect the telomere length up to 30 generations of  
23  
24 culturing while it showed modest derepression of several *var* gene expression.  
25  
26  
27  
28  
29  
30  
31

## 32 Results

### 33 Biophysical characterization of the AP2\_G4 motif

34  
35 Here, we chose a putative G4FS (AP2\_G4), to identify the proteins that interact with  
36  
37 the G4. This AP2\_G4 motif is located at 18 bp upstream, on the non-coding strand, of  
38  
39 the transcription start site (TSS) of a dynamically expressing putative AP2  
40  
41 transcription factor (Pf3D7\_0934400) (Figure. 1a) (Aurrecochea *et al.*, 2009). This  
42  
43 motif is predicted as potential G4FS with a high propensity score of 2.55 by G4Hunter  
44  
45 (Gazanion *et al.*, 2020). G4Hunter is a G4 prediction tool that gives the quadruplex  
46  
47 propensity score of a given nucleic acid sequence based on G-richness and G-  
48  
49 skewness of a sequence (Bedrat *et al.*, 2016). Next, we determined the ability of the  
50  
51 AP2\_G4 to form G4 using different biophysical experiments including thermal  
52  
53 differential spectra (TDS), UV melting profile, and circular dichroism (CD). In  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 addition, we tested a control sequence (mut\_G4), where the central Guanine (G) of  
4 AP2\_G4 is substituted by Adenine (A) to disrupt the G4 formation (Figure 1a). The  
5 TDS of AP2\_G4 displayed two positive peaks at 240 nm and 275 nm, and a negative  
6 peak at 295 nm, corresponding to the TDS signature feature of G4 formation (Figure  
7 1b) (Mergny *et al.*, 2005). In contrast, mut\_G4 did not display any significant peaks of  
8 G4 formation. Similarly, using UV melting profiles of the probes at 295 nm, we  
9 observed that AP2\_G4 displays an inverted transition with a  $T_m$  of  $68 \pm 1$  °C,  
10 characteristics of stable G4 while mut\_G4 does not form any G4 (Figure S1) (Mergny  
11 *et al.*, 1998). Having confirmed that AP2\_G4 forms a G4, we next evaluated the G4  
12 topology using CD spectra recordings at 20°C with 100 mM KCl. The CD spectra of  
13 the AP2\_G4 showed a positive peak at 260 nm and a negative peak at 240 nm, which  
14 are distinctive of parallel G4 (Figure 1c) (del Villar-Guerra *et al.*, 2018). Taken  
15 together, the data showed that AP2\_G4 forms a stable parallel G4 whereas mut\_G4  
16 does not form a G4 under same conditions.

### 37 **Identification of PfGBP2 as G4 binding protein by DNA pull-down based approach**

38  
39 We next used the biotinylated G4-forming AP2\_G4 or control mut\_G4  
40 oligonucleotides as a bait in DNA pull-down based approach to identify G4-BPs from  
41 nuclear lysate of intra-erythrocytic parasites (Figure S2a). We identified 29 potential  
42 G4-BPs based on the label-free quantification (LFQ) intensity score (Supplementary  
43 Table 1). Among these candidates, we identified p1/s1 nuclease and Replication A1  
44 protein, which interacts with G4 in other organisms, hence reinforcing the value of our  
45 approach (Qureshi *et al.*, 2012; Zhou *et al.*, 2013). The gene ontology (GO) analysis of  
46 the obtained candidates revealed most of these proteins are involved in nuclear-based  
47 processes such as translation, transcription, replication, and telomere maintenance  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 (Figure S2b). Since a majority of G4FS are identified in telomeric regions of  
4 Plasmodium genome that can form stable G4 *in vitro* (De Cian *et al.*, 2008;  
5 Smargiasso *et al.*, 2009; Calvo and Wasserman, 2016; Stanton *et al.*, 2016; Gazanion  
6 *et al.*, 2020), hence, we focused on one of the obtained candidates PfGBP2, shown to  
7 bind to a telomeric sequence of *P. falciparum in vitro* (Calvo and Wasserman, 2015).  
8  
9  
10  
11  
12  
13  
14

### 15 **Domain characterization of PfGBP2**

16  
17  
18 PfGBP2 is a putative single-strand telomere-binding protein, which was named after  
19 its homolog *Cryptosporidium parvum* CpGBP (Calvo and Wasserman, 2015).  
20  
21 Multiple sequence alignment of PfGBP2 revealed that PfGBP2 contains two highly  
22 conserved RNA recognition motifs (RRMs) of approximately 80 amino acid residues  
23 long, which binds to single-stranded RNAs (Keene and Query, 1991). However,  
24 unlike its homologs, PfGBP2 contains a long arginine-rich (30 %) hinge region  
25 between the RRM domains. This hinge region contains two RGG motifs, which  
26 interact with nucleic acids and protein (FigureS3a) (Thandapani *et al.*, 2013). We next  
27 sought to examine the evolutionary conservation of PfGBP2 across closely related  
28 Plasmodium species, called *Laverania* species. The *Laverania* species are a subgenus  
29 of Plasmodium that can infect apes and cause malaria. The bioinformatics analysis  
30 showed that the PfGBP2 is highly conserved among closely related species except for  
31 *P. billcollinsi G01*, suggesting the crucial role of GBP2 across the *Laverania* species  
32 (Figure S3b).  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49

### 50 **Characterization of *in vitro* interaction between recombinant PfGBP2 and AP2\_G4**

51  
52 We expressed and purified N-terminal His-tagged PfGBP2 protein from *Escherichia*  
53 *coli* to confirm *in vitro* interaction between PfGBP2 and G4-forming AP2\_G4 (Figure  
54 S4a). Using standard filter-binding assays (Wong and Lohman, 1993), we showed  
55  
56  
57  
58  
59  
60



1  
2  
3 that PfGBP2 binds to AP2\_G4 ( $K_d = 0.7698 \pm 0.62 \mu\text{M}$ ), which is 23-fold tighter than  
4  
5 binding to mut\_G4 ( $K_d = 17.72 \pm 14.45 \mu\text{M}$ ) (Figure 2a and S4b). The selective  
6  
7 binding of PfGBP2 to AP2\_G4 was further confirmed by electrophoretic mobility shift  
8  
9 assay (EMSA)(Figure S4c). Altogether, these binding studies confirmed that PfGBP2  
10  
11 preferentially binds to the AP2\_G4 *in vitro*.  
12  
13  
14

### 15 **Genome-wide identification of PfGBP2 binding sites in *P. falciparum***

16  
17  
18 To study the role of PfGBP2 *in vivo*, we successfully generated endogenously  
19  
20 expressing HA-tagged PfGBP2 parasite line (iKO-PfGBP2) based on CRISPR-Cas9  
21  
22 gene editing (Figure S5a-c) (Knuepfer *et al.*, 2017). The immunofluorescence assay  
23  
24 (IFA) showed that PfGBP2 is expressed in all the intra-erythrocytic stages (rings,  
25  
26 trophozoites, and schizonts) of parasites (Figure 3a). Moreover, PfGBP2 displays a  
27  
28 diffuse cytoplasmic pattern along with a punctuate pattern in the nuclei, predominantly  
29  
30 in schizonts.  
31  
32  
33

34  
35 To determine the genome-wide binding sites of the PfGBP2, we performed two  
36  
37 independent chromatin immunoprecipitation (ChIP) assays followed by high  
38  
39 throughput sequencing (ChIP-Seq) from HA-tagged PfGBP2 expressing parasite line  
40  
41 (iKO-PfGBP2) with  $\alpha$ -HA or IgG antibodies. We used the overlapping peaks with  
42  
43  $\log Q > 50$  obtained from both the experiments in our further studies (Figure S6a and b)  
44  
45 (Supplementary Table 2). The ChIP-Seq data revealed that PfGBP2 binds to multiple  
46  
47 sites throughout the genome of the parasite with the enriched peaks localized on the  
48  
49 edges of the chromosomes (Figure 3b), which are enriched with G4FS in *P.*  
50  
51 *falciparum* (Smargiasso *et al.*, 2009; Stanton *et al.*, 2016; Gazanion *et al.*, 2020). We  
52  
53 next searched for putative G4FS within the PfGBP2 binding sites by the G4Hunter  
54  
55 web application with a window size of 25 nucleotides and a threshold score of 1.2  
56  
57  
58  
59  
60

1  
2  
3 (Brázda *et al.*, 2019). The G4 prediction analysis showed that our dataset contains  
4  
5 1,949 putative G4FS (Supplementary Table 3), while the whole genome of *P.*  
6  
7 *falciparum*, with a GC content of 19.3%, harbors 6,033 putative G4FS. Therefore, the  
8  
9 data showed that PfGBP2 binding sites have an increased GC content of 26.8% and  
10  
11 15-fold enrichment of G4FS as compared to the *P. falciparum* genome  
12  
13 (hypergeometric p-value = 2,994 e-8), indicating *in vivo* binding of PfGBP2 to G4 or  
14  
15 G4FS. To further explore the binding specificity of PfGBP2, we looked for consensus  
16  
17 binding motif of PfGBP2 among its binding regions using the MEME suite tool with a  
18  
19 window size of 12 nucleotides wide (Machanick and Bailey, 2011). As anticipated, the  
20  
21 obtained top motif corresponds to the telomeric repeat of *P. falciparum* (GGGTTYA,  
22  
23 where Y is T or C) (Vernick and McCutchan, 1988; Scherf, 1996), confirming that  
24  
25 PfGBP2 is *bona fide* telomere binding protein (Figure 3c).  
26  
27  
28  
29  
30

### 31 **Effect of deletion of the PfGBP2 gene on parasite proliferation, telomere length and *var*** 32 **gene expression** 33 34 35

36  
37 To study the physiological role of PfGBP2 on parasite viability, we excised the  
38  
39 PfGBP2 gene in the iKO-PfGBP2 parasite line with rapamycin treatment via  
40  
41 DiCre/Lop recombination system, thus resulting into PfGBP2-knockout ( $\Delta$ PfGBP2)  
42  
43 parasite line (Figure S7a-c). The growth phenotype analysis showed that there is a  
44  
45 subtle effect on the growth of the  $\Delta$ PfGBP2 parasite line to control parasite lines over  
46  
47 two consecutive cycles (Figure S7d), thus highlighting the dispensability of PfGBP2  
48  
49 for the intra-erythrocytic cycle of the parasite. As we have confirmed that PfGBP2  
50  
51 interacts with G4FS, we next examined whether the presence of the highly selective  
52  
53 G4-stabilizing ligand pyridostatin (PDS) would affect the growth of parasites lacking  
54  
55 PfGBP2. We observed that  $\Delta$ PfGBP2 parasite line did not survive as compared to the  
56  
57  
58  
59  
60

1  
2  
3 control parasites when cultured with PDS (Figure 4a), thus indicating that parasites are  
4 highly sensitive to PDS in the absence of PfGBP2.  
5  
6

7  
8 Having previously shown that PfGBP2 binds to the telomere, we further investigated  
9 whether PfGBP2 depletion affects telomere length of the parasites. To address this, we  
10 isolated the genomic DNA from  $\Delta$ PfGBP2 parasite line to perform telomere  
11 restriction fragment (TRF) length using Southern blot analysis with a biotinylated  
12 probe against the telomeric repeats of *P. falciparum*. The data demonstrated that the  
13 TRF length profile is comparable in both  $\Delta$ PfGBP2 and control parasite lines (Figure  
14 4b). Moreover, the mean telomere length of these parasites is consistent with the  
15 average length of telomeres (~ 1.5 kb long) in *P. falciparum* (Figueiredo *et al.*, 2002),  
16 hence, suggesting that deletion of PfGBP2 does not affect telomere length for up to 30  
17 generations.  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31

32 Previous studies have shown that several G4s are present in the regulatory regions of  
33 *var* genes and are associated in the regulation of *var* gene expression (Smargiasso *et*  
34 *al.*, 2009; Stanton *et al.*, 2016; Claessens *et al.*, 2018; Gazanion *et al.*, 2020). Mutually  
35 exclusive expression of *var* genes contribute to antigenic virulence and immune  
36 evasion in *P. falciparum*. Therefore, we examined the effect of loss of PfGBP2 protein  
37 on the expression of *var* genes by qRT-PCR. The analysis showed that the same *var*  
38 gene (PF3D7\_0711700) was expressed at a similar level in both  $\Delta$ PfGBP2 and  
39 control parasite lines even after 15 generations (Figure 4c). This indicates that a  
40 mutually exclusive *var* gene expression is retained in the parasites lacking PfGBP2. In  
41 contrast, we observed slight derepression in the expression of remaining *var* genes in  
42 the  $\Delta$ PfGBP2 parasite line, although the overall level of expression is low (Figure  
43 4c). Altogether, the results indicate that PfGBP2 could play a role in the regulation of  
44 *var* gene expression.  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## Discussion

Several computational and experimental studies unveiled the existence of G4 in the genomes of *P. falciparum* (De Cian *et al.*, 2008; Smargiasso *et al.*, 2009; Calvo and Wasserman, 2016; Stanton *et al.*, 2016; Anas *et al.*, 2017; Harris *et al.*, 2018; Marsico *et al.*, 2019; Gazanion *et al.*, 2020). However, there is a lack of information on the proteins that can interact or support the G4 formation within these parasites.

Here, we report the first comprehensive study on the identification of G4-BPs that can interact with G4s and modulate their function. Using a DNA pull-down based approach, we identified 29 potential G4-BPs in *P. falciparum*. The gene ontology (GO) analysis of the obtained candidates revealed that the most of these proteins are involved in nuclear-based processes such as translation, transcription, replication, and telomere maintenance.

A previous study has shown that PfGBP2 binds to the telomeric sequence of *P. falciparum* *in vitro*, however, no functional implications have been elucidated (Calvo and Wasserman, 2015). Based on ChIP-Seq data, we confirmed that PfGBP2 is a bona fide telomere binding protein of *P. falciparum*. Besides, we showed that PfGBP2 binding sites are enriched with putative G4FS and PfGBP2 binds to G4 forming AP2\_G4 motif *in vitro*, indicating that PfGBP2 is also a bona fide G4 binding protein in *P. falciparum*.

In a recent study, Gazanion *et al* showed that PDS-treated parasites displays large scale perturbations in the expression of plasmodial genes (Gazanion *et al.*, 2020). Interestingly, the data showed that there is an upregulation in the expression of PfGBP2 in PDS-treated parasites. In addition, we showed that the parasites lacking PfGBP2 are highly sensitive to PDS treatment. Altogether, the data suggest that PDS-treated parasites upregulates the expression of PfGBP2 that can recognize and bind to these folded G4 to cope with the

1  
2  
3 increased G4 formation while lacking PfGBP2 are unable to survive. Thus, we speculate that  
4  
5 PfGBP2 plays a critical role in modulating G4-related phenotypes.  
6  
7

8  
9 Several telomere-associated proteins are shown to interact with telomeric G4 and regulate  
10  
11 telomere maintenance (Wang *et al.*, 2012; Pagano *et al.*, 2015). For instance, a variant of  
12  
13 heterogeneous nuclear ribonucleoprotein A2 (hnRNP A2 \*) unfolds the telomeric G4 and  
14  
15 promotes telomere extension by providing access to telomerase while even though human  
16  
17 POT1 unfolds the telomeric G4, it inhibits telomere extension by blocking the access of  
18  
19 telomerase (Kelleher *et al.*, 2005). In our study, we observed that telomere binding protein  
20  
21 PfGBP2 interacts with G4, but the loss of protein does not affect telomere length for up to 30  
22  
23 generations. Likewise, the deletion of Rlf6p/GBP2, the yeast homolog of PfGBP2, did not  
24  
25 affect the telomere length in *S. cerevisiae* (Konkel *et al.*, 1995). This led us to speculate that  
26  
27 PfGBP2 might contribute to telomere structure complex or the subtle effect of PfGBP2-KO  
28  
29 on telomere length could be due to presence of additional proteins that have partially or  
30  
31 completely redundant telomere functions with PfGBP2.  
32  
33  
34  
35  
36

37  
38 In *Plasmodium falciparum*, telomere-associated proteins such as histone deacetylases  
39  
40 PfSir2A and PfSir2B are shown to regulate the transcription of *var* genes in *P. falciparum*  
41  
42 (Figueiredo and Scherf, 2005; Tonkin *et al.*, 2009). The knocking out of both PfSir2A and  
43  
44 PfSir2B affected the expression of *var* gene while the absence of PfSir2A, but not PfSir2B  
45  
46 affected the telomere lengths. Likewise, we observed that the loss of PfGBP2, modestly  
47  
48 derepress the expression of several *var* genes after 15 generations. Given that several G4FS  
49  
50 are present in the regulatory regions of *var* genes (Stanton *et al.*, 2016; Claessens *et al.*, 2018;  
51  
52 Gazanion *et al.*, 2020), it could be possible that the PfGBP2 regulates the expression of *var*  
53  
54 genes by interacting with G4. Gazanion et al showed that G4 found in the non-coding strand  
55  
56 of *var* promoters modulates reporter gene expression (Gazanion *et al.*, 2020). Thus, it would  
57  
58  
59  
60

1  
2  
3 be interesting to re-examine the effect of PfGBP2 on *var* gene expression by culturing these  
4  
5 PfGBP2 lacking parasites for a longer duration or with PDS.  
6  
7

8  
9 On the side note, PfGBP2 might be involved in RNA metabolism due to the presence  
10  
11 of the RNA recognition motif (RRM) region (Maris *et al.*, 2005) and its localization in  
12  
13 the cytoplasmic region. In fact, the homologs in yeast and human have been described  
14  
15 to regulate RNA metabolic processes (Weighardt *et al.*, 1996; Windgassen and  
16  
17 Krebber, 2003), thus further studies are need to be carried out to study the role of  
18  
19 PfGBP2 in RNA related functions.  
20  
21

22  
23  
24 In conclusion, we here report the first snapshot of the G4-interactome in *P.*  
25  
26 *falciparum*. We identified several G4-BPs and further investigated one of the  
27  
28 candidates, PfGBP2, during the intra-erythrocytic development of *P. falciparum*. We  
29  
30 showed that PfGBP2 is a *bona fide* G4-binding protein and essential for parasites  
31  
32 when cultured in G4- stabilizing conditions. However, the mechanism of G4 binding  
33  
34 needs to be addressed in future studies. Moreover, our study further supports the  
35  
36 existence of G4 in *P. falciparum* and opens new avenues for dissecting the molecular  
37  
38 mechanisms underlying the G4 functions in these malarial parasites.  
39  
40  
41  
42

## 43 **Experimental procedures**

### 44 **Parasite culturing**

45  
46  
47 Parasites were cultured in A<sup>+</sup> human erythrocytes in RPMI 1640 medium (Gibco Life  
48  
49 Technologies, 52400 RPMI 1640, HEPES) supplemented with 5% human serum and 0.5%  
50  
51 Albumax II, 0.2 mM hypoxanthine (C.C.Pro GmbH) and 25 µg/mL gentamicin (Sigma). The  
52  
53 cultures were kept at 37°C under a controlled trigaz atmosphere (3% CO<sub>2</sub>, 5% O<sub>2</sub>, and 92%  
54  
55 N<sub>2</sub>). Synchronization of parasites was done by a sequential combination of Percoll (Saul *et*  
56  
57  
58  
59  
60

1  
2  
3 *al.*, 1982) and sorbitol treatment. (Lambros and Vanderberg, 1979) Parasite development was  
4  
5 monitored by Giemsa staining. (Moll *et al.*, 2013)  
6  
7

### 8 **Construction of plasmid and Transfection**

9  
10  
11 To generate an inducible Knockout parasite line, we have employed the combined approach  
12 of CRISPR-Cas9 and DiCre/Loxp recombination system.(Knuepfer *et al.*, 2017) The  
13 modified pLN vector (consisting of multiple cloning site and HA<sub>3</sub>loxp region ) was used as a  
14 template for donor plasmid (pLN-don-HA<sub>3</sub>loxp) where 5' homology region, modified gene  
15 coding region, and 3' homology regions were cloned using infusion cloning. PCR amplified  
16 400 bp upstream of the start codon to the coding region of the respective gene until the  
17 insertion site of loxpintrom was cloned into *ApaI* and *EcoRV*; Gblock consisting of loxpintrom  
18 and recodonized gene until stop codon was synthesized by geneart gene synthesis  
19 (ThermoFisher) and cloned into *BseR1* and *MluI*; PCR amplified 3' HR containing 3'UTR of  
20 the gene was cloned into *SpeI* and *XhoI*. All the PCR reactions were performed using  
21 *PfuUltra II* fusion DNA polymerase and resulting plasmids were verified using Sanger  
22 sequencing. The gRNA was selected for the respective gene using the CHOPCHOP tool and  
23 was cloned at the *BbsI* site of pDC- Cas9 U6-hdhfr/yfcu vector (gifted by Ellen knuepfer).  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42

43 Synchronized ring stage of the Pf3D7 p230pDiCre line was transfected with the plasmid  
44 DNA mixture consists of 60 µg of pDC2 Cas9/gRNA/hDHFR/yFCU plasmid and 60 µg of  
45 linearized pLN-loxp donor plasmid (linearized with *ApaI* and *XhoI*) by electroporation.  
46 Electroporation was performed using a Bio-Rad Gene Pulser at settings of 310 V, 950 µF,  
47 and 200 Q in 0.2 cm cuvette. The obtained time constants were 9-12 ms. These electroporated  
48 samples were immediately mixed with culture media and grown in agitation. After 4 hours of  
49 post-transfection, the culture medium was replaced with fresh media supplemented with 2.5  
50 nM of WR99210, which was withdrawn after 5 days. Once the parasites were observed on  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 Giemsa staining, limiting dilution cloning was performed to obtain the individual clones with  
4 desired endogenous locus modification. Parasites were genotyped to check for correct  
5 modification at the gene locus using primers. All the primers used in the construction of the  
6 plasmids and genotyping are listed in this ESI.  
7  
8  
9  
10  
11  
12

13 In order to induce the DiCre mediated knockout of the gene, early ring-stage parasites were  
14 treated with rapamycin (20 nM) upto 12 h. After treatment, parasites were washed and  
15 returned to culture.  
16  
17  
18  
19  
20

### 21 **DNA Pull-down assay and Mass Spectrometry**

22  
23

24 The biotinylated oligonucleotides, AP2\_G4 (bio-GGGATTTGGGAGGGGGGG) and  
25 Mut\_G4 (bio-GAGATTTGAGAGGGAGGG) were synthesized by Eurofins Genomics.  
26 These oligos (1200 pmol /500  $\mu$ L) were heated at 95°C, followed by cooling at RT for  
27 overnight in the presence of 150 mM KCl. The nuclear lysate was prepared from *P.*  
28 *falciparum* NF54 as previously described. (Chêne *et al.*, 2012)  
29  
30  
31  
32  
33  
34  
35  
36

37 The subtractive-based pull-down assay was performed where the nuclear lysate was first pre-  
38 cleared with mut\_G4 and then incubated with AP2\_G4. Further steps were carried out as  
39 described in the published protocol. (Jutras *et al.*, 2012) Eluted fractions were dialyzed  
40 against the dialysis buffer (25mM Tris-Cl, 6mM DTT, and 1mM iodoacetamide) using  
41 Vivaspin 10MWCO tubes.  
42  
43  
44  
45  
46  
47  
48

49 Thereafter, these eluted proteins were digested in solution. Briefly, each sample was diluted  
50 (final volume 100  $\mu$ L) in TEAB 100 mM. One microliter of DTT 1 M was added and  
51 incubation was performed for 30 min at 60 °C. A volume of 10  $\mu$ L of IAA 0.5 M was added  
52 (incubation for 30 min in the dark). Enzymatic digestion was performed by the addition of  
53 1  $\mu$ g trypsin (Gold, Promega, Madison USA) in TEAB 100 mM and incubation overnight at  
54  
55  
56  
57  
58  
59  
60



1  
2  
3 30 °C. After digestion, peptides were purified and concentrated using OMIX (Agilent  
4 Technologies Inc.) according to the manufacturer's specifications. Peptides were dehydrated  
5 in a vacuum centrifuge. After resuspension in formic acid (0.1%, buffer A) samples were  
6 loaded onto a 50 cm reversed-phase column (75 mm inner diameter, Acclaim Pepmap 100®  
7 C18, Thermo Fisher Scientific) and separated with an Ultimate 3000 RSLC system (Thermo  
8 Fisher Scientific) coupled to a Q Exactive HF (Thermo Fisher Scientific) via a nano-  
9 electrospray source, using a 125 min gradient of 2 to 40% of buffer B (80% ACN, 0.1%  
10 formic acid) and a flow rate of 300 nL/min.  
11  
12

13 MS/MS analyses were performed in a data-dependent mode. Full scans (375-1,500 m/z) were  
14 acquired in the Orbitrap mass analyzer with a 60,000 resolution at 200 m/z. For the full scans,  
15 3e6 ions were accumulated within a maximum injection time of 60 ms and detected in the  
16 Orbitrap analyzer. The twelve most intense ions with charge state  $\geq 2$  were sequentially  
17 isolated ( $1e5$ ) with a maximum injection time of 45 ms and fragmented by HCD (Higher-  
18 energy collisional dissociation) in the collision cell (normalized collision energy of 28%) and  
19 detected in the Orbitrap analyzer at 30,000 resolutions. Raw spectra were processed using the  
20 MaxQuant environment (Cox and Mann, 2008) and Andromeda for database search with  
21 label-free quantification (LFQ) and match between runs (Cox *et al.*, 2011). The MS/MS  
22 spectra were matched against the UniProt Reference proteomes (Proteome ID UP000001450  
23 and UP000005640) of *Plasmodium falciparum* and Human and 250 frequently observed  
24 contaminants as well as reversed sequences of all entries (MaxQuant contaminant database).  
25 Enzyme specificity was set to trypsin/P, and the search included cysteine  
26 carbamidomethylation as a fixed modification and oxidation of methionine and acetylation  
27 (protein N-term) as variable modifications. Up to two missed cleavages were allowed for  
28 protease digestion. FDR was set at 0.01 for peptides and proteins and the minimal peptide  
29 length at 7.  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## Biophysical confirmation of G4 folding of oligonucleotides

Oligonucleotides were purchased from Eurogentec (Belgium) and were resuspended at a strand concentration of 4  $\mu\text{M}$ , unless stated, in the cacodylic acid buffer (10mM cacodylic acid, pH 7.2 adjusted with LiOH), containing 100mM KCl, followed by carrying out the measurements for biophysical experiments in 1cm quartz cells.

### Absorbance Spectroscopy

Absorption spectra and thermal denaturation profiles (absorbance as a function of temperature) were acquired on a UVIKON XS spectrophotometer. For thermal denaturation profiles, the absorbance was recorded at different wavelengths (such as 245, 260, 273, 295, and 335 nm), while cooling and then heating the samples between 5°C and 95°C at a rate of 0.2°C/min. Melting temperatures ( $T_m$ ) were graphically calculated as the intercept between the melting curves and the median lines between low and high-temperature absorbance linear baselines. Thermal difference spectra (TDS) was obtained by subtracting the absorption spectra at 2°C (low temperature) from the one at 90°C (high temperature), the spectrum were recorded after annealing from 90°C to 2°C at 0.2°C/min. (Mergny *et al.*, 2005) UV/Vis spectra were recorded on a SAFAS spectrometer.

### Circular Dichroism

Circular dichroism (CD) spectra were recorded by using a Jasco J-810 spectropolarimeter at 20°C or 5°C, after annealing from 90°C to 2°C at 0.2°C/min. The oligonucleotide was dissolved at a strand concentration of 3 or 6  $\mu\text{M}$  in the cacodylic acid buffer (as mentioned above).

### Bioinformatic analysis

1  
2  
3 The fasta sequence of PfGBP2 (Q8IJX3) and its homologs including *Chlamydomonas*  
4 *reinhardtii* CrGBP1p (Q39568), *Cryptosporidium parvum* CpGBP (Q86PS0),  
5  
6 *Saccharomyces cerevisiae* ScGBP2 (P25555), and *Homo sapiens* HnRNP-A2 (P22626) and  
7  
8 Laverania species: *P. billcollinsi* G01 (PBILCG01\_1003800), *P. blacklocki* G01  
9  
10 (PBLACG01\_1004900), *P. praefalciparum strain* G01 (PPRFG01\_1008000), *P. reichenowi*  
11  
12 CDC (PRCDC\_1006200), *P. adleri* G01 (PADL01\_1005300) and *P. gaboni* G01  
13  
14 (PGABG01\_1004800) were obtained from PSI-BLAST, Plasmodb, and UniProt. Multiple  
15  
16 sequence alignment of PfGBP2 and its homologs were carried out using the default  
17  
18 MUSCLE algorithm at [www.phylogeny.fr](http://www.phylogeny.fr).  
19  
20  
21  
22  
23  
24

### 25 **Cloning, expression, and purification of recombinant protein**

26  
27  
28 The recodonized PfGBP2 coding region was synthesized and cloned into the NdeI and  
29  
30 BamHI sites of a pET15b expression vector (Novagen, Germany) using infusion cloning. The  
31  
32 resulting construct was confirmed by restriction digestion and sequencing. The N-terminal 6×  
33  
34 His tagged protein was expressed in BL21 (DE3) pLysS cells grown in medium  
35  
36 supplemented with 100 µg/mL Ampicillin and 25 µg/mL of chloramphenicol. Expression  
37  
38 was induced at an optical density of 0.7-0.9 at 600 nm, with 0.5- 1 mM isopropyl β-D-  
39  
40 thiogalactoside (IPTG) at 25 °C for 5h.  
41  
42  
43  
44

45 All the purification steps were carried out at 4 °C. Cell lysis was performed in lysis buffer (50  
46  
47 mM Tris-HCl, pH 8.0, 300 mM NaCl, 5% (v/v) glycerol and 20 mM of imidazole) by  
48  
49 sonication (0.5s ON, 0.5s OFF at 60% pulse intensity, 3 min) using a Branson Digital  
50  
51 Sonifier 450-D. After centrifugation of the lysate (14000 rpm, 30 min), the supernatant was  
52  
53 loaded onto HisPur Cobalt Spin column and purification was carried out according to the  
54  
55 manufacturer's instruction. The eluates were run through Superdex 200 16/600 (10 – 60kDa)  
56  
57 size exclusion chromatography. The positive fraction was combined and concentrated using  
58  
59  
60

1  
2  
3 Amicon Ultra-15 centrifugal filters. Expression and purification of His-tagged protein were  
4 confirmed by coomassie staining and western blot analysis with the  $\alpha$ -his antibody.  
5  
6  
7

### 8 **Double filter binding assay**

9  
10  
11 Double filter binding assay was performed as described earlier with some  
12 modifications.(Götz *et al.*, 2019) DNA (20 pmol) was 5' labeled with 25  $\mu$ Ci [ $\gamma$ -<sup>32</sup>P] ATP by  
13 T4 polynucleotide kinase (NEB, Ipswich,UK ), followed by purification of AP2\_G4 and  
14 mut\_G4 structures from 7 % SDS-PAGE. DNA protein binding was analyzed by double-  
15 filter binding assays (Wong and Lohman, 1993) using a 96-well Bio-Dot SF apparatus (Bio-  
16 Rad, Hercules, CA, USA). 0.1  $\mu$ M DNA in binding buffer (50 mM Tris/HCl pH 8.0, 125 mM  
17 KCl, 5 mM DTT, 10% (v/v) glycerol (Paeschke *et al.*, 2005) was incubated with the  
18 increased concentration of recombinant His-tagged PfGBP2 from 0 to 25  $\mu$ M for 20 min at  
19 21°C. After incubation, the reactions were filtered through nitrocellulose and a positively  
20 charged nylon membrane, followed by three washing steps with binding buffer and drying of  
21 the membranes. The membranes were analyzed by phosphorimaging on a Typhoon FLA  
22 7000 (GE Healthcare). The DNA bound percentage values of PfGBP2 were determined using  
23 ImageQuant and the graph was plotted curve fitting using linear regression on GraphPad  
24 Prism.  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44

### 45 **EMSA**

46  
47 Electromobility shift assay (EMSA) was performed using biotinylated oligos and expressed  
48 recombinant protein as described earlier with some modifications (Takahama *et al.*, 2011).  
49 Previously used biotin-labeled oligonucleotides (AP2\_G4 and Mut\_G4) were diluted to 100  
50 nM in 10 mM Tris-HCl, pH 7.5 with or without 150 mM KCl. The quadruplex formation was  
51 performed by heating the samples at 95 °C and then cooling it at room temperature overnight  
52 in the presence of potassium ions. DNA-protein binding reaction was carried out in a final  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 volume of 15 uL using 100 nM oligonucleotide with a varying concentration of recombinant  
4  
5 protein in a binding buffer (50 mM Tris-HCl pH 7.5, 0.5 mM EDTA, 0.5 mM DTT, 0.1  
6  
7 mg/mL bovine serum albumin, and 100 mM KCl ). After the samples were incubated for 1h  
8  
9 at 25 °C, they were resolved on a 10% native PAGE gel with or without 20 mM KCl. The  
10  
11 native PAGE gel was pre-ran at 100 V for 1h at 4 °C. Electrophoresis was performed at 100V  
12  
13 for 1h at 4 °C in 0.5× TBE buffer supplemented with or without 20 mM KCl. The DNA  
14  
15 bound proteins were transferred to the Amersham Hybond- N<sup>+</sup> nylon membrane by  
16  
17 electroblotting at 130 mA for 30 min using Trans-Blot Semi-dry blot transfer cell (Biorad).  
18  
19 The interactions were detected using a Thermo Scientific™ Pierce™ LightShift  
20  
21 chemiluminescent EMSA kit following the manufacturer's guidelines.  
22  
23  
24  
25

### 26 27 **ChIP-Seq**

28  
29  
30 Infected RBCs were cross-linked using 1% formaldehyde for 10 min at 37°C. 0.125 M  
31  
32 glycine was added for quenching the cross-linking reaction. The samples were washed thrice  
33  
34 using 1X PBS (chilled) before proceeding with saponin lysis. Followed by the steps as  
35  
36 mentioned in the previously published ChIP-Seq protocol with some modifications (Lopez-  
37  
38 rubio *et al.*, 2013). The lysed samples were sonicated in Bioruptor (Bioruptor® Plus ,  
39  
40 Diagenode) with settings (24 cycles of sonication 10 sec on/off) to obtain the chromatin size  
41  
42 of 200-400 bp. Two independent experiments were carried out, where in the first experiment  
43  
44 (rep1), ChIP samples underwent a pre-clearing step of adding IgG prior to HA pull-down step  
45  
46 whereas in the second experiment (rep2), an IgG control was prepared independently from  
47  
48 the IP sample.  
49  
50  
51

52  
53  
54 ChIP-sequencing libraries for all the samples were prepared from purified DNA using TruSeq  
55  
56 ChIP Library Preparation Kit (Illumina), as described by the manufacturer. Libraries were  
57  
58 then quality control checked by a bioanalyzer and qPCR prior to pooling and sequencing on  
59  
60

1  
2  
3 an Illumina NovaSeq 6000 system, flowcell SP (ID AHJHNGDRXX). Sequencing of 150 bp  
4 paired-ends yielded on average  $70 \times 10^6$  reads per library. Fastq files were obtained by  
5 demultiplexing the data using bcl2fastq software (Illumina), prior to downstream analysis.  
6  
7 Mapping was performed using BWA mem and peaks were calculated using the command  
8 “callpeak” from MACS2 and the input as “control” sample. The intersection of peaks from  
9  
10 the different replicate samples using Bedtools produced the final dataset. Motif surveys were  
11 performed using MEME (version 5.1.1 (Release date: Wed Jan 29 15:00:42 2020 -0800) The  
12  
13 R package ChipQC and ggplots2 were used to plot the various datasets.  
14  
15  
16  
17  
18  
19  
20  
21

### 22 **Growth phenotype assay**

23  
24  
25 This assay was performed in trophozoites stage of synchronized parasite culture with an  
26 initial parasitaemia of 0.01% in 5% hematocrit. The parasites were treated with or without 20  
27 nM rapamycin, followed by culturing for 2 consecutive cycles (for 6 days), where media was  
28 changed, and slides were prepared every 48 h. The assay was done in triplicates and the  
29 parasitemia was counted under microscopy. Data were represented as mean and SD of  
30 triplicates. To analyze the effect of Pyridostatin on the growth of iKO-PfGBP2 parasites line,  
31 ring-staged parasites were treated for 12 h with 1  $\mu$ M Pyridostatin. Gazanion et al showed  
32 that PDS is toxic for *P. falciparum* with  $IC_{50} = 5.2 \pm 0.9 \mu$ M, (Gazanion *et al.*, 2020) thus we  
33 chose 1  $\mu$ M PDS that has moderate effect on parasite growth for our growth phenotype  
34 analysis.  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48

### 49 **Telomere Restriction Fragment (TRF) Southern Blotting**

50  
51  
52 To perform TRF southern blotting, the iKO-PfGBP2 parasite line was cultured and collected  
53 after 30 and 60 days of with or without rapamycin treatment. The genomic DNA was  
54 extracted from saponin-lysed parasites using the MN NucleoSpin Blood QuickPure kit  
55 according to the manufacturer’s guidelines.  
56  
57  
58  
59  
60

1  
2  
3 Two micrograms of genomic DNA was digested with 10 units of each enzyme: AluI, DdeI,  
4 MboII, RsaI overnight at 37°C. Digestions were precipitated and run in 1% agarose gels at 70  
5 V for 5 h and transferred to a Hybond N+ membrane overnight through capillarity (Sambrook  
6 et al. Molecular Cloning). The DNA was then cross-linked to the membrane in a UV oven for  
7 1 min and pre-hybridized in 6X SSC 0.1% SDS 2.5% skimmed milk for 1 h at 42°C. 10 pmol  
8 of biotinylated telomeric probes (PG184-5' GGGTTTAGGGTTTAGGGTTTAGGGTTTA 3'  
9 and PG185-5'GGGTTTCAGGGTTTCAGGGTTTCAGGGTTCA 3') were added and incubated  
10 overnight, followed by four times 30 min washes of 6X SSC with 0.1% SDS. The biotin-  
11 labeled probes were detected using the LightShift Chemiluminescent EMSA Kit and imaging  
12 in a BioRad ChemiDoc.

### 26 **Quantitative reverse transcription-PCR (qRT-PCR) on *var* genes**

27  
28 To perform qRT-PCR, samples were prepared from the iKO-PfGBP2 parasite line that is  
29 treated with or without rapamycin for 30 days. Using saponin treatment, parasites were  
30 harvested at 14-17h, where peak expression of *var* genes is observed. The obtained parasite  
31 pellet was resuspended in Trizol to extract the RNA, as described in the published  
32 protocol.(Moll *et al.*, 2013)

33  
34 Quantitative PCR was performed on cDNA using specific primers for each *var* gene, as  
35 previously described (Salanti *et al.*, 2003) with few modifications from Dzikowski *et al*  
36 (Dzikowski *et al.*, 2006). DNase-treated RNA samples were reverse-transcribed into cDNA,  
37 using SuperScript III first-strand synthesis SuperMix (Invitrogen), according to  
38 manufacturer's instructions. Alternatively, RNA samples were run without RT enzyme, to  
39 check for genomic DNA contamination. Target genes in cDNA samples were quantified  
40 using PowerUp SYBR Green Master Mix (Applied Biosystems) and normalized using the  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 housekeeping gene-Serine-tRNA-ligase. The results were expressed as relative copy number  
4  
5 with SD.  
6  
7

### 8 **Western blotting**

9  
10  
11 Saponin harvested parasites (equivalent to  $10^8$  parasites per lane) were lysed using  $1\times$   
12  
13 Laemmli Buffer supplemented with  $1\times$  protease inhibitors and 50 mM DTT. The proteins  
14  
15 were run on a 10 % SDS-PAGE gel and transferred onto a PVDF membrane. The membrane  
16  
17 was blocked with 3 % BSA/TBST for 1h at RT. The immunoblot was probed with primary  
18  
19 antibody (rat  $\alpha$ -HA 3F10) in 3% BSA/TBST for 1h at RT, followed by incubation with HRP-  
20  
21 conjugated secondary antibody. The immunoblot was developed with Clarity Western ECL  
22  
23 substrate (Biorad).  
24  
25  
26  
27

### 28 **Immunofluorescence assay**

29  
30  
31 Immunofluorescence assay was performed on smears of infected erythrocytes. The smear was  
32  
33 fixed with 4% PFA for 15 min at RT, followed by neutralization with 0.1M glycine/PBS for  
34  
35 10 min. Cells were permeabilized with 0.1% Triton-X-100/PBS for 15 min. Washing with  $1\times$   
36  
37 PBS was performed after every step. Cells were blocked with 1.5 % BSA/PBS for 60 min,  
38  
39 followed by incubation with primary antibody (rat  $\alpha$ -HA 3F10) in 0.15% BSA/PBS for either  
40  
41 1h at RT or overnight at 4 °C. After three washes, cells were incubated with Alexa  
42  
43 conjugated secondary antibodies, diluted in 0.15% BSA/PBS for 1h at RT. Slides were  
44  
45 washed to remove the unbound antibodies and were stained with Hoechst (diluted in 1: 4000  
46  
47 in  $1\times$  PBS) for 10 min. The cells were mounted with ProLong™ Gold antifade reagent  
48  
49 (Invitrogen, P10144) and a coverslip was placed. Images were captured and processed by  
50  
51 Zen Blue software (Zeiss).  
52  
53  
54  
55  
56  
57  
58  
59  
60



## Acknowledgements

We acknowledge the MRI Imaging facility of the University of Montpellier for providing access to their microscopy imaging facility. This work has been supported by the Atip-Avenir program, ANR « investissements d'avenir » program (ANR-11-LABX-0024-01 « PARAFRAP ») and by the Fondation pour la Recherche Médicale (ARF20150934098).

## Conflict of Interest

There are no conflicts to declare

## References

Anas, M., Sharma, R., Dhamodharan, V., Pradeepkumar, P.I., Manhas, A., Srivastava, K., *et al.* (2017) Investigating Pharmacological Targeting of G-Quadruplexes in the Human Malaria Parasite. *Biochemistry* **56**: 6691–6699.

Aurrecoechea, C., Brestelli, J., Brunk, B.P., Dommer, J., Fischer, S., Gajria, B., *et al.* (2009) PlasmoDB: A functional genomic database for malaria parasites. *Nucleic Acids Res* **37**: 539–543.

Bedrat, A., Lacroix, L., and Mergny, J.-L. (2016) Re-evaluation of G-quadruplex propensity with G4Hunter. *Nucleic Acids Res* **44**: 1746–1759.

Belmonte-Reche, E., Martínez-García, M., Guédin, A., Zuffo, M., Arévalo-Ruiz, M., Doria, F., *et al.* (2018) G-Quadruplex Identification in the Genome of Protozoan Parasites Points to Naphthalene Diimide Ligands as New Antiparasitic Agents. *J Med Chem* **61**: 1231–1240.

Brázda, V., Cerveň, J., Bartas, M., Mikysková, N., Coufal, J., and Pečinka, P. (2018) The

1  
2  
3 amino acid composition of quadruplex binding proteins reveals a shared motif and predicts  
4 new potential quadruplex interactors. *Molecules* **23**: 2341–2357.  
5  
6

7  
8  
9 Brázda, V., Hároníková, L., Liao, J.C.C., and Fojta, M. (2014) DNA and RNA quadruplex-  
10 binding proteins. *Int J Mol Sci* **15**: 17493–17517.  
11  
12

13  
14  
15 Brázda, V., Kolomazník, J., Lýsek, J., Bartas, M., Fojta, M., Šťastný, J., *et al.* (2019)  
16 G4Hunter web application: A web server for G-quadruplex prediction. *Bioinformatics* **35**:  
17 3493–3495.  
18  
19

20  
21  
22 Calvo, E.P., and Wasserman, M. (2015) PfGBP : A Plasmodium falciparum telomere binding  
23 protein. *Quim Org y Bioq* **44**: 5–10.  
24  
25

26  
27  
28 Calvo, E.P., and Wasserman, M. (2016) G-Quadruplex ligands: Potent inhibitors of  
29 telomerase activity and cell proliferation in Plasmodium falciparum. *Mol Biochem Parasitol*  
30 **207**: 33–38.  
31  
32

33  
34  
35 Chêne, A., Vembar, S.S., Rivière, L., Lopez-Rubio, J.J., Claes, A., Siegel, T.N., *et al.* (2012)  
36 PfAlbas constitute a new eukaryotic DNA/RNA-binding protein family in malaria parasites.  
37 *Nucleic Acids Res* **40**: 3066–3077.  
38  
39

40  
41  
42 Cian, A. De, Grellier, P., Mouray, E., Depoix, D., Bertrand, H., Monchaud, D., *et al.* (2008)  
43 Plasmodium telomeric sequences: structure, stability and quadruplex targeting by small  
44 compounds. *Chembiochem* **9**: 2730–2739.  
45  
46

47  
48  
49 Claessens, A., Harris, L.M., Stanojic, S., Chappell, L., Stanton, A., Kuk, N., *et al.* (2018)  
50 RecQ helicases in the malaria parasite Plasmodium falciparum affect genome stability, gene  
51 expression patterns and DNA replication dynamics. *PLoS Genet* **14**: 1–32.  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 Cox, J., and Mann, M. (2008) MaxQuant enables high peptide identification rates,  
4 individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat*  
5  
6 *Biotechnol* **26**: 1367–1372.  
7  
8

9  
10  
11 Cox, J., Neuhauser, N., Michalski, A., Scheltema, R.A., Olsen, J. V., and Mann, M. (2011)  
12  
13 Andromeda: A peptide search engine integrated into the MaxQuant environment. *J Proteome*  
14  
15 *Res* **10**: 1794–1805.  
16  
17

18  
19  
20 Dzikowski, R., Frank, M., and Deitch, K. (2006) Mutually exclusive expression of virulence  
21  
22 genes by malaria parasites is regulated independently of antigen production. *PLoS Pathog* **2**:  
23  
24 0184–0194.  
25  
26

27  
28 Figueiredo, L., and Scherf, A. (2005) Plasmodium telomeres and telomerase: The usual  
29  
30 actors in an unusual scenario. *Chromosom Res* **13**: 517–524.  
31  
32

33  
34 Figueiredo, L.M., Freitas-Junior, L.H., Bottius, E., Olivo-Marin, J.C., and Scherf, A. (2002)  
35  
36 A central role for Plasmodium falciparum subtelomeric regions in spatial positioning and  
37  
38 telomere length regulation. *EMBO J* **21**: 815–824.  
39  
40

41  
42  
43 Gazanion, E., Lacroix, L., Alberti, P., Gurung, P., Wein, S., Cheng, M., *et al.* (2020) Genome  
44  
45 wide distribution of G-quadruplexes and their impact on gene expression in malaria parasites.  
46  
47 *PLoS Genet* **16**: e1008917.  
48  
49

50  
51  
52 González, V., Guo, K., Hurley, L., and Sun, D. (2009) Identification and characterization of  
53  
54 nucleolin as a c-myc G-quadruplex-binding protein. *J Biol Chem* **284**: 23622–23635.  
55  
56

57  
58  
59 Götz, S., Pandey, S., Bartsch, S., Juranek, S., and Paeschke, K. (2019) A novel G-quadruplex  
60  
binding protein in yeast-Slx9. *Molecules* **24**: 14–18.

1  
2  
3 Hacht, A. Von, Seifert, O., Menger, M., Schütze, T., Arora, A., Konthur, Z., *et al.* (2014)  
4 Identification and characterization of RNA guanine-quadruplex binding proteins. *Nucleic*  
5  
6 *Acids Res* **42**: 6630–6644.  
7  
8

9  
10  
11 Harris, L.M., Monsell, K.R., Noulin, F., Toyin Famodimu, M., Smargiasso, N., Damblon, C.,  
12  
13 *et al.* (2018) G-quadruplex DNA motifs in the malaria parasite plasmodium falciparum and  
14  
15 their potential as novel antimalarial drug targets. *Antimicrob Agents Chemother* **62**: 1–32.  
16  
17

18  
19  
20 Jutras, B.L., Verma, A., and Stevenson, B. (2012) Identification of Novel DNA-Binding  
21  
22 proteins Using DNA-Affinity Chromatography/Pull Down. *Curr Protoc Microbiol* 1–13.  
23  
24

25  
26 Keene, D., and Query, C. (1991) Nuclear RNA-binding Proteins. *PNAS* **41**: 179–202.  
27

28  
29  
30 Kelleher, C., Kurth, I., and Lingner, J. (2005) Human Protection of Telomeres 1 (POT1) Is a  
31  
32 Negative Regulator of Telomerase Activity In Vitro. *Mol Cell Biol* **25**: 808–818.  
33  
34

35  
36 Knuepfer, E., Napiorkowska, M., Ooij, C. Van, and Holder, A.A. (2017) Generating  
37  
38 conditional gene knockouts in Plasmodium - A toolkit to produce stable DiCre recombinase-  
39  
40 expressing parasite lines using CRISPR/Cas9. *Sci Rep* **7**: 3881–3893.  
41  
42

43  
44  
45 Konkil, L.M.C., Enomoto, S., Chamberlain, E.M., Mccune-Zierath, P., Iyadurai, S.J.P., and  
46  
47 Berman, J. (1995) A class of single-stranded telomeric DNA-binding proteins required for  
48  
49 Rap1p localization in yeast nuclei. *Proc Natl Acad Sci U S A* **92**: 5558–5562.  
50

51  
52  
53 Lambros, C., and Vanderberg, J.P. (1979) Synchronization of Plasmodium falciparum  
54  
55 Erythrocytic Stages in Culture. *J Parasitol* **65**: 418.  
56  
57

58  
59 Lopez-rubio, J., Siegel, T.N., and Artur (2013) *Genome-wide Chromatin*  
60

1  
2  
3 *Immunoprecipitation-Sequencing in Plasmodium* . .  
4

5  
6  
7 Machanick, P., and Bailey, T.L. (2011) MEME-ChIP: Motif analysis of large DNA datasets.  
8  
9 *Bioinformatics* **27**: 1696–1697.  
10

11  
12  
13 Maris, C., Dominguez, C., and Allain, F.H.T. (2005) The RNA recognition motif, a plastic  
14  
15 RNA-binding platform to regulate post-transcriptional gene expression. *FEBS J* **272**: 2118–  
16  
17 2131.  
18

19  
20  
21 Marsico, G., Chambers, V.S., Sahakyan, A.B., McCauley, P., Boutell, J.M., Antonio, M. Di,  
22  
23 and Balasubramanian, S. (2019) Whole genome experimental maps of DNA G-quadruplexes  
24  
25 in multiple species. *Nucleic Acids Res* **47**: 3862–3874.  
26  
27

28  
29  
30 Martadinata, H., and Phan, A.T. (2013) Structure of human telomeric RNA (TERRA):  
31  
32 Stacking of two G-quadruplex blocks in K<sup>+</sup> solution. *Biochemistry* **52**: 2176–2183.  
33  
34

35  
36 Mcrae, E.K.S., Booy, E.P., Padilla-Meier, G.P., and Mckenna, S.A. (2017) On Characterizing  
37  
38 the Interactions between Proteins and Guanine Quadruplex Structures of Nucleic Acids. *J*  
39  
40 *Nucleic Acids* **2017**: 1–11.  
41  
42

43  
44 Mergny, J.-L., Phan, A.-T., and Lacroix, L. (1998) Following G-quartet formation by UV-  
45  
46 spectroscopy. *FEBS J* **435**: 74–78.  
47  
48

49  
50 Mergny, J.L., Li, J., Lacroix, L., Amrane, S., and Chaires, J.B. (2005) Thermal difference  
51  
52 spectra: A specific signature for nucleic acid structures. *Nucleic Acids Res* **33**: 1–6.  
53  
54

55  
56 Moll, K., Kaneko, A., Scherf, A., and Wahlgren, M. (2013) *Methods in malaria reserach\_6th*  
57  
58 *edition* . .  
59  
60

1  
2  
3 Paeschke, K., Simonsson, T., Postberg, J., Rhodes, D., and Lipps, H.J. (2005) Telomere end-  
4 binding proteins control the formation of G-quadruplex DNA structures in vivo. *Nat Struct*  
5  
6  
7  
8 *Mol Biol* **12**: 847–854.

9  
10  
11 Pagano, B., Margarucci, L., Zizza, P., Amato, J., Iaccarino, N., Cassiano, C., *et al.* (2015)  
12  
13 Identification of novel interactors of human telomeric G-quadruplex DNA. *Chem Commun*  
14  
15  
16 **51**: 2964–2967.

17  
18  
19  
20 Qureshi, M.H., Ray, S., Sewell, A.L., Basu, S., and Balci, H. (2012) Replication protein A  
21  
22 unfolds G-quadruplex structures with varying degrees of efficiency. *J Phys Chem B* **116**:  
23  
24  
25 5588–5594.

26  
27  
28 Rhodes, D., and Lipps, H.J. (2015) Survey and summary G-quadruplexes and their regulatory  
29  
30  
31 roles in biology. *Nucleic Acids Res* **43**: 8627–8637.

32  
33  
34 Salanti, A., Staalsoe, T., Lavstsen, T., Jensen, A.T.R., Sowa, M.P.K., Arnot, D.E., *et al.*  
35  
36 (2003) Selective upregulation of a single distinctly structured var gene in chondroitin  
37  
38 sulphate A-adhering Plasmodium falciparum involved in pregnancy-associated malaria. *Mol*  
39  
40  
41 *Microbiol* **49**: 179–191.

42  
43  
44  
45 Saul, A., Myler, P., Elliott, T., and Kidson, C. (1982) Purification of mature schizonts of  
46  
47 Plasmodium falciparum on colloidal silica gradients. *Bull World Health Organ* **60**: 755–759.

48  
49  
50  
51 Scherf, A. (1996) Plasmodium telomeres and telomere proximal gene expression. *Semin Cell*  
52  
53  
54 *Dev Biol* **7**: 49–57.

55  
56  
57 Scherf, A., Lopez-Rubio, J.J., and Riviere, L. (2008) Antigenic Variation in Plasmodium  
58  
59  
60 falciparum. *Annu Rev Microbiol* **62**: 445–470.

1  
2  
3 Silanes, I.L. De, D'Alcontres, M.S., and Blasco, M.A. (2010) TERRA transcripts are bound  
4 by a complex array of RNA-binding proteins. *Nat Commun* **1**: 1–10.  
5  
6  
7

8  
9 Sissi, C., Gatto, B., and Palumbo, M. (2011) The evolving world of protein-G-quadruplex  
10 recognition: A medicinal chemist's perspective. *Biochimie* **93**: 1219–1230.  
11  
12  
13

14  
15 Smargiasso, N., Gabelica, V., Damblon, C., Rosu, F., Pauw, E. De, Teulade-Fichou, M.-P.P.,  
16 *et al.* (2009) Putative DNA G-quadruplex formation within the promoters of *Plasmodium*  
17 *falciparum* var genes. *BMC Genomics* **10**: 1–12.  
18  
19  
20  
21

22  
23 Stanton, A., Harris, L.M., Graham, G., and Merrick, C.J. (2016) Recombination events  
24 among virulence genes in malaria parasites are associated with G-quadruplex-forming DNA  
25 motifs. *BMC Genomics* **17**: 1–16.  
26  
27  
28  
29  
30

31  
32 Takahama, K., Kino, K., Arai, S., Kurokawa, R., and Oyoshi, T. (2011) Identification of  
33 Ewing's sarcoma protein as a G-quadruplex DNA- and RNA-binding protein. *FEBS J* **278**:  
34 988–998.  
35  
36  
37  
38  
39

40 Thandapani, P., O'Connor, T.R., Bailey, T.L., and Richard, S. (2013) Defining the RGG/RG  
41 Motif. *Mol Cell* **50**: 613–623.  
42  
43  
44  
45

46  
47 Tonkin, C.J., Carret, C.K., Duraisingh, M.T., Voss, T.S., Ralph, S.A., Hommel, M., *et al.*  
48 (2009) Sir2 paralogues cooperate to regulate virulence genes and antigenic variation in  
49 *Plasmodium falciparum*. *PLoS Biol* **7**: 0771–0788.  
50  
51  
52  
53

54  
55 Vernick, K.D., and McCutchan, T.F. (1988) Sequence and structure of a *Plasmodium*  
56 *falciparum* telomere. *Mol Biochem Parasitol* **28**: 85–94.  
57  
58  
59  
60

1  
2  
3 Villar-Guerra, R. del, Trent, J.O., and Chaires, J.B. (2018) G-Quadruplex Secondary  
4 Structure Obtained from Circular Dichroism Spectroscopy. *Angew Chemie - Int Ed* **57**: 7171–  
5 7175.  
6  
7  
8  
9

10  
11 Wang, F., Tang, M.L., Zeng, Z.X., Wu, R.Y., Xue, Y., Hao, Y.H., *et al.* (2012) Telomere-  
12 and telomerase-interacting protein that unfolds telomere G-quadruplex and promotes  
13 telomere extension in mammalian cells. *Proc Natl Acad Sci U S A* **109**: 20413–20418.  
14  
15  
16  
17  
18

19  
20 Weighardt, F., Biamonti, G., and Riva, S. (1996) The roles of heterogeneous nuclear  
21 ribonucleoproteins (hnRNP) in RNA metabolism. *BioEssays* **18**: 747–756.  
22  
23  
24  
25

26  
27 Windgassen, M., and Krebber, H. (2003) Identification of Gbp2 as a novel poly(A)<sup>+</sup> RNA-  
28 binding protein involved in the cytoplasmic delivery of messenger RNAs in yeast. *EMBO*  
29 *Rep* **4**: 278–283.  
30  
31  
32  
33

34  
35 Wong, I., and Lohman, T.M. (1993) A double-filter method for nitrocellulose-filter binding:  
36 Application to protein-nucleic acid interactions. *Proc Natl Acad Sci U S A* **90**: 5428–5432.  
37  
38  
39

40  
41 Zhou, Z., Zhu, J., Zhang, L., Du, Y., Dong, S., and Wang, E. (2013) G-quadruplex-based  
42 fluorescent assay of S1 nuclease activity and K<sup>+</sup>. *Anal Chem* **85**: 2431–2435.  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



## Figures

### Figure 1: Biophysical characterization of the AP2\_G4.

a) Schematic representation of localization of the selected G4 motif (AP2\_G4) (blue box). The AP2\_G4 (GGGATTTGGGAGGGGGGG) is present 18 bp upstream of the transcription start site (TSS) of AP2 transcription factor (TF) (Pf3D7\_0934400) on the non-coding strand. b) Thermal differential spectra of AP2\_G4 and control mut\_G4. The TDS of AP2\_G4 displays the signature characteristics of G4 while mut\_G4 does not show any peaks for G4 formation. c) Circular dichroism (CD) spectra of AP2\_G4. The obtained CD spectra of AP2\_G4 indicates the formation of parallel G4. The AP2\_G4 (GGGATTTGGGAGGGGGGG) and mut\_G4 (GAGATTTGAGAGGGAGGG) oligonucleotides was used in these experiments.

### Figure 2: *In vitro* binding studies for PfGBP2 and G4 interaction

a) Quantification of PfGBP2 binding to AP2\_G4 and mut\_G4 by double filter-binding assay. The fraction of DNA bound to PfGBP2 is plotted against the PfGBP2 concentration. The graph was plotted for two replicates using linear regression (GraphPad Prism, San Diego, CA, USA).

### Figure 3: Genome-wide distribution of binding sites of PfGBP2

a) Immunofluorescence assay of HA-tagged PfGBP2. The PfGBP2 is expressed in both cytoplasm and nucleus, throughout the intraerythrocytic cycle of the parasites. DIC, differential interference contrast. HA tagged PfGBP2 were detected by  $\alpha$ -HA antibody 3F10. Nuclei is stained with

1  
2  
3 Hoescht. The scale bar = 1  $\mu\text{m}$ . b) Integrative genomic view of ChIP-Seq  
4 peaks with  $-\log(Q) > 50$  of PfGBP2 in all the chromosomes (1-14). The peaks  
5 (black lines) of PfGBP2 is distributed throughout all the chromosomes with  
6 enriched peaks near the chromosomal ends. The height of peak corresponds  
7 to the enrichment of the PfGBP2 peaks over IP input. c) The binding motif  
8 analysis of PfGBP2 ChIP-Seq dataset [ $\log(Q) > 50$ ] by MEME Suite. The most  
9 significant binding motif for PfGBP2 corresponds to the Plasmodium  
10 telomere consensus motif [GGGTT(T/C)A]. The e-value describes the  
11 statistical significance of the motif while the sites represents the number of  
12 sites that contributed to the construction of the motif.

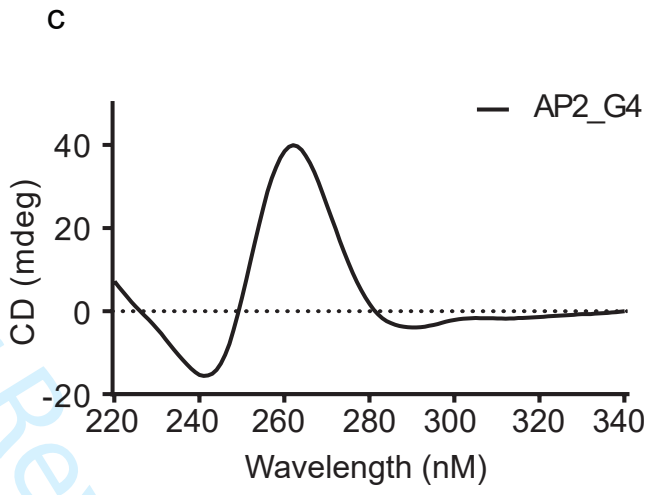
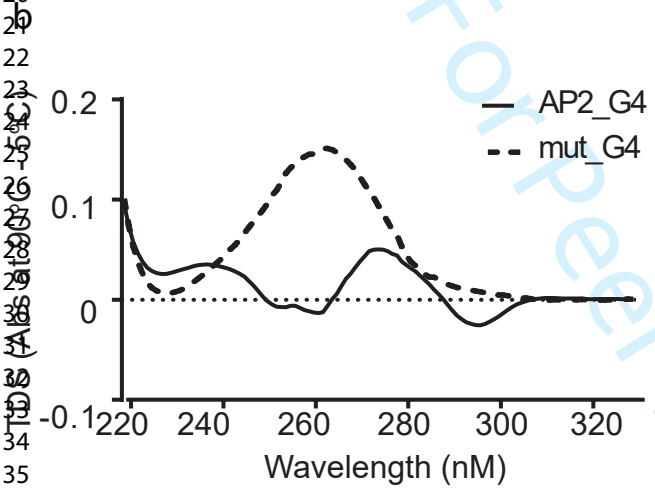
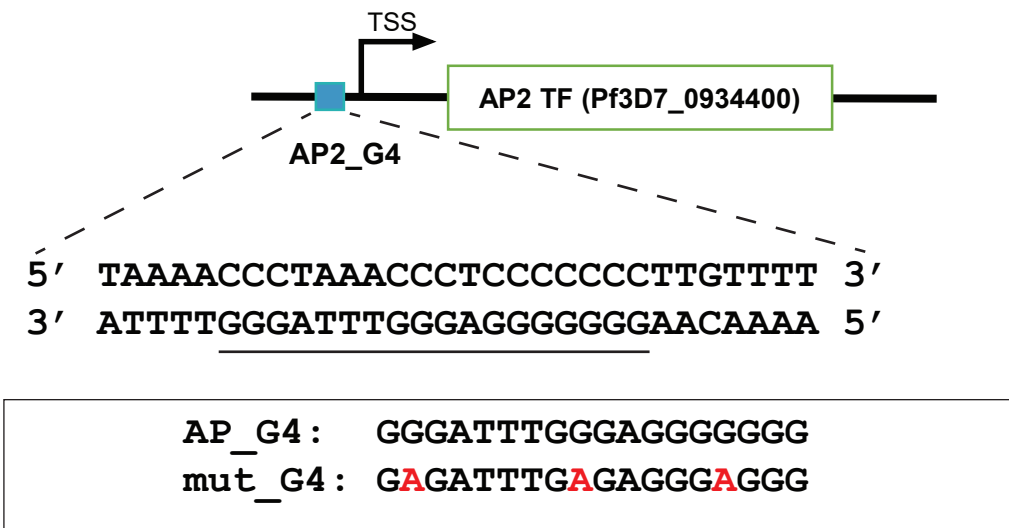
23  
24  
25  
26  
27  
28 **Figure 4: Effect of loss of PfGBP2 protein on parasite proliferation,**  
29 **telomere length and *var* gene expression**

30  
31  
32  
33 a) Growth phenotype assay of  $\Delta$ PfGBP2 ( $\Delta$ PfGBP2) and control (PfGBP2)  
34 parasite lines were performed over two consecutive cycles in the presence of  
35 1  $\mu\text{M}$  G4-stabilizing ligand- Pyridostatin (PDS). Upon PDS treatment,  
36  $\Delta$ PfGBP2 parasite line showed growth defect as compared to other control  
37 lines. Means and standard error are displayed for three independent replicates.  
38  
39 b) TRF Southern blot analysis of  $\Delta$ PfGBP2 parasite line ( $\Delta$ PfGBP2) and  
40 control (PfGBP2) lines for 30 and 60 days to examine the effect of loss of  
41 PfGBP2 on the telomere length of the parasites. The telomere length of the  
42  $\Delta$ PfGBP2 parasite line and control parasites are observed to be comparable.  
43  
44 c) qRT-PCR analysis of expression of *var* genes in the absence of PfGBP2  
45 protein ( $\Delta$ PfGBP2) for 30 days. The bar graphs display the relative  
46 expression of *var* gene expression, which is normalized using the  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

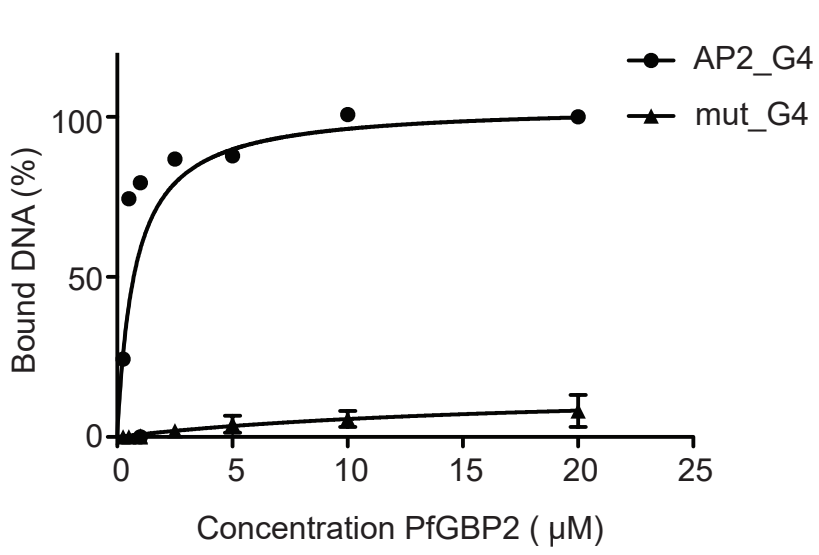
1  
2  
3 housekeeping gene- Serine-tRNA-ligase. The identical var gene is expressed  
4  
5 in both  $\Delta$ PfGBP2 and control parasite lines while slight derepression is  
6  
7 observed for the majority of the var genes. Data are represented as mean  $\pm$  SD  
8  
9  
10 (n =2).  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

For Peer Review

a  
1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

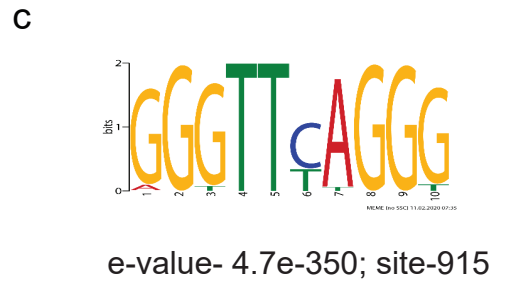
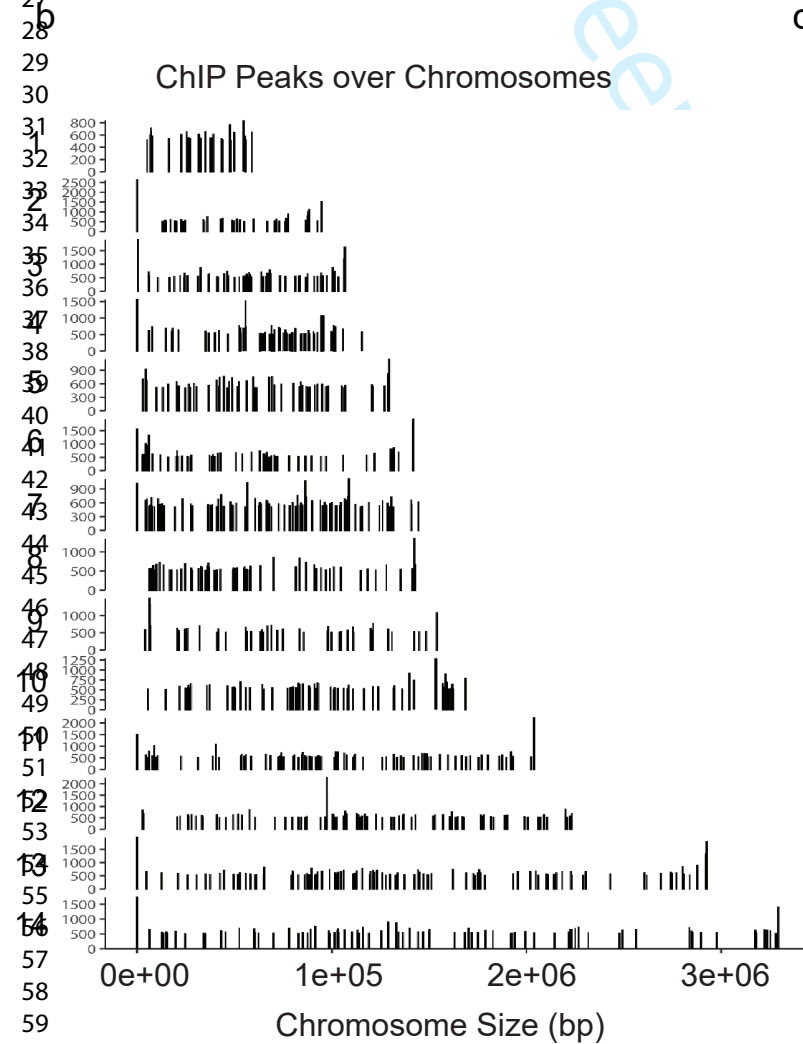
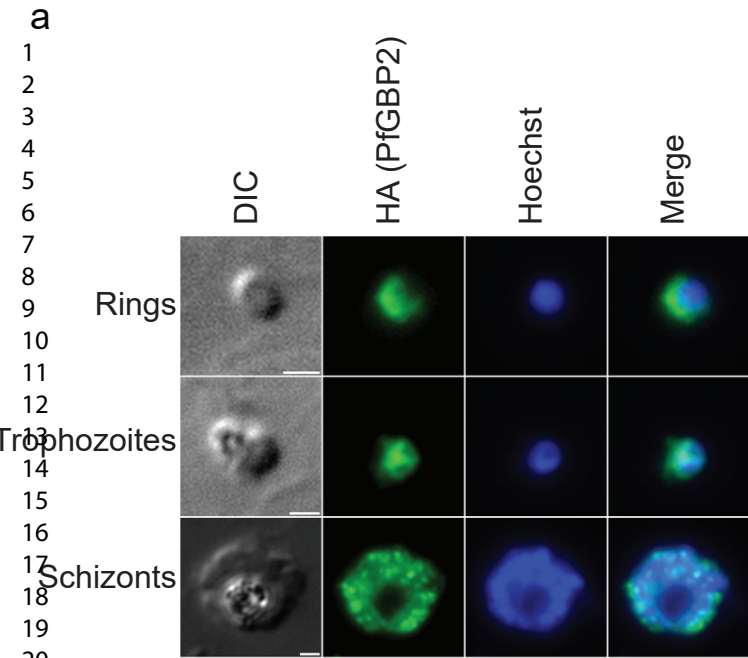


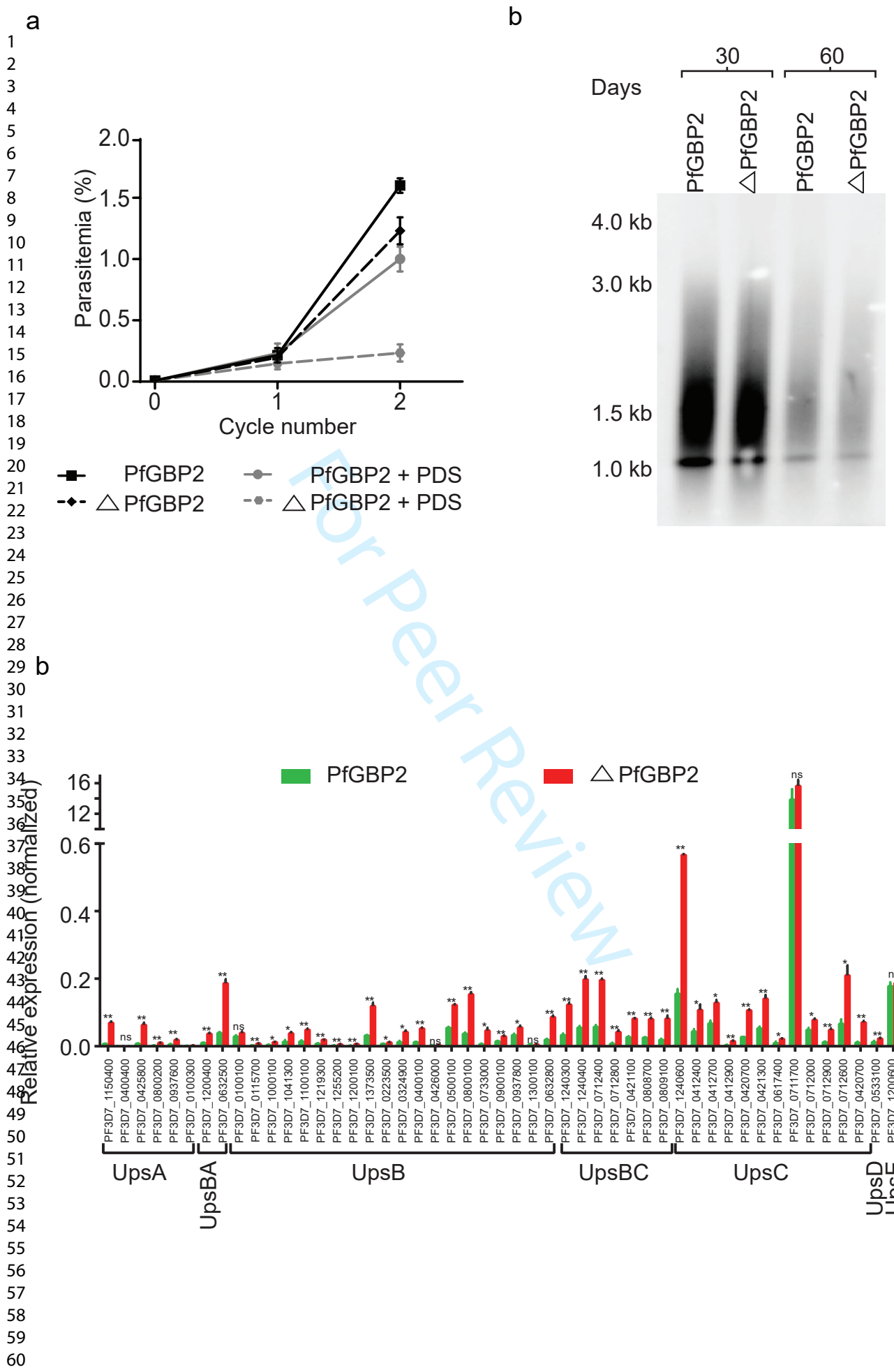
FOR PEER REVIEW



For Peer Review

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60





## Supplementary figures

### Figure S1: Thermal melting profile of the putative G4 motif (AP2\_G4)

The thermal melting profile of AP2\_G4 and mut\_G4 at 295 nM. The AP2\_G4 shows an inverted transition curve, characteristic of G4 structure while mut\_G4 does not form any G4.

### Figure S2: Identification of G4-BP using DNA pull-down based approach.

a) Outline of DNA pull-down approach. The nuclear lysate of *P. falciparum* is first incubated with the biotinylated mut\_G4 to pre-clear the lysate. This pre-cleared lysate is then incubated with the biotinylated AP2\_G4. The protein bound to these biotinylated AP2-G4 and mut\_G4 were eluted and analyzed by LC-MS/MS. Both the biotinylated oligonucleotides (AP2/mut\_G4) are heated at 95°C for 5 min followed by overnight cooling at room temperature to fold them into G4 or secondary structure. b) Pie chart of the gene ontology (GO) analysis of potential G4-BP that have nucleic acid binding function

### Figure S3: Bioinformatic analysis of domain organization of PfGBP2 and its homologs.

a) Schematic domain architecture of PfGBP2 protein along with the multiple sequence alignment with its homologs. The analysis shows that PfGBP2 contains highly conserved RNA-recognition motif domains (light grey blocks labelled as RRM1 and RRM2) and RGG motif (red bar). Aligned protein



1  
2  
3 sequence includes *Plasmodium falciparum* PfGBP2 (Q8IJX3),  
4  
5 *Chlamydomonas reinhardtii* CrGBP1p (Q39568), *Cryptosporidium parvum*  
6  
7 CpGBP (Q86PS0), *Saccharomyces cerevisiae* ScGBP2 (P25555), and *Homo*  
8  
9 *sapiens* HnRNP-A2 (P22626). b) Multiple sequence alignment of PfGBP2  
10  
11 and its homologs in *Laverania* species including *P. falciparum* 3D7  
12  
13 (Pf3D7\_1006800), *P. billcollinsi* G01 (PBILCG01\_1003800), *P. blacklocki*  
14  
15 G01 (PBLACG01\_1004900), *P. praefalciparum* G01 (PPRFG01\_1008000),  
16  
17 *P. reichenowi* CDC (PRCDC\_1006200), *P. adleri* G01 (PADL01\_1005300)  
18  
19 and *P. gaboni* G01 (PGABG01\_1004800). The analysis shows that the  
20  
21 PfGBP2 is highly conserved among *Laverania* species except *P. billcollinsi*  
22  
23 G01. Multiple sequence alignment is obtained from the MUSCLE algorithm  
24  
25 at [www.phylogeny.fr](http://www.phylogeny.fr). (\*) indicates identical amino acid while (;) indicates  
26  
27 conserved amino acid residues.  
28  
29  
30  
31  
32  
33  
34  
35

#### 36 **Figure S4: *In vitro* interaction studies for PfGBP2 and G4.**

37  
38 a) The coomassie stained SDS-PAGE and  $\alpha$ -His probed immunoblot of  
39  
40 recombinant PfGBP2 protein. The N-terminal His-tagged PfGBP2 protein  
41  
42 was produced in *E. coli* and purified by His-Pur Cobalt Spin column, followed  
43  
44 by size exclusion chromatography. The expected ~30.53 kDa PfGBP2 protein  
45  
46 is detected with  $\alpha$ -His antibody. The non-specific band detected by  $\alpha$ -His  
47  
48 belongs to *E. coli*, which is confirmed by MS analysis (data not shown). NEB  
49  
50 pre-stained protein ladder was used as protein size marker. b) Double-filter  
51  
52 binding assay for PfGBP2 and G4 interaction. The radiolabelled probes  
53  
54 (AP2\_G4 or mut\_G4) are incubated with increasing concentration of PfGBP2  
55  
56 (0-25  $\mu$ M). The top membrane (nitrocellulose) shows PfGBP2 bound to the  
57  
58  
59  
60

1  
2  
3 probes (AP2\_G4 or mut\_G4) while the bottom membrane (nylon+) shows  
4 unbound probes. The nitrocellulose membrane shows that PfGBP2 binds  
5 unbound probes. The nitrocellulose membrane shows that PfGBP2 binds  
6 preferentially to AP2\_G4 than mut\_G4.c) EMSA blot for PfGBP2 and G4  
7 interaction. The biotinylated AP2\_G4 or mut\_G4 was incubated with the  
8 increasing concentration of the PfGBP2 and subjected to run on 10 % native  
9 PAGE and visualized by chemiluminescence. The biotinylated AP2\_G4 (100  
10 nM) without PfGBP2 displays two different forms: folded G4 formation  
11 (upper band) and the linearized form (lower band) (lane 1) while an additional  
12 G4-PfGBP2 complex band appears with the increase in the concentration of  
13 PfGBP2 protein (0 - 818 nM) (lane 2-6). In contrast, PfGBP2 did not show  
14 any binding to mut\_G4 (100nM) even at the highest concentration of PfGBP2  
15 protein (lane 7-8).

### Figure S5. Confirmation of iKO-PfGBP2 transgenic line.

16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35 a) Schematic illustration of the strategy to generate iKO-PfGBP2 parasite  
36 line. In this approach, we have replaced the endogenous copy of PfGBP2 with  
37 a modified PfGBP2 gene cassette containing two loxp sites and a C-terminal  
38 hemagglutinin (HA) tag in *P. falciparum* p230p line. A donor plasmid  
39 containing recodonized PfGBP2 gene (white-shaded box) fused to a 3HA tag  
40 (green box) and flanked by homologous regions 1 and 2 (grey colored box).  
41 Two loxp regions are represented by red color arrowhead. This cassette is  
42 introduced into the WT locus of PfGBP2 gene in the *P. falciparum* p230p  
43 strain by Cas9, which recognizes the gRNA (red flash symbol) present in the  
44 third exon. The cas9 and gRNA is cloned into the pDC-Cas9 U6-hdfr/yfcu  
45 vector. The primers (blue colored arrows) are used for verification of the  
46 intergration of the cassette into the gene locus. b) The PCR analysis of the  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

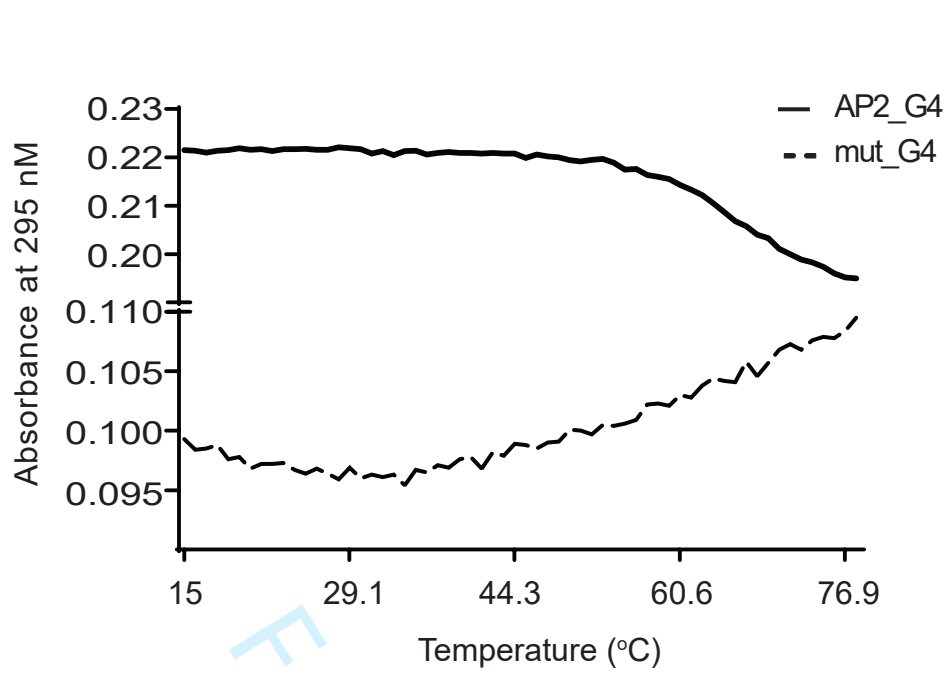
1  
2  
3 genomic DNA of iKO-PfGBP2 parasite line and the control parental Pf3D7  
4 p230p line (3D7-WT) with primer pairs 1+2 (box1) and 3+4 (box 2). The  
5  
6 analysis confirms the successful modification at endogenous locus of PfGBP2  
7  
8 exclusively in the iKO-PfGBP2 parasite line with the expected amplicons of  
9  
10 0.81 kb (box1) and 0.94 kb (box 2). The NEB 1kb ladder is used as DNA size  
11  
12 marker (lane1). c) Western blot analysis of the endogenous expression of  
13  
14 PfGBP2-HA protein in iKO-PfGBP2 parasite line with  $\alpha$ -HA antibody. The  
15  
16 immunoblot shows that the HAtagged PfGBP2 exclusively expressed in iKO-  
17  
18 PfGBP2 parasite line as compared to the parental Pf 3D7 p230p line. The anti-  
19  
20 H3 core is used as a loading control.  
21  
22  
23  
24  
25  
26  
27  
28

### 29 **Figure S6. Optimization of ChIP-Seq assay**

30  
31 a) Calculation of percentage of reads that map within called peaks obtained  
32  
33 in two experiments. In experiment 1 (exp 1) the PfGBP2 samples underwent  
34  
35 a pre-clear step while in experiment 2 (exp 2) an IgG control was prepared  
36  
37 independently from the IP sample. The bar graph shows that the more reads  
38  
39 maps within the called peaks in exp 1 as compared to exp 2, suggesting the  
40  
41 effective of pre-clear step to reduce the noise generated by the non-specific  
42  
43 binding of antibody. b) Intersection of ChIP-Seq peaks obtained in two  
44  
45 experiments (exp 1 and exp 2). The data shows that 89% of the exp 2 peaks  
46  
47 overlaps with the exp 1 peaks, confirming the reproducibility of the ChIP-Seq  
48  
49 replicates.  
50  
51  
52  
53  
54  
55

### 56 **Figure S7. Conditional deletion of PfGBP2 in iKO-PfGBP2 transgenic** 57 58 **line** 59 60

1  
2  
3 a) Schematic illustration of the strategy to induce the deletion of PfGBP2 in  
4 iKO-PfGBP2 transgenic parasite line. The gene is excised upon the  
5 rapamycin induction, which provokes the dimerisation of dicre-recombinase  
6 subunits (blue-orange shape). The dimerised DiCre-recombinase recognizes  
7 the loxp sites to promoter recombination that results into the excision of the  
8 respective gene. b) The PCR analysis of iKO-PfGBP2 parasite line treated  
9 with or without rapamycin using primer pair (1 and 4). The data shows the  
10 truncation of gene encoding PfGBP2-HA in rapamycin treated parasites while  
11 the gene remains intact in control parasites. c) Western blot analysis of lysates  
12 obtained from iKO-PfGBP2 parasite line treated with or without rapamycin  
13 to detect the expression of HA-tagged PfGBP2 using  $\alpha$ -HA antibody. The  
14 depletion of the PfGBP2 protein (~32.5 kDa) is detected in rapamycin (rap +)  
15 treated parasites (i.e.  $\rho$ PfGBP2 parasite line) as compared to the control (rap  
16 -) parasites.  $\alpha$ - H3 (~15.4 kDa) was used as loading control. d) Growth  
17 phenotype assay of parental Pf 3D7 p230p line and iKO-PfGBP2 parasite line  
18 was performed over two consecutive cycles with or without rapamycin  
19 treatment. The loss of PfGBP2 protein showed slight effect on the parasite  
20 growth as compared to the control and parental parasites. Means and standard  
21 error are displayed for three independent replicates.  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



For Peer Review

2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

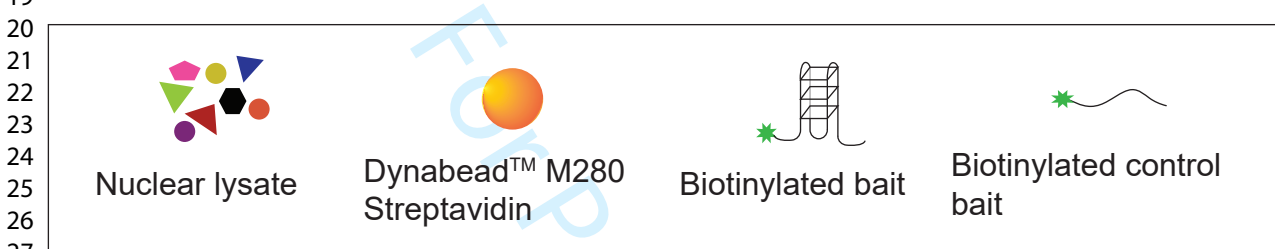
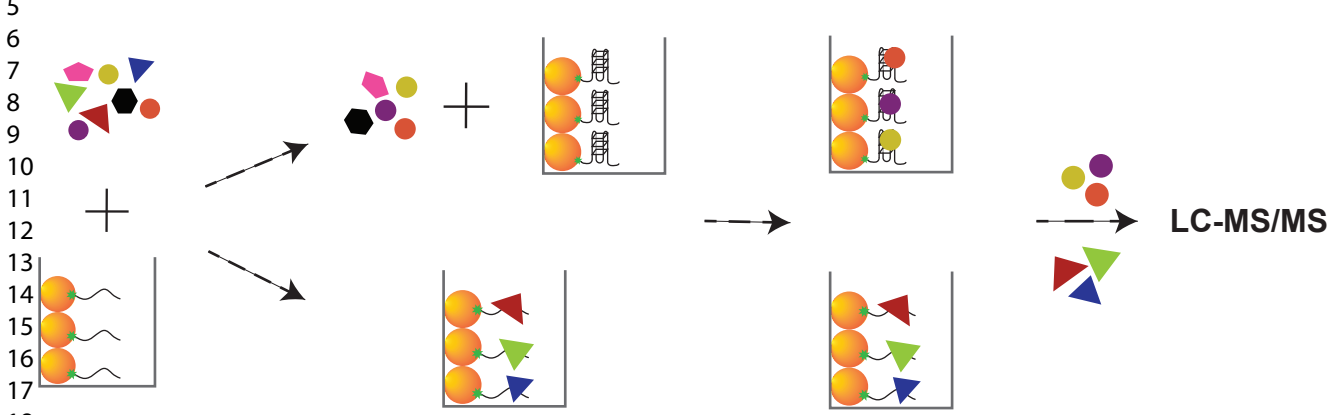
1a

2

3) Incubation of nuclear lysate with control bait

2) Incubation of pre-cleared nuclear lysate with bait

3) Elution of proteins bound to bait or control bait



3b

37

38

39

40

41

42

43

44

45

46

47

48

49

50

51

52

53

54

55

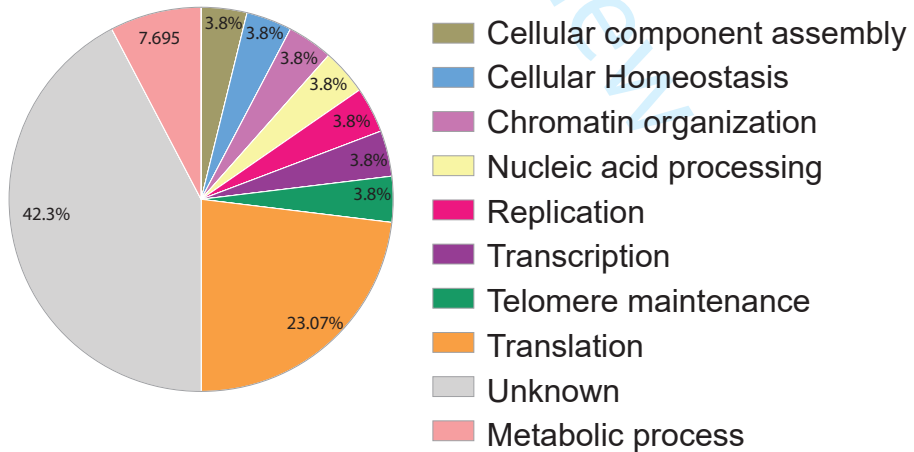
56

57

58

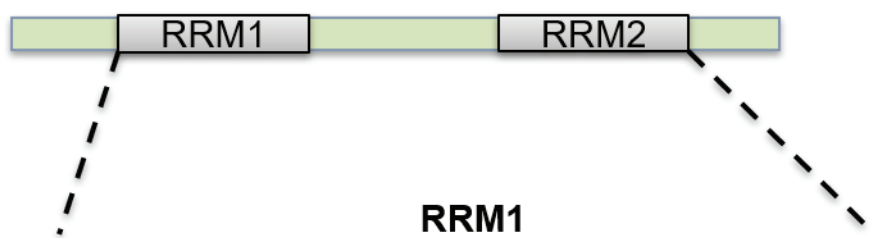
59

60



a  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
b  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

**PfGBP2**



**RRM1**

Pf_GBP2	CRVYVGNLPWKVTWPI LKNHMKKA GDVVRVDI FEDT-QGRSKGCGIVEYATYEEAQAIAISS
Cr_GBP1p	KRCFVGNLAWKT SWQDLKDKFREC GNVVYTNVMRDD-DGRSKGWGIVEFESPEEALHAIQ
Cp_GBP	CRVYVGNLPWKAKWHD LKDHMRQAGNVIRADVFEDV-VGRSRGCGVVEYSFPPEEAQRAIN
Sc_GBP2	FEVFIINLPYSMNWQS LKDMFKECGHVLRADELDF-NGFSRGGFVSIYPTEDEMIRAIID
Hs_HNRNP-A2	RKLFIGGLSFETTEES LRNYEYQWGLTDCVVMRDPASKRSRGGFVTFSSMAEVDAAAMA

. : . \* : . . \* : . : \* : . : \* \* \* \* : \* \* :

**RRM1**

Pf_GBP2	SLNSKLEDRLIFVREDREENSGNFEKRFKFNVRKDKFYESRRRRDYDYRKEYRRDDYRR
Cr_GBP1p	TLNGAELGGRRLIVREDREDRDVQADGAPGAPAERAPR---A-----R
Cp_GBP	ELNNTTLLDRLIFVREDREDESSRYGRRSN-----
Sc_GBP2	TFNGMEVEGRVLEVRGRFNKRKNNDRYNQ--RREDL-EDT-RGT-----
Hs_HNRNP-A2	ARP-HSIDGRVVEPKRAVA-----

: . \* : . :

**RRM2**

Pf_GBP2	DFRDEFRRGGGEFRRDYRRDEFRRGGGEFRSSSKRNCTLIVYNLPPQVTWKE LKDLFRK
Cr_GBP1p	-----GAGRGAANGGRGAGR-GGRGAGAPEATGESSGLQVVVQGI PWAYTWRELKDMFAE
Cp_GBP	-----KWNNRGYGMRTTRTHAPRPPLEENKKGQV FVTNLAWKTTQEDLAKAFNE
Sc_GBP2	---EPGLAQDAA-VHIDETAAKFT---EGVNPGGDRNCFIYCSNLPFSTARSDFDLFGP
Hs_HNRNP-A2	---REESG-----KP--GAHVTVKKLFVGGIKEDTEEHLRDIYFEE

: . : \* . \*

**RRM2**

Pf_GBP2	HGRVVRADLKNE-DNSSKELIGVIMENEYEAKN-AIDALNFCNFDGYILKVNYENNE
Cr_GBP1p	VGGVDRADVVVG-YDGRSRGYGTVKFTTKEAAEA-AVARYHESELEGRRLAVFIDRYQ
Cp_GBP	IGPLESCEVYFYF-EDGRSRGIATIVFTDPNHAQL-AVEKLNDRIDGREILVRIDQ--
Sc_GBP2	IGKINNAELKQP-ENGQPTGVAVVEYENLVDAF-CIQKLNINYGGCSLQISYARRD
Hs_HNRNP-A2	YGKIDTIEIITDRQSGKKRGGFVTFDDHDPVDKIVLQKYHTINGHNAEVRKALSQRE

\* : : : . . . : . : : : : . : . :

**RRM1**

Pf3D7_1006800	CRVYVGNLPWKVTWPI LKNHMKKA GDVVRVDI FEDTQGRSKGCGIVEYATYEEAQAIAISS
PBILCG01_1003800	CKLIFVGNLSSNAREEDVRRRFEQYGDIMHMQW-----KRRFAFIEY--YKASH----A
PBLACG01_1004900	CRVYVGNLPWKVTWPI LKNHMKKA GDVVRVDI FEDTQGRSKGCGIVEYATYEEAQAIAISS
PPRF01_1008000	CRVYVGNLPWKVTWPI LKNHMKKA GDVVRVDI FEDTQGRSKGCGIVEYATYEEAQAIAISS
PRCDC_1006200	CRVYVGNLPWKVTWPI LKNHMKKA GDVVRVDI FEDTQGRSKGCGIVEYATYEEAQAIAISS
PADL01_1005300	CRVYVGNLPWKVTWPI LKNHMKKA GDVVRVDI FEDTQGRSKGCGIVEYATYEEAQAIAISS
PGABG01_1004800	CRVYVGNLPWKVTWPI LKNHMKKA GDVVRVDI FEDTQGRSKGCGIVEYATYEEAQAIAISS

\* : : \* \* \* : . . . : : \* \* : : : : :

Pf3D7_1006800	LNSKLEDRLIFVREDREENSGNFEKRFKFNVRKDKFYESRRRRDYDYRKEYRRDDYRRD
PBILCG01_1003800	LNALEKENGKMPFGEE LNVQEHQYNMNSYG-----YKNDNRSYPHFK-----
PBLACG01_1004900	LNSKLEDRLIFVREDREENSGNFEKRFKFNVRKDKFYESRRRRDYDYRKEYRRDDYRRD
PPRF01_1008000	LNSKLEDRLIFVREDREENSGNFEKRFKFNVRKDKFYESRRRRDYDYRKEYRRDDYRRD
PRCDC_1006200	LNSKLEDRLIFVREDREENSGNFEKRFKFNVRKDKFYESRRRRDYDYRKEYRRDDYRRD
PADL01_1005300	LNSKLEDRLIFVREDREENSGNFEKRFKFNVRKDKFYESRRRRDYDYRKEYRRDDYRRD
PGABG01_1004800	LNSKLEDRLIFVREDREENSGNFEKRFKFNVRKDKFYESRRRRDYDYRKEYRRDDYRRD

\*\* : \* : : \* . \* : : . : : . . . \* . . \* : . :

Pf3D7_1006800	FRRDEFRRGGGEFRRDYRRDEFRRGGGEFRSSSKRNCTLIVYNLPPQVTWKE LKDLFRKH
PBILCG01_1003800	-----MSRNYSPPNFGNESIE-RRNSLR---IVVKNIDEKASWQDLKDFGREV
PBLACG01_1004900	FRRDEFRRGGGEFRRDYRRDEFRRGGGEFRSSSKRNCTLIVYNLPPQVTWKE LKDLFRKH
PPRF01_1008000	FRRDEFRRGGGEFRRDYRRDEFRRGGGEFRSSSKRNCTLIVYNLPPQVTWKE LKDLFRKH
PRCDC_1006200	FRRDEFRRGGGEFRRDYRRDEFRRGGGEFRSSSKRNCTLIVYNLPPQVTWKE LKDLFRKH
PADL01_1005300	FRRDEFRRGGGEFRRDYRRDEFRRGGGEFRSSSKRNCTLIVYNLPPQVTWKE LKDLFRKH
PGABG01_1004800	FRRDEFRRGGGEFRRDYRRDEFRRGGGEFRSSSKRNCTLIVYNLPPQVTWKE LKDLFRKH

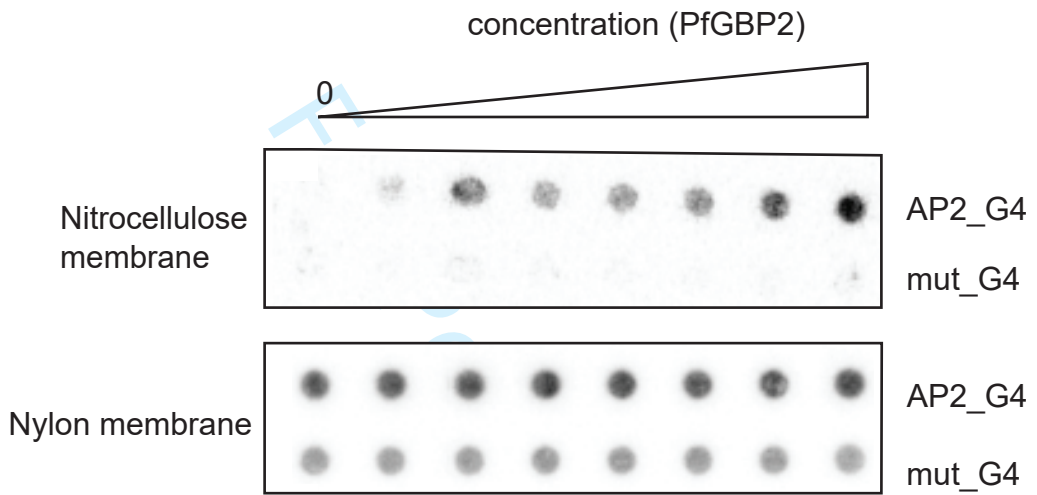
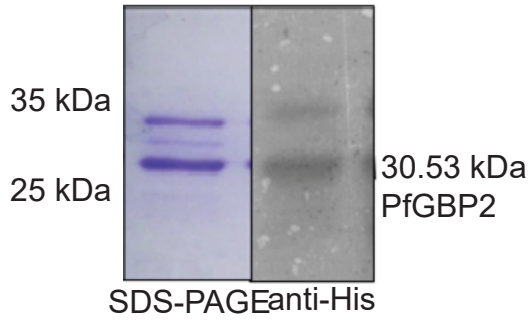
: \* \* : \* . \* \* : : \* \* : : : \* \* : :

**RRM2**

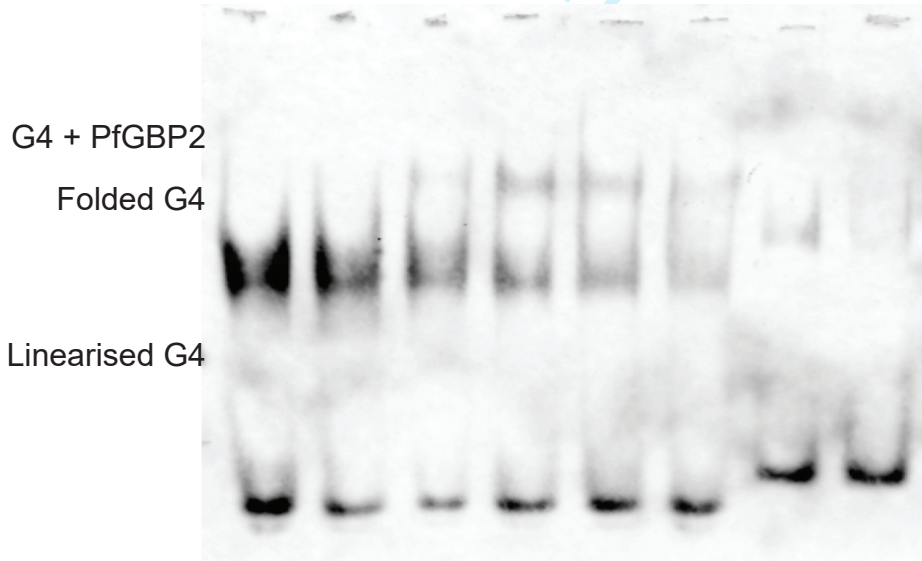
Pf3D7_1006800	GRVVRADLKNE DNSSKE LIGVIMENEYEAKN AIDALNFCNFDGYILKVNYENNE
PBILCG01_1003800	GSVSYANIV-DDYHSKEKFGIIEFYNHENAKDAI-----
PBLACG01_1004900	GRVVRADLKNE DNSSKE LIGVIMENEYEAKN AIDALNFCNFDGYILKVNYENNE
PPRF01_1008000	GRVVRADLKNE DNSSKE LIGVIMENEYEAKN AIDALNFCNFDGYILKVNYENNE
PRCDC_1006200	GRVVRADLKNE DNSSKE LIGVIMENEYEAKN AIDALNFCNFDGYILKVNYENNE
PADL01_1005300	GRVVRADLKNE DNSSKE LIGVIMENEYEAKN AIDALNFCNFDGYILKVNYENNE
PGABG01_1004800	GRVVRADLKNE DNSSKE LIGVIMENEYEAKN AIDALNFCNFDGYILKVNYENNE

\* \* \* : : \* \* \* : : : \* : \* \* \* \* :

1a  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



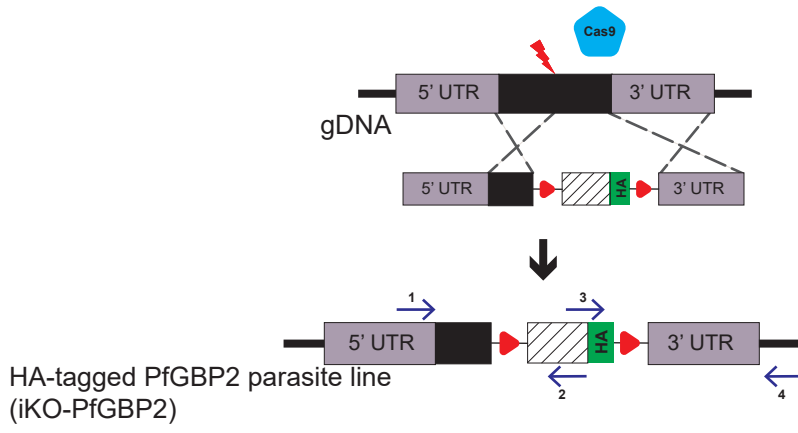
AP2_G4 (100nM)	+	+	+	+	+	+	-	-
mut_G4 (100nM)	-	-	-	-	-	-	+	+
PfGBP2 (nM)	0	65	196	262	327	818	0	818



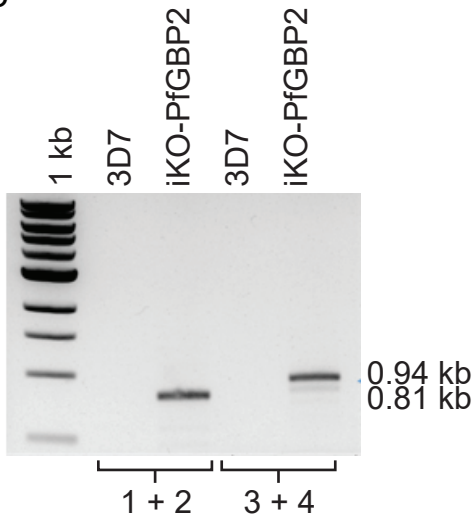


1a

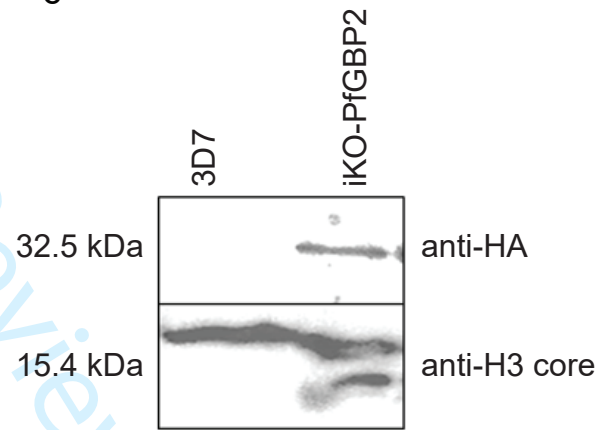
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



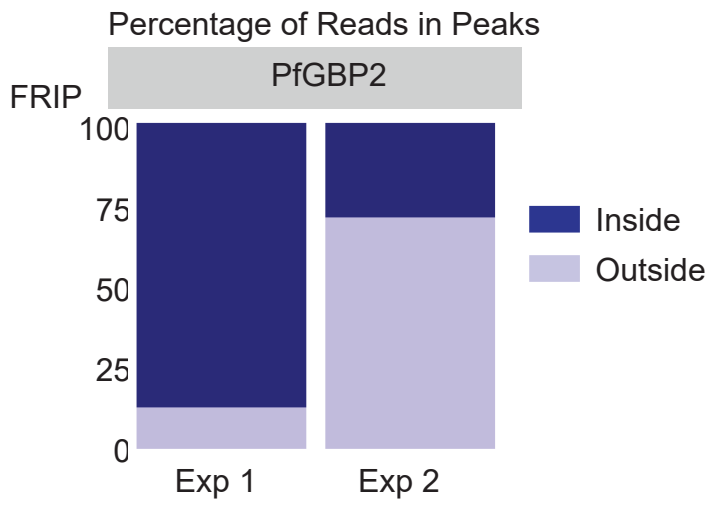
b



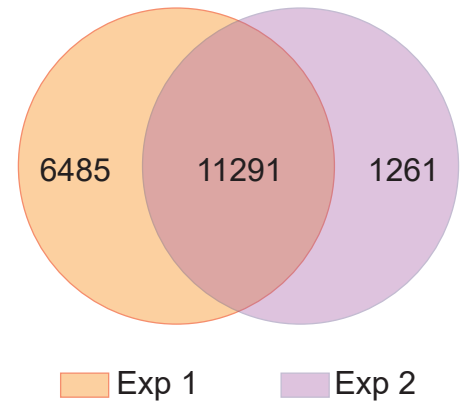
c



a  
1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



b



For Peer Review

a  
1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
b  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
c  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

