



Supporting Online Material for

Regulation of Sexual Development of *Plasmodium* by Translational Repression

Gunnar R. Mair, Joanna A. M. Braks, Lindsey S. Garver, Joop C. A. G. Wiegant,
Neil Hall, Roeland W. Dirks, Shahid M. Khan, George Dimopoulos, Chris J. Janse,
Andrew P. Waters*

*To whom correspondence should be addressed. E-mail: waters@lumc.nl

Published 4 August 2006, *Science* **313**, 667 (2006)
DOI: 10.1126/science.1125129

This PDF file includes:

Materials and Methods
Figs. S1 to S9
Tables S1 to S4
References

Material and Methods

Oligonucleotides. For a complete list of sequences please refer to Supplementary Table S4.

Characterization of *pbdozi*. In the proteome of female gametocytes (1) peptides were present that recognised the *Plasmodium berghei* gene model PB000603.01.0 which is annotated as a 'putative ATP-dependent RNA helicase' (www.genedb.org). Based on its null mutant phenotype we propose to name the protein development of zygote inhibited (DOZI). *pbdozi* is part of the ~32kb large contig5325 (www.sanger.ac.uk). Through Reverse Transcription of total RNA extracted from Nycodenz-gradient purified gametocytes with oligo d(T) followed by PCR with primers 2290×2340 and subsequent sequence analysis of TOPO-cloned amplicons the 2-exon/single intron structure of *pbdozi* was confirmed and the ORF extended at the 5' end. The *pbdozi* ORF comprises 1302 bps and encodes a protein with a predicted molecular weight of 49.4 kDa.

Multiple sequence alignments. *Plasmodium falciparum* RNA helicases were retrieved from www.plasmodb.org combining two text queries (*geneName query rna helicase from Gene Name*) union (*geneName query rna helicase from Curated Annotation*) yielding 33 putative RNA helicases, some of them truncated. S.

cervisiae RNA helicases are from www.yeastgenome.org. ClustalW and BOOTSTRAP analysis were done at <http://www.ddbj.nig.ac.jp/search/clustalw-e.html>. Output trees were drawn with tree-view from taxonomy.zoology.gla.ac.uk/rod/treeview.html.

Generation of parasite line, PbDOZI::GFP, expressing GFP-tagged DOZI.

C-terminal GFP-tagging of DOZI was performed using a construct that integrates through single cross-over homologous recombination into the *pbdozi*-locus and contains the *tgdhfr/ts* selectable marker. Primers 2290×2381*Bam*HI were used to amplify the 2153 bps targeting region of *pbdozi*. The amplicon was cut with *Eco*RV and *Bam*HI, cloned into pL0110 followed by removal of the SSU yielding pL1059. The plasmid was linearised with *Spe*I and 5 µg were used for transfection of *P. berghei* clonal line cl15cy1 of the ANKA strain using published methods (2). Transfected parasites were cloned by limiting dilution and correct integration of the construct was analysed by FIGE and in 5 independent, GFP-expressing clones (683cl1-5) by the following PCRs: 2290×1753 (5' integration; amplicon size = 2321 bps), 1419×2241 (3' integration; 4078 bps), 1419×1420 (*tgdhfr/ts*; 628 bps), 2290×2241 (wild type *pbdozi*; 3117 bps) and control primers 537×538, amplifying a fragment of *p28*. Cloned line 683cl4 was used for all further analyses and named PbDOZI::GFP.

Reverse transcriptase (RT) PCR analysis of *pbdozi::gfp*. To confirm use of splice sites in the fusion gene *pbdozi::gfp* identical to the ones employed in

pbdzi, RT of 5 µg gametocyte total RNA with oligo d(T) was performed followed by PCR with primers 2339 and 1214, TOPO-cloning of the product and sequence analysis. As a control, genomic DNA was amplified with the same primer combination.

***In vitro* fertilisation/zygote development assays.** Standard *in vitro* assays for ookinete development of *P. berghei* were used to analyse gamete formation fertilization and development of zygotes into ookinetes (1,3). Analysis of cross-fertilization between gametes of different mutant parasite lines was performed in these assays as described previously (1). Male gamete production is analysed by counting exflagellating male gametocytes in a cell counter. Female gametes, zygotes and ookinetes are counted in Giemsa stained thin blood films of the culture material. Zygotes are distinguished from unfertilized females in Giemsa stained slides as described (2). Ookinete production (ookinete conversion rates) is defined as the percentage of female gametes that develop into mature ookinetes.

Immunoprecipitation of complexes containing GFP::DOZI.

Immunoprecipitations (IP) were done essentially as described (4). Purified gametocytes (see above) obtained from 1ml of blood from one mouse (5) were washed 3 times with ice-cold 'enriched' PBS, lysed in 250 µl NET-2 (50 mM NaCl, 150 mM TrisHCl [pH7.4], 0.5% NP40) containing 2 mM DTT and supplemented with protease inhibitors (Complete Mini; Roche) and RNasin (2 units/µl;

PROMEGA). After lysis the extract was spun for 10 minutes at 14000 RPM to remove the pelleted, insoluble fraction. Fifty μ l extract were used per IP with anti-GFP-antibody (1 μ g, Roche), anti-c-myc-antibody (2.5 μ l, Sigma) or beads-only and incubated on ice for 1 hour with occasional mixing. Recovery of proteins was done with protein G-sepharose beads (20 μ l packed bead volume per IP; a kind gift from Dr. Cornelis H. Hokke) in a total volume of 200 μ l for 1 hour on ice with occasional mixing. IPs were washed 4 times with 400 μ l NET-2 and divided in 2 equal volumes. One pellet was resuspended in SDS-PAGE loading buffer and used for western analysis; a second pellet was resuspended in 500 μ l TRIzol reagent for RNA isolation. Equivalent amounts of protein lysate or total RNA were loaded.

Generation of *pbdozi* null mutant *P. berghei* parasites. *Pbdozi* null mutants were generated by double cross-over homologous recombination replacing the orf of *pbdozi* with the *tgdhfr/ts* selectable marker. The 5' targeting region (TR) (PCR 2238EcoRI \times 2239HindIII) and the 3' TR (PCR 2240SmaI \times 2241BamHI) were cloned sequentially into plasmid pL0001, upstream and downstream, respectively, of *tgdhfr/ts*, resulting in construct pL1052. This plasmid was linearised with *KpnI* and *KspI*, and 5 μ g were used for transfection of *P. berghei* (cl15cy1 of the ANKA strain) in two independent experiments (exp. 560cl1-5, and 611cl1-4) using published methods (2,6). Transfected parasites were cloned by limiting dilution and correct integration of the construct was analysed from both experiments by FIGE and the following PCRs: 2290 \times 695 (5' integration; 890 bps), 1662 \times 2387 (3' integration; 1800 bps), 1419 \times 1420 (*tgdhfr/ts*; 628 bps), 2339 \times 2381 (wild

type *pbdozi*; 1486 bps) and control primers 537×538, amplifying a fragment of *p28*. Cloned line 560cl2 and 611cl1 were used for all further experiments and named *pbdozi* null mutants.

The absence of transcript in the *pbdozi* null mutants was confirmed by northern analysis of RNA from gametocytes using as a probe the PCR-amplicon 2339×2381 representing the entire ORF of *pbdozi*.

Localisation of DOZI and *p25* or *p28* transcripts in female gametocytes by FISH. Fluorescent In Situ Hybridisation (FISH) to detect *p25* and *p28* transcripts was done as described (7). Probes were generated by PCR for *p25* (1462×1463) and *p28* (537×538), and direct Cy3-labeling was achieved by nick-translation of 1 µg of template DNA. For co-localization of mRNAs with DOZI::GFP, slides containing gametocytes were first processed for FISH, followed by incubation with a mouse monoclonal anti-GFP-antibody and a fluorescein-labeled goat-anti-mouse antibody. Parasites were analyzed using an Epi-fluorescence microscope (Leica DM-RA, Leica, Wetzlar, Germany) equipped with a cooled CCD camera (Photometrics, Tuscon, Arizona). In addition, live DOZI::GFP expressing gametocytes were observed and photographed using GFP filter settings on a Leica fluorescence MDR microscope and a DC500 digital camera.

Immunoblots. Antibodies used in Western blot analyses were: a monoclonal mouse anti-GFP antibody (Roche; 1:10000), a polyclonal rabbit anti-P47 antiserum (8), a polyclonal rabbit anti-*Artemia*-eEF1A antiserum (9) (1:5000),

rabbit polyclonal anti-P28 antisera (881 and 882; 1:200) and mouse monoclonal *P.yoelii* anti-P25 (1:200; a generous gift from Takafumi Tsuboi). Secondary antibodies were an anti-rabbit HRP-linked antibody (donkey; Amersham Biosciences) or an anti-mouse HRP-linked antibody (sheep; Amersham Biosciences). Visualisation was with the ECL Plus Western Blotting Detection System from Amersham Biosciences.

Generation of probes for Northern analysis. DNA templates used in KLENOW driven labeling with random hexamers and α -³²P-ATP were generated by PCR from genomic DNA* or cDNA[#] for the following genes: *pbdozi*[#], *eef1a**, *gfp**, *nima kinase*[#] (*nek4*), *p25**, *p28**, *p47**, *pb000245.02.0** (www.genedb.org), *pb000633.00.0** (www.genedb.org), *pbccp1** and *warp**. Oligonucleotides 87R and L647 were end-labeled with γ -³²P-ATP and T4 polynucleotide kinase.

Microarray analysis. Sixty-mer oligonucleotides for the 5283 predicted *P. berghei* transcripts were designed using the Oligo Picky software (10). Oligonucleotide sequences were designed to be complementary to regions within 1 kb from the 3' UTR of transcripts and had a minimal sequence identity overlap with non-target transcript sequences. Microarrays were constructed through *in situ* synthesis of oligonucleotides on glass slides by Agilent Technologies (www.agilent.com).

Purified gametocytes were obtained from 2 mice infected with wild type *P. berghei* (cl15cy1), and from 2 mice infected with the *pbdozi* null mutant (560cl2)

as described (5). RNA from all parasites was extracted with TRIzol according to the manufacturer's instructions. Fluorochrome cRNA labeled probes were synthesized from 2 µg RNA using the Agilent Technologies low input linear amplification RNA labeling kit according to the manufacturer's instructions. Probe quantity was estimated with a Beckman DU640 spectrophotometer and 16 hours hybridizations were done with the Agilent Technologies *in situ* hybridization kit according to manufacturer's instructions. After washes, microarrays were scanned with an Axon GenePix 4200AL scanner. Scan images were analyzed and raw data for Cy5 and Cy3 signal and ratio values were generated with Genepix Pro 6.0 software (Axon Instruments, Union City, CA). The minimum signal intensity was set to 100 fluorescent units and the signal to background ratio cut-off was set to 2.0 for both Cy5 and Cy3 channels. Two independent *pbdozi* null mutant clones were assayed in 2 and 3 technical replicates, respectively. The background subtracted median fluorescent values were normalized according to a LOWESS normalization method and Cy5/Cy3 ratios from replica assays were subjected to T-Tests with a 0.05 P-value using the TIGR MIDAS and MeV software (www.tigr.org). Microarray gene expression data are presented in Supplementary Table S1.

Microarray assayed gene expression was further validated with quantitative RT-PCR for 15 genes and showed a high degree of correlation (Pearson correlation coefficient $p = 0.86$; best-fit linear-regression $R^2 = 0.801$; and the slope of the regression line $m = 1.10$) for 15 tested genes (Supplementary Figure S8; Supplementary Table S3).

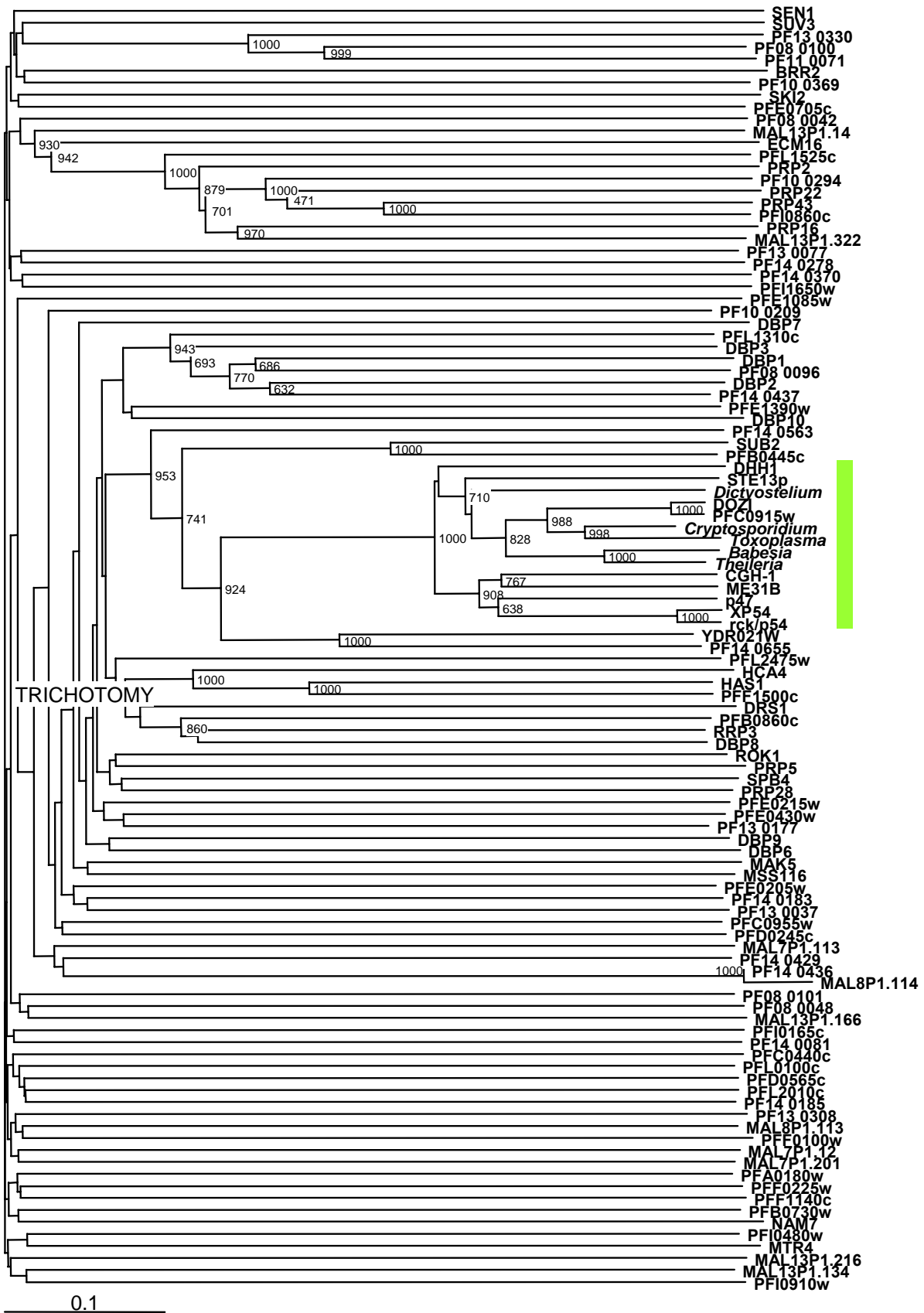
Reverse Transcription quantitative PCR (RTqPCR). RNA samples were reverse-transcribed using Superscript III (Invitrogen) with random hexamers. Real-time quantification was performed using the QuantiTect SYBR Green PCR Kit (Qiagen) and ABI Detection System ABI Prism 7000. All PCR reactions were performed in triplicate. Specificity of the PCR reactions was assessed by analysis of melting curves for each data point. A non-differentially expressed gene, pb301087.00.0, and comparative data from northern blots were used for normalization of cDNA templates.

Reverse Transcription-PCR, RTqPCR and FISH analyses of RNAs recovered from immunoprecipitations (IP). 5 uls total RNA from each IP-eluate, the corresponding wash supernatants and the input material were treated in the presence of 40 units RNasin (PROMEGA) for 50 minutes at 37°C with 1 unit DNaseI (PROMEGA) followed by addition of 2 uls stop solution and a 10 minute incubation at 65°C. To each sample was then added 100 pmoles of oligo d(T) primer and incubated at 65°C for 10 minutes followed by cooling on ice for 2 minutes. Samples were then divided in 2 aliquots (RT-negative and RT-positive), and only to the RT-positive sample added SuperscriptIII Reverse Transcriptase (200 units, GIBCO). RT-negative samples were treated otherwise exactly as RT-positive samples. RT was carried out at 46°C for 60 minutes followed by enzyme inactivation at 70°C for 15 minutes; finally 2 units RNaseH (GIBCO) were added for 20 minutes at 37°C. For PCR 2 uls were diluted $1/50$ in water and 2.5 uls were

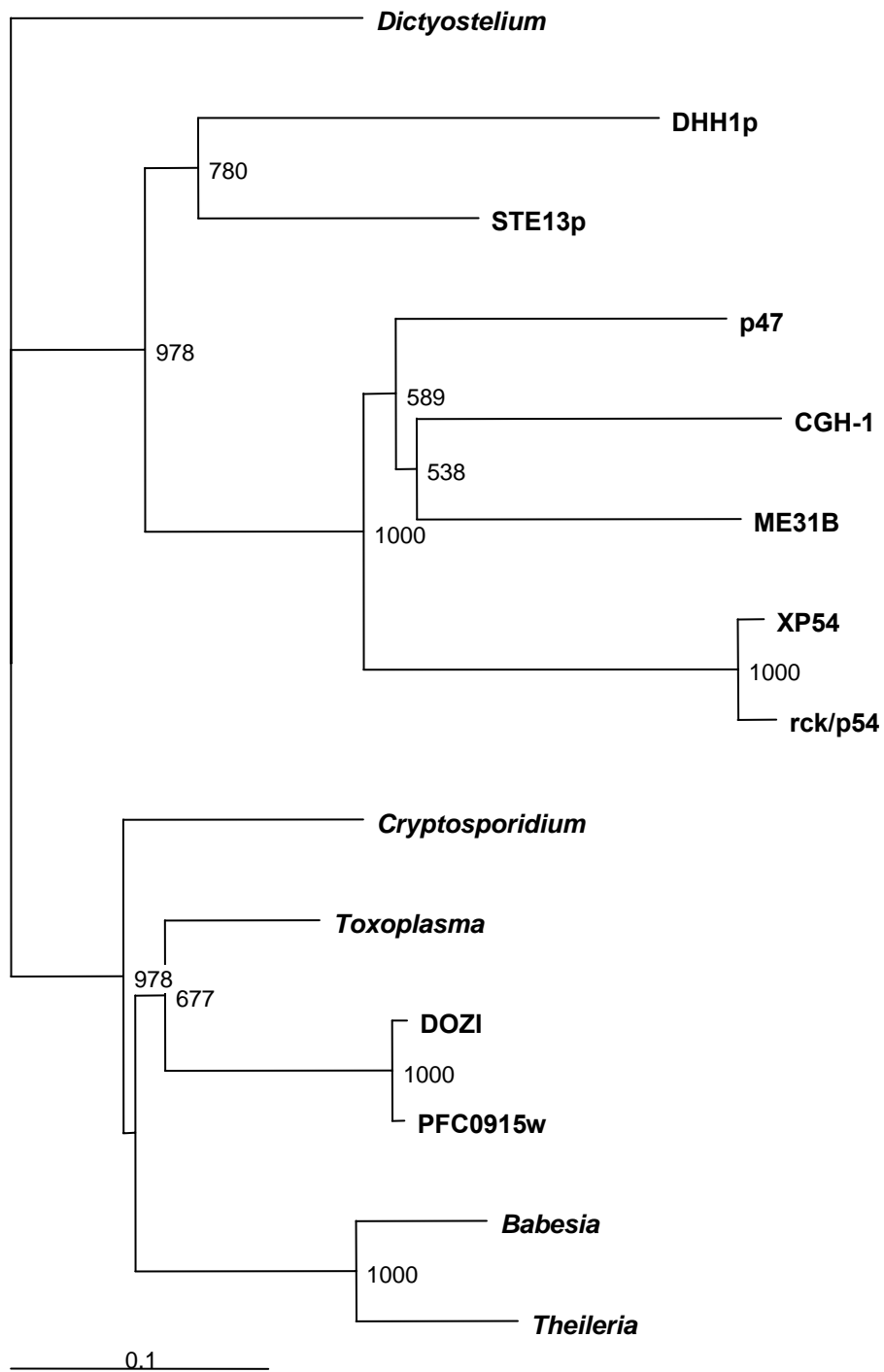
used in PCR with primers shown in Table S4F. No amplicons were identified with any of the RT-negative templates.

RT-qPCR assays were performed as described above; primers used are shown in Table S4E.

Templates for Cy3 or Cy5-labelling of probes used in FISH analysis were generated by PCR using cDNA as template (for primers see Table S4F). Labelling of probes by nick-translation and FISH of DOZI::GFP parasites followed by IFA detection of DOZI::GFP was as described above, except that label incorporation for the *pb300567.00.0*-probe was achieved during PCR. Each amplicon was also TOPO-TA cloned and sequenced.



Supplementary Figure S1 Neighbor-joining tree of 33 *S. cerevisiae* and 32 *P. falciparum* putative RNA helicases together with DOZI and a selection of RNA helicases known to be involved in translational repression, and some high scoring homologs. The bootstrap values for 1000 trials are shown.



Supplementary Figure S2 Neighbor-joining tree of DOZI and a selection of RNA helicases known to be involved in translational repression, and some high scoring homologs. The bootstrap values for 1000 trials are shown.

DHH1p 1 -----MGSINNNTNNSNTDLDLDRWKTALNPKKDTTPQTDVLTNK
 STE13p 1 -----MAESLIQKLENANLNDRESEKGGQAKAQDMDRPFKEDVTKTK
 DOZI 1 -----MSYKTNCVANSANTYALNNSNNLNKIEDNVIPDPAWKKKLEBPKDPMKTEDVTKTK
 PFC0915w 1 -----MSYKTNCVANSANTNTLNSSSNYNKIDDNILDEBPKKKLEBPKDPMKTEDVTKTK
 Cryptosporidium 1
 Toxoplasma 1 MSAAYYTRRPLGSLGSPYPQSQNSGGERNANQGQGPAPYGYPPQQRMQHFSQPQNSVPSQPANPAHVDPNSNPEGDGQWKSXVLPKPKDNVTKTEDVTKTK
 Babesia 1 -----MVKDRDTKPSDSAGVSDRSNGRRSSGDRFVDPWKKKLEBPKDPMKTEDVTKTK
 Theileria 1 -----MSLSLSDTSCKNKNTSVDPPTWKRDLDKNADKRYKTEDVTKTK
 Dictyostelium 1 -----MNSQONKLPQQQPPSSIVDDDNWKSQKLEPPEBRKOTFEDVTKTK
 CGH-1 1 -----MSGAEQQQIVPANNGBENWKAAGLNLEPAKDRRFKADVTDTK
 ME31B 1 -----MMTEKLNSGHTNLTSGKGIINDLQIAGNTSDDMGQWKSXVLPKPKDNVTKTEDVTKTK
 XP54 1 -----MSTARTENPVLGMSSQNGQLRGLPKPSAGPGGGGTQOQIN-QLKNASTINSGSSQQQAQSMSSIKKPGDWWKKTLKLEPDKLRKTSKSDVTSK
 rck/p54 1 -----MGLSSQNGQLRGPVKPTGGPGGGGTQOQQMNLKNTNTINNGSSQQQAQSMSTTIKPGDWWKKTLKLEPDKLRKTSKSDVTSK
 p47 1 -----MASVKVDHAINHSSPATNDLKGDAHDQG--WKTKLTPPKDTRVTKSDVNTNK

DHH1p 45 GNMFEDFYLRKRELLMGIFEGFEKPSPIQEEAIPVAITGRDILARAKNGTGKTAAFVITPLEKMKPRLNKQALIMVPTRELALQTSQVVRTLGKHCGL
 STE13p 43 GNMFEDFYLRKRELLMGIFEGFEKPSPIQEEAIPVAITGRDILARAKNGTGKTAAFVITPLEKMKPRLNKQALIMVPTRELALQTSQVVRTLGKHCGL
 DOZI 49 GNMFEDFYLRKRELLMGIFEGFEKPSPIQEEAIPVAITGRDILARAKNGTGKTAAFVITPLEKMKPRLNKQALIMVPTRELALQTSQVVRTLGKHCGL
 PFC0915w 59 GNMFEDFYLRKRELLMGIFEGFEKPSPIQEEAIPVAITGRDILARAKNGTGKTAAFVITPLEKMKPRLNKQALIMVPTRELALQTSQVVRTLGKHCGL
 Cryptosporidium 1 -----MGIPEKGFEPSPPIQEEAIPVAITGRDILARAKNGTGKTAAFVITPLEKMKPRLNKQALIMVPTRELALQTSQVVRTLGKHCGL
 Toxoplasma 101 GNMFEDFYLRKRELLMGIFEGFEKPSPIQEEAIPVAITGRDILARAKNGTGKTAAFVITPLEKMKPRLNKQALIMVPTRELALQTSQVVRTLGKHCGL
 Babesia 59 GNMFEDFYLRKRELLMGIFEGFEKPSPIQEEAIPVAITGRDILARAKNGTGKTAAFVITPLEKMKPRLNKQALIMVPTRELALQTSQVVRTLGKHCGL
 Theileria 42 GNMFEDFYLRKRELLMGIFEGFEKPSPIQEEAIPVAITGRDILARAKNGTGKTAAFVITPLEKMKPRLNKQALIMVPTRELALQTSQVVRTLGKHCGL
 Dictyostelium 48 GNMFEDFYLRKRELLMGIFEGFEKPSPIQEEAIPVAITGRDILARAKNGTGKTAAFVITPLEKMKPRLNKQALIMVPTRELALQTSQVVRTLGKHCGL
 CGH-1 42 GNMFEDFYLRKRELLMGIFEGFEKPSPIQEEAIPVAITGRDILARAKNGTGKTAAFVITPLEKMKPRLNKQALIMVPTRELALQTSQVVRTLGKHCGL
 ME31B 57 GNMFEDFYLRKRELLMGIFEGFEKPSPIQEEAIPVAITGRDILARAKNGTGKTAAFVITPLEKMKPRLNKQALIMVPTRELALQTSQVVRTLGKHCGL
 XP54 94 GNMFEDFYLRKRELLMGIFEGFEKPSPIQEEAIPVAITGRDILARAKNGTGKTAAFVITPLEKMKPRLNKQALIMVPTRELALQTSQVVRTLGKHCGL
 rck/p54 84 GNMFEDFYLRKRELLMGIFEGFEKPSPIQEEAIPVAITGRDILARAKNGTGKTAAFVITPLEKMKPRLNKQALIMVPTRELALQTSQVVRTLGKHCGL
 p47 52 GNMFEDFYLRKRELLMGIFEGFEKPSPIQEEAIPVAITGRDILARAKNGTGKTAAFVITPLEKMKPRLNKQALIMVPTRELALQTSQVVRTLGKHCGL

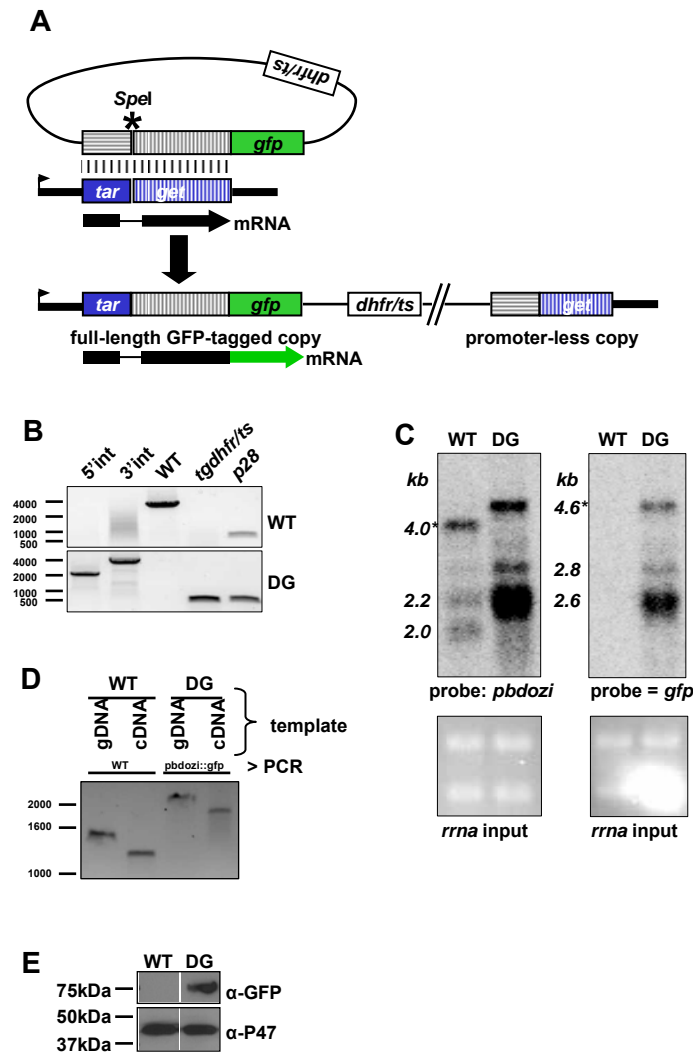
DHH1p 144 SGMVTTGGTSLRDDILRLNNTVHILVGTGPRVLDLASEKVVADLSDSGLFMDEADKLLSRDFEKTIEOILSFLPPTHOSLLESATFPVTVKSEFMVKHLHK
 STE13p 142 KVMVTTGGTSLRDDILRLNNTVHILVGTGPRVLDLASEKVVADLSDSGLFMDEADKLLSRDFEKTIEOILSFLPPTHOSLLESATFPVTVKSEFMVKHLHK
 DOZI 158 QCMVTTGGTSLRDDIMRLNRYNPHVHILCGTPGRILDLANKDVANLSCGCHIMVDEADKLLSPEFQPIVEBLLKFLPKERQILLYSATFPVTVKSEFMVKHLHK
 PFC0915w 158 QCMVTTGGTSLRDDIMRLNRYNPHVHILCGTPGRILDLANKDVANLSCGCHIMVDEADKLLSPEFQPIVEBLLKFLPKERQILLYSATFPVTVKSEFMVKHLHK
 Cryptosporidium 86 QCMVTTGGTSLRDDILRLNNTVHILVGTGPRVLDLASEKVVADLSDSGLFMDEADKLLSPEFQPIVEBLLKFLPKERQILLYSATFPVTVKSEFMVKHLHK
 Toxoplasma 200 QCMVTTGGTSLRDDIMRLNRYNPHVHILCGTPGRILDLANKDVANLSCGCHIMVDEADKLLSPEFQPIVEBLLKFLPKERQILLYSATFPVTVKSEFMVKHLHK
 Babesia 158 QCMVTTGGTSLRDDIMRLNRYNPHVHILCGTPGRILDLANKDVANLSCGCHIMVDEADKLLSPEFQPIVEBLLKFLPKERQILLYSATFPVTVKSEFMVKHLHK
 Theileria 141 QCMVTTGGTSLRDDIMRLNRYNPHVHILCGTPGRILDLANKDVANLSCGCHIMVDEADKLLSPEFQPIVEBLLKFLPKERQILLYSATFPVTVKSEFMVKHLHK
 Dictyostelium 147 QCMVTTGGTSLRDDIMRLNRYNPHVHILVGTGPRVLDLASEKVVADLSDSGLFMDEADKLLSPEFQPIVEBLLKFLPKERQILLYSATFPVTVKSEFMVKHLHK
 CGH-1 141 KVMVTTGGTSLRDDIMRLNRYNPHVHILVGTGPRVLDLASEKVVADLSDSGLFMDEADKLLSPEFQPIVEBLLKFLPKERQILLYSATFPVTVKSEFMVKHLHK
 ME31B 156 KVMVTTGGTSLRDDILRLNRYNPHVHILCGTPGRILDLANKDVANLSCGCHIMVDEADKLLSPEFQPIVEBLLKFLPKERQILLYSATFPVTVKSEFMVKHLHK
 XP54 194 KVMVTTGGTSLRDDIMRLNRYNPHVHILVGTGPRVLDLASEKVVADLSDSGLFMDEADKLLSPEFQPIVEBLLKFLPKERQILLYSATFPVTVKSEFMVKHLHK
 rck/p54 184 KVMVTTGGTSLRDDIMRLNRYNPHVHILVGTGPRVLDLASEKVVADLSDSGLFMDEADKLLSPEFQPIVEBLLKFLPKERQILLYSATFPVTVKSEFMVKHLHK
 p47 151 KVMVTTGGTSLRDDIMRLNRYNPHVHILVGTGPRVLDLASEKVVADLSDSGLFMDEADKLLSPEFQPIVEBLLKFLPKERQILLYSATFPVTVKSEFMVKHLHK

DHH1p 244 PVEINLMDELTLLKGVTOYYAFVBEROKHCLNTLFSKLOINOAIIFCNSITRVELLAKKITELGSCFYIHAKMVOAHRNRVPHDFRNGACRLVSSDLF
 STE13p 242 PVEINLMDELTLLKGVTOYYAFVBEROKHCLNTLFSKLOINOAIIFCNSITRVELLAKKITELGSCFYIHAKMVOAHRNRVPHDFRNGACRLVSSDLF
 DOZI 258 ABEINLMDELTLLKGVTOYYAFVBEROKHCLNTLFSKLOINOAIIFCNSITRVELLAKKITELGSCFYIHAKMVOAHRNRVPHDFRNGACRLVSSDLF
 PFC0915w 258 ABEINLMDELTLLKGVTOYYAFVBEROKHCLNTLFSKLOINOAIIFCNSITRVELLAKKITELGSCFYIHAKMVOAHRNRVPHDFRNGACRLVSSDLF
 Cryptosporidium 186 ABEINLMDELTLLKGVTOYYAFVBEROKHCLNTLFSKLOINOAIIFCNSITRVELLAKKITELGSCFYIHAKMVOAHRNRVPHDFRNGACRLVSSDLF
 Toxoplasma 300 ABEINLMDELTLLKGVTOYYAFVBEROKHCLNTLFSKLOINOAIIFCNSITRVELLAKKITELGSCFYIHAKMVOAHRNRVPHDFRNGACRLVSSDLF
 Babesia 258 ABEINLMDELTLLKGVTOYYAFVBEROKHCLNTLFSKLOINOAIIFCNSITRVELLAKKITELGSCFYIHAKMVOAHRNRVPHDFRNGACRLVSSDLF
 Theileria 241 ABEINLMDELTLLKGVTOYYAFVBEROKHCLNTLFSKLOINOAIIFCNSITRVELLAKKITELGSCFYIHAKMVOAHRNRVPHDFRNGACRLVSSDLF
 Dictyostelium 247 ABEINLMDELTLLKGVTOYYAFVBEROKHCLNTLFSKLOINOAIIFCNSITRVELLAKKITELGSCFYIHAKMVOAHRNRVPHDFRNGACRLVSSDLF
 CGH-1 241 PVEINLMDELTLLKGVTOYYAFVBEROKHCLNTLFSKLOINOAIIFCNSITRVELLAKKITELGSCFYIHAKMVOAHRNRVPHDFRNGACRLVSSDLF
 ME31B 256 PVEINLMDELTLLKGVTOYYAFVBEROKHCLNTLFSKLOINOAIIFCNSITRVELLAKKITELGSCFYIHAKMVOAHRNRVPHDFRNGACRLVSSDLF
 XP54 294 PVEINLMDELTLLKGVTOYYAFVBEROKHCLNTLFSKLOINOAIIFCNSITRVELLAKKITELGSCFYIHAKMVOAHRNRVPHDFRNGACRLVSSDLF
 rck/p54 284 PVEINLMDELTLLKGVTOYYAFVBEROKHCLNTLFSKLOINOAIIFCNSITRVELLAKKITELGSCFYIHAKMVOAHRNRVPHDFRNGACRLVSSDLF
 p47 251 PVEINLMDELTLLKGVTOYYAFVBEROKHCLNTLFSKLOINOAIIFCNSITRVELLAKKITELGSCFYIHAKMVOAHRNRVPHDFRNGACRLVSSDLF

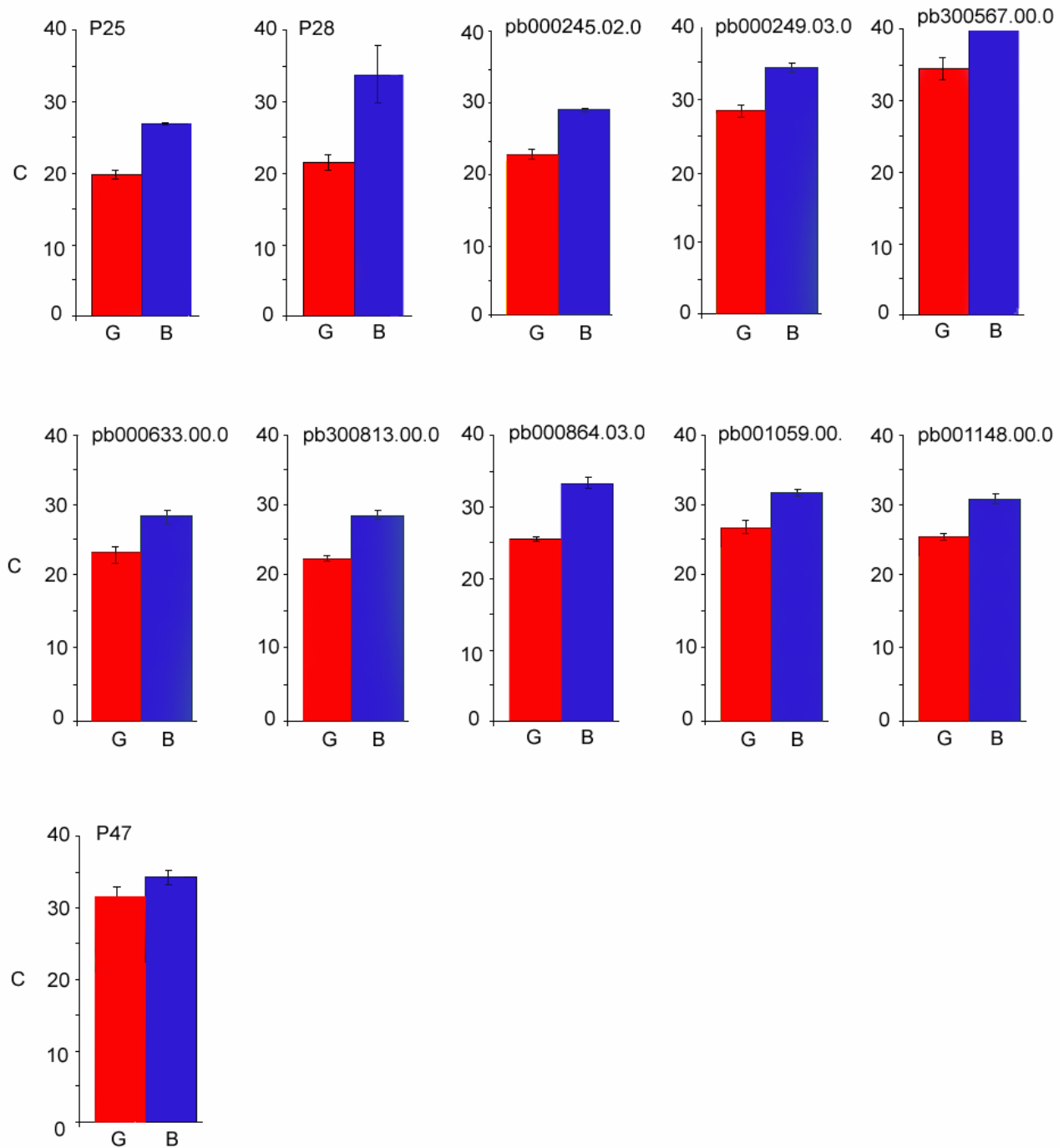
DHH1p 344 TRGIDIQAVNVVINFDFPKNAETYLHRIGRSGRFGHLGLAINLITYEDRFNLYRIESELGTEIQPIPAHVDPALYS
 STE13p 342 TRGIDIQAVNVVINFDFPKNAETYLHRIGRSGRFGHLGLAINLITYEDRFNLYRIESELGTEIQPIPAHVDPALYS
 DOZI 358 TRGIDIQAVNVVINFDFPKNAETYLHRIGRSGRFGHLGLAINLITYEDRFNLYRIESELGTEIQPIPAHVDPALYS
 PFC0915w 358 TRGIDIQAVNVVINFDFPKNAETYLHRIGRSGRFGHLGLAINLITYEDRFNLYRIESELGTEIQPIPAHVDPALYS
 Cryptosporidium 286 TRGIDIQAVNVVINFDFPKNAETYLHRIGRSGRFGHLGLAINLITYEDRFNLYRIESELGTEIQPIPAHVDPALYS
 Toxoplasma 400 TRGIDIQAVNVVINFDFPKNAETYLHRIGRSGRFGHLGLAINLITYEDRFNLYRIESELGTEIQPIPAHVDPALYS
 Babesia 358 -----
 Theileria 341 TRGIDIQAVNVVINFDFPKNAETYLHRIGRSGRFGHLGLAINLITYEDRFNLYRIESELGTEIQPIPAHVDPALYS
 Dictyostelium 347 TRGIDIQAVNVVINFDFPKNAETYLHRIGRSGRFGHLGLAINLITYEDRFNLYRIESELGTEIQPIPAHVDPALYS
 CGH-1 341 TRGIDIQAVNVVINFDFPKNAETYLHRIGRSGRFGHLGLAINLITYEDRFNLYRIESELGTEIQPIPAHVDPALYS
 ME31B 356 TRGIDIQAVNVVINFDFPKNAETYLHRIGRSGRFGHLGLAINLITYEDRFNLYRIESELGTEIQPIPAHVDPALYS
 XP54 394 TRGIDIQAVNVVINFDFPKNAETYLHRIGRSGRFGHLGLAINLITYEDRFNLYRIESELGTEIQPIPAHVDPALYS
 rck/p54 384 TRGIDIQAVNVVINFDFPKNAETYLHRIGRSGRFGHLGLAINLITYEDRFNLYRIESELGTEIQPIPAHVDPALYS
 p47 351 TRGIDIQAVNVVINFDFPKNAETYLHRIGRSGRFGHLGLAINLITYEDRFNLYRIESELGTEIQPIPAHVDPALYS

DHH1p 444 QQQLPSQQQFAIPPQQHHPQFVPPSHQQQQAYPPPQMPSSQQGYPPQEHFMAMPFGQSQPOY-----
 STE13p 421 -----GDYQIPRPLTASADQVLAQAQKQGEYHNRPNNNRGGHPRGGGNGRGGYQSNRQPRYRGGQKAD
 DOZI
 PFC0915w
 Cryptosporidium
 Toxoplasma
 Babesia
 Theileria
 Dictyostelium
 CGH-1
 ME31B 456 NVSK-----
 XP54
 rck/p54
 p47

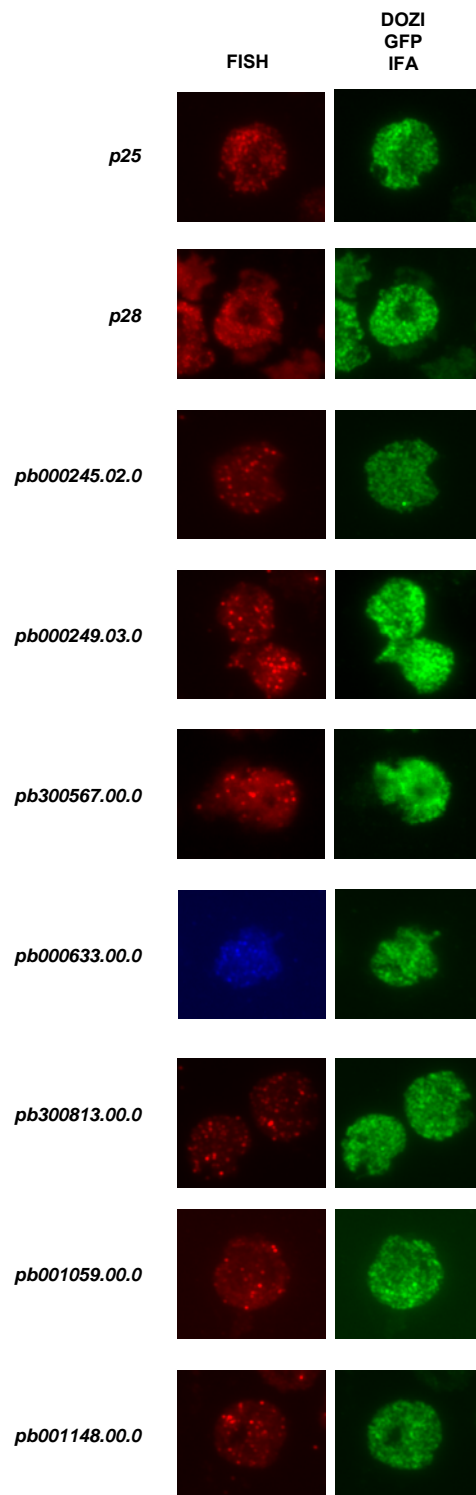
Supplementary Figure S3 ClustaW alignment of RNA helicases shown in Supplementary Figure 2.



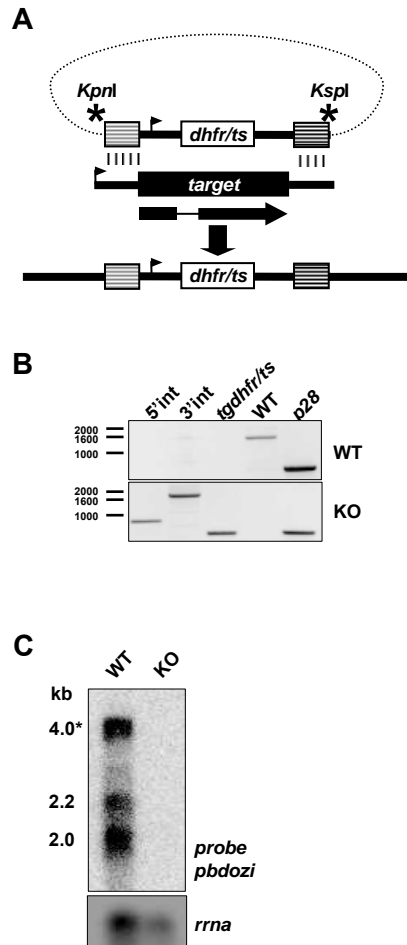
Supplementary Figure S4 Generation and characterization of parasite, PbDOZI::GFP, expressing GFP-tagged DOZI. A Schematic representation of GFP-tagging of the *pbdozi*-locus using a construct that integrated through single cross-over homologous recombination and contained the *tgdhfr/ts* selectable marker. **B** Diagnostic PCRs of mutant population. PCRs showing 5' and 3' integration of the plasmid construct into the *pbdozi* locus (lanes 5' and 3' int), lack of WT *pbdozi*, presence of *tgdhfr/ts* and the *p28* locus as a control. **C** Northern analysis of transcripts of PbDOZI::GFP gametocytes showing the presence of *pbdozi::gfp* transcripts, absence of wild type *pbdozi* mRNA. Asterisk indicates the *pbdozi* transcript specific to gametocytes. **D** RT-PCR confirming correct use of splice sites in the *pbdozi::gfp* pre-mRNA. **E** Western blot analyses of gametocyte proteins of PbDOZI::GFP showing the correct size (~75kDa) for the DOZI::GFP protein. P47 was used as a control.



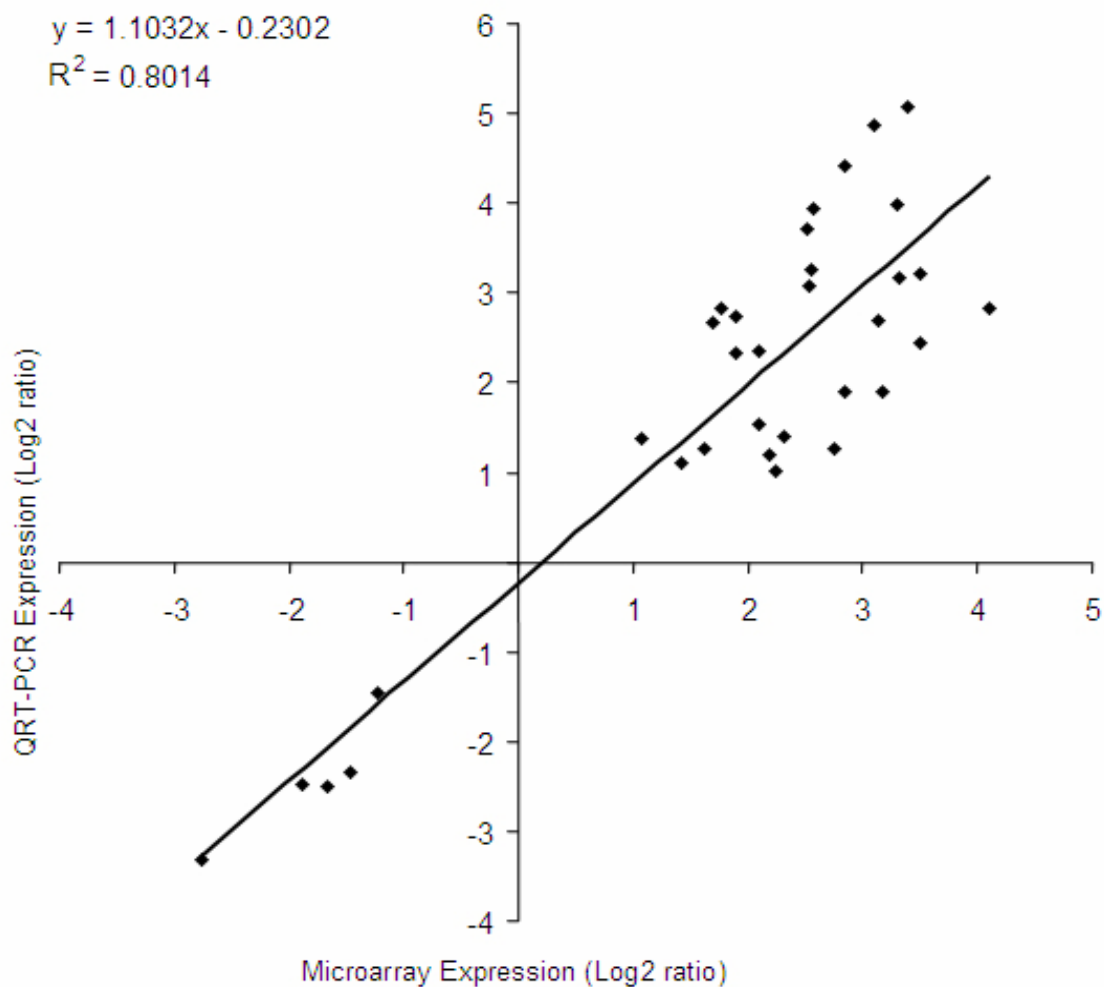
Supplementary Figure S5 RT-qPCR of immunoprecipitation eluates. The cycle number (**C**) at which a threshold of 0.2 relative fluorescent units (RFU) was reached is indicated for each gene since a normalization of cDNA templates with a control gene was not feasible for these types of assays. Enrichment of translationally repressed genes in the IP-GFP fraction (**G**) versus the uncoated bead fraction (**B**) was reflected by the larger difference in cycle numbers needed to reach the 2.0 RFU for these transcripts compared to the P47 control transcript. The maximum cycle number used in these assays was 40. Amplification of gene *pb300567.00.0* exceeded this 40 cycle limit in the uncoated bead eluate. Data is representative of 3 replicas. Error bars reflect standard deviation between replicas for each gene and p-values ($p < 0.05$) were determined by single-factor ANOVA analysis.



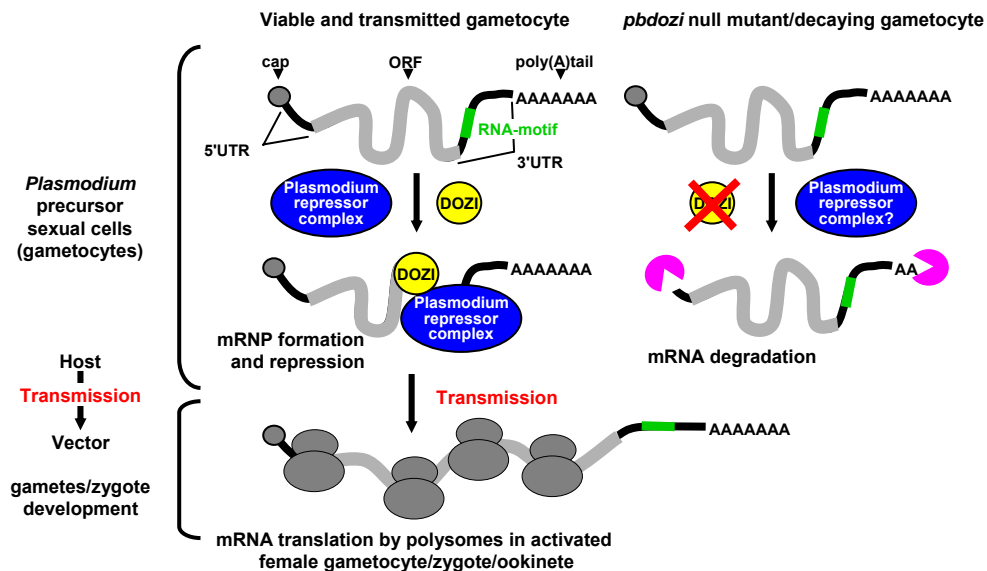
Supplementary Figure S6 FISH analyses of transcripts translationally repressed in *Plasmodium* blood stage gametocytes.



Supplementary Figure S7 Generation and characterization of *pbdozi*⁻ null mutant parasites. **A.** *Pbdozi*⁻ parasites were generated by double crossover homologous recombination replacing *pbdozi* with the *tgdhfr/ts* selectable marker. **B.** Diagnostic PCRs of mutant population. PCRs show 5' and 3' integration of the targeting plasmid into the *pbdozi* locus (lanes 5' and 3' int), lack of wild type *pbdozi*, presence of *tgdhfr/ts* and *p28* as a control. **C.** Northern blot analysis of RNA from *pbdozi*⁻ null parasites showing the lack of *pbdozi* transcripts (*rrna* was included as input control).



Supplementary Figure S8 Correlation of quantitative real-time RT-PCR and microarray assayed gene expression. Thirteen genes, shown to be differentially regulated by microarray analysis, were verified using quantitative real-time PCR. Log₂ transformed gene expression ratios derived from microarray (X-axis) were plotted against log₂ transformed ratios derived from QRT-PCR (Y-axis) for the same gene. The best-fit linear regression analysis ($R^2 = .801$) and the slope of the regression line ($m = 1.10$) indicate a high degree of consistency between microarray and real time PCR analyses. Detailed gene expression data for these assays are presented in Supplementary Table S1.



Supplementary Figure S9 Proposed role for DOZI in translational repression in sexual cells (gametocytes) of *Plasmodium*. In the presence of the RNA helicase DOZI specific transcripts are assembled into translationally quiescent mRNPs in the female gametocytes that are the precursors of the female gametes. These mRNAs are stored for later translation after gamete formation and fertilisation, which occurs in the mosquito. The stable assembly of such translational repressor complexes for the transcripts *p25* and *p28*, encoding surface proteins of the zygote, is absolutely dependent on the activity of DOZI; in the absence of DOZI these transcripts are neither stored in silent mRNPs nor transported to translating polysomes but instead are specifically degraded.

Supporting Online Material

Supplementary Table S1 | Microarray gene expression data presented as normalized wild type/*pbd0zi* null mutant signal intensity ratios.

Only genes exhibiting 2-fold or greater differential expression between wild type and mutant parasites are presented. Pf *Plasmodium*

falciparum, PC *Plasmodium chabaudi*, PY *Plasmodium yoelii*, orth ortholog, SP signal peptide, tm transmembrane

Gene ID	Homology	wild type <i>pbd0zi</i>	Pf orth	PC orth	PY orth	Annotation	SP	# tm domains
PB000296.02.0	no homology	10,45	MAL13P1.312	PC000614.00.0	EAA19665.1	conserved hypothetical protein	No	0
PB301316.00.0	heat shock protein 86 family protein, putative	10,25	PF07_0031	PC000578.02.0	—	heat shock protein 86 family protein, putative	No	0
PB001102.01.0	no homology	8,27	—	PC001324.02.0	—	hypothetical protein	No	0
PB000432.00.0	no homology	8,15	PFE0220w	PC000865.02.0	EAA15615.1	conserved hypothetical protein	Yes	0
PB000157.01.0	no homology	7,56	MAL13P1.102	PC301521.00.0	EAA21145.1	conserved hypothetical protein	No	0
PB000504.02.0	LCCL domain-containing protein CCP2, putative	7,53	PF14_0532	PC000363.04.0	EAA20927.1	LCCL domain-containing protein CCP2, putative	Yes	0
PB300774.00.0	no homology	7,51	PFI1455c	PC000635.00.0	—	conserved hypothetical protein	No	0
PB001257.02.0	Pfemp3-like protein, putative	7,48	PF13_0314	PC000020.03.0	EAA20473.1	Pfemp3-like protein, putative	No	0
PB000653.03.0	no homology	7,38	PF13_0220	PC000127.00.0	EAA16438.1	conserved hypothetical protein	No	0
PB401889.00.0	no homology	7,35	—	PC105904.00.0	—	hypothetical protein	No	1
PB000673.02.0	calcium-dependent protein kinase, putative	7,26	PF14_0227	PC000274.03.0	EAA19816.1	calcium-dependent protein kinase, putative	No	0
PB300912.00.0	no homology	7,18	—	PC000811.03.0	—	hypothetical protein	No	0
PB000768.00.0	no homology	7,11	—	PC000192.02.0	—	hypothetical protein	No	0
PB000272.02.0	ankyrin 3, putative	6,98	PF14_0222	PC000123.02.0	EAA21215.1	ankyrin 3, putative	No	0
PB000603.01.0	ATP-dependent RNA helicase, putative	6,89	PFC0915w	PC000373.03.0	EAA21264.1	ATP-dependent RNA helicase, putative	No	0
PB000214.02.0	no homology	6,89	PFI0705w	PC000014.04.0	EAA21833.1	conserved hypothetical protein	No	0
PB001243.00.0	Pb-fam-5 protein	6,83	PF13_0168	PC001214.02.0	—	Pb-fam-5 protein	No	0
PB000446.02.0	no homology	6,82	MAL6P1.177	PC000064.00.0	EAA21767.1	conserved hypothetical protein	No	6
PB000388.00.0	no homology	6,80	MAL12P1.351	—	EAA18278.1	conserved hypothetical protein	No	0
PB001059.00.0	Rad51 homolog, putative	6,64	PF11_0087	PC000445.01.0	EAA15553.1	Rad51 homolog, putative	No	0
PB000594.00.0	no homology	6,35	PF14_0544	—	—	conserved hypothetical protein	No	0
PB000266.01.0	25 kDa ookinete surface antigen precursor	6,26	PF10_0303	PC000894.02.0	EAA16663.1	25 kDa ookinete surface antigen precursor	No	1
PB300567.00.0	no homology	6,21	MAL12P1.482	PC001038.02.0	—	conserved hypothetical protein	No	0
PB000477.02.0	protein kinase, putative	6,20	PF11_0464	PC000083.05.0	EAA18028.1	protein kinase, putative	No	0
PB000623.02.0	no homology	6,16	MAL12P1.290	PC000606.02.0	EAA19418.1	conserved hypothetical protein	No	0
PB001125.00.0	no homology	6,08	PF08_0033	PC001387.02.0	EAA16375.1	conserved hypothetical protein	No	0
PB000298.01.0	no homology	6,05	—	PC000727.04.0	—	hypothetical protein	No	0
PB001203.02.0	no homology	5,95	MAL7P1.74	—	EAA22392.1	conserved hypothetical protein	No	0
PB000851.02.0	no homology	5,94	PFA0490w	PC000051.02.0	EAA15734.1	conserved hypothetical protein	No	0
PB300170.00.0	no homology	5,85	PFE0770w	—	—	conserved hypothetical protein	No	0
PB000648.02.0	no homology	5,84	MAL12P1.292	PC000484.04.0	EAA15538.1	conserved hypothetical protein	No	0
PB001099.01.0	no homology	5,83	MAL12P1.102	PC301624.00.0	EAA15199.1	conserved hypothetical protein	No	0
PB001250.02.0	blood stage antigen 41-3 precursor, putative	5,83	MAL12P1.77	PC000287.01.0	EAA15984.1	blood stage antigen 41-3 precursor, putative	No	0
PB300353.00.0	high mobility group protein, putative	5,81	MAL8P1.72	PC000449.02.0	—	high mobility group protein, putative	No	0
PB000981.00.0	no homology	5,78	PF11_0160	PC000782.00.0	EAA18466.1	conserved hypothetical protein	No	0
PB000846.02.0	no homology	5,77	PF13_0064	PC301610.00.0	EAA22300.1	conserved hypothetical protein	No	0
PB000572.01.0	no homology	5,76	PF11_0442	PC000148.03.0	EAA15892.1	conserved hypothetical protein	No	0
PB300652.00.0	no homology	5,67	MAL13P1.316	PC300180.00.0	—	conserved hypothetical protein	No	0
PB000408.03.0	no homology	5,57	PFB0176c	PC000507.01.0	EAA19063.1	conserved hypothetical protein	No	0

Gene ID	Homology	wild_type pabozi	Pf orth	PC orth	PY orth	Annotation	SP	# tm domains
PB301072.00.0	no homology	5,55	MAL6P1.308	-	-	conserved hypothetical protein	No	0
PB104209.00.0	no homology	5,45	-	PC107824.00.0	EAA21123.1	conserved hypothetical protein	Yes	0
PB401599.00.0	no homology	5,45	-	PC402872.00.0	-	hypothetical protein	No	0
PB001237.00.0	no homology	5,44	PF0545w	PC000122.01.0	EAA16803.1	conserved hypothetical protein	No	0
PB000817.01.0	kinesin-like protein, putative	5,41	PF07_0104	PC001107.02.0	EAA19086.1	kinesin-like protein, putative	No	0
PB001069.00.0	myosin a, putative	5,39	PF13_0233	PC001166.02.0	EAA20523.1	myosin a, putative	No	0
PB108208.00.0	no homology	5,31	-	-	EAA22128.1	hypothetical protein	Yes	0
PB403516.00.0	no homology	5,29	-	PC000042.05.0	-	hypothetical protein	No	0
PB000264.01.0	no homology	5,26	-	PC104505.00.0	EAA18530.1	hypothetical protein	No	0
PB000051.02.0	no homology	5,21	PF13_0103	PC000454.03.0	EAA22833.1	conserved hypothetical protein	No	0
PB000866.03.0	no homology	5,16	PF08_0073	PC000984.01.0	EAA17348.1	conserved hypothetical protein	No	0
PB000370.02.0	oxoglutarate/malate translocator protein,	5,16	PF08_0031	PC000403.04.0	EAA21506.1	oxoglutarate/malate translocator protein,	No	2
PB000682.02.0	no homology	5,16	PF13_0050	PC000152.04.0	EAA21762.1	conserved hypothetical protein	No	0
PB000020.03.0	von willebrand factor a-domain-related protein	5,12	PF08_0136b	PC000790.01.0	EAA20187.1	von willebrand factor a-domain-related protein,	Yes	0
PB300813.00.0	no homology	5,10	PFC0185w	PC000124.01.0	-	conserved hypothetical protein	No	0
PB000955.03.0	no homology	5,06	PFI0185w	PC000205.00.0	-	conserved hypothetical protein	No	0
PB000915.01.0	no homology	5,06	PFE0550w	PC000037.02.0	-	conserved hypothetical protein	No	0
PB000553.03.0	no homology	5,05	PF13_0161	PC000775.02.0	EAA22151.1	conserved hypothetical protein	No	0
PB000242.01.0	no homology	5,04	-	PC000301.05.0	EAA22381.1	hypothetical protein	No	0
PB000407.03.0	no homology	5,01	PFB0177c	PC000508.01.0	EAA19064.1	conserved hypothetical protein	No	0
PB001166.02.0	Pfs77 protein, putative	5,01	MAL6P1.213	PC000665.02.0	EAA16164.1	Pfs77 protein, putative	No	0
PB000685.03.0	no homology	4,95	PF14_0683	PC000149.02.0	-	conserved hypothetical protein	No	0
PB000096.01.0	no homology	4,95	PF10_0039	PC000425.04.0	EAA16185.1	conserved hypothetical protein	No	0
PB000662.02.0	no homology	4,93	MAL8P1.29	PC000674.01.0	EAA17720.1	conserved hypothetical protein	No	0
PB000261.02.0	no homology	4,92	-	PC300547.00.0	-	hypothetical protein	No	0
PB000750.01.0	no homology	4,92	PF10_0069	PC001377.02.0	EAA20079.1	conserved hypothetical protein	No	0
PB300940.00.0	no homology	4,92	MAL12P1.233	PC000756.00.0	-	conserved hypothetical protein	No	0
PB300798.00.0	no homology	4,91	PF0395c	PC000191.00.0	-	conserved hypothetical protein	No	0
PB405922.00.0	no homology	4,89	MAL12P1.73	PC103204.00.0	-	conserved hypothetical protein	Yes	1
PB000061.02.0	Ser/Thr protein kinase, putative	4,89	PFB0665w	PC301925.00.0	EAA18603.1	Ser/Thr protein kinase, putative	No	0
PB001153.02.0	ADP-ribosylation factor-like protein, putative	4,81	PF10_0337	PC000658.03.0	EAA17498.1	ADP-ribosylation factor-like protein, putative	No	0
PB000226.02.0	no homology	4,80	MAL12P1.372	PC000142.05.0	EAA19215.1	conserved hypothetical protein	No	0
PB000034.01.0	no homology	4,79	PF11_0325	PC102268.00.0	-	conserved hypothetical protein	No	0
PB107920.00.0	no homology	4,77	-	PC101833.00.0	-	hypothetical protein	No	0
PB001624.02.0	no homology	4,76	PF11305w	PC000738.01.0	EAA22080.1	conserved hypothetical protein	No	6
PB000461.01.0	membrane skeletal protein, putative	4,72	PF11_0431	PC301320.00.0	EAA15257.1	membrane skeletal protein, putative	No	0
PB300665.00.0	acid phosphatase, putative	4,71	PFI0880c	PC000941.02.0	-	acid phosphatase, putative	Yes	2
PB102048.00.0	no homology	4,68	-	PC107916.00.0	EAA19374.1	hypothetical protein	No	0
PB000865.01.0	no homology	4,68	MAL6P1.30	-	EAA21534.1	conserved hypothetical protein	No	0
PB001394.02.0	RNA pseudouridylylate synthase, putative	4,65	MAL12P1.270	-	EAA15805.1	RNA pseudouridylylate synthase, putative	No	0
PB000346.00.0	no homology	4,61	PFC0176c	PC000326.04.0	EAA16468.1	conserved hypothetical protein	No	0
PB000634.00.0	no homology	4,61	PF14_0521	PC000968.01.0	EAA18975.1	conserved hypothetical protein	No	0
PB000907.00.0	glycolipid transfer protein, putative	4,56	PFI0775w	PC000093.02.0	EAA19956.1	glycolipid transfer protein, putative	No	0
PB000836.00.0	no homology	4,53	PFI0995w	-	-	conserved hypothetical protein	No	0
PB000423.00.0	no homology	4,49	PF11_0348	PC000484.00.0	EAA17321.1	conserved hypothetical protein	No	0
PB300967.00.0	Pb-reticulocyte binding protein	4,47	MAL6P1.309	-	-	Pb-reticulocyte binding protein	No	0
PB000191.01.0	no homology	4,46	PF10_0037	PC300709.00.0	EAA19440.1	conserved hypothetical protein	No	0
PB000721.01.0	no homology	4,44	PF14_0470	PC000166.03.0	EAA15508.1	conserved hypothetical protein	No	0
PB000693.03.0	no homology	4,40	PFE0785c	PC000749.01.0	EAA21807.1	conserved hypothetical protein	No	10
PB000638.02.0	choline kinase, putative	4,38	PF14_0020	PC000127.01.0	EAA20668.1	choline kinase, putative	No	0

Gene ID	Homology	wild type pabozi	Pf orth	PC orth	PY orth	Annotation	SP	# tm domains
PB000859.01.0	no homology	4,37	PF10_0046	PC000705.02.0	EAA18913.1	conserved hypothetical protein	No	0
PB000220.00.0	no homology	4,34	PFE1260c	-	-	conserved hypothetical protein	No	0
PB000506.00.0	no homology	4,33	MAL13P1.350	PC000351.00.0	EAA18658.1	conserved hypothetical protein	No	0
PB001084.00.0	no homology	4,31	PF14_0491	PC000679.01.0	EAA20973.1	conserved hypothetical protein	No	0
PB001233.02.0	aspartyl protease, putative	4,30	PF10_0329	PC000289.01.0	EAA16175.1	aspartyl protease, putative	No	0
PB000245.02.0	no homology	4,30	MAL12P1.463	PC000594.03.0	-	conserved hypothetical protein	No	0
PB300559.00.0	no homology	4,29	MAL13P1.332	PC000524.00.0	-	conserved hypothetical protein	No	0
PB000507.03.0	Pb-fam-4 protein	4,27	PF08_0118	PC000588.00.0	EAA17595.1	Pb-fam-4 protein	No	0
PB000829.00.0	ADP-ribosylation factor-like protein, putative	4,25	PF11005w	PC000187.02.0	-	ADP-ribosylation factor-like protein, putative	No	0
PB000229.00.0	no homology	4,25	PF11_0261	PC000364.00.0	EAA15230.1	conserved hypothetical protein	No	0
PB000209.00.0	ribose 5-phosphate epimerase, putative	4,24	PFE0730c	PC000527.02.0	EAA19557.1	ribose 5-phosphate epimerase, putative	No	0
PB001168.00.0	erythrocyte membrane-associated antigen,	4,24	PF10_0177	PC000212.04.0	EAA22037.1	erythrocyte membrane-associated antigen,	No	0
PB000621.03.0	Formin 2, putative	4,24	MAL12P1.186	PC000698.03.0	-	Formin 2, putative	No	0
PB000566.02.0	protein kinase, putative	4,22	MAL13P1.84	PC000418.03.0	EAA18782.1	protein kinase, putative	No	0
PB000737.02.0	no homology	4,20	PFD0520c	-	EAA21745.1	conserved hypothetical protein	No	0
PB000824.02.0	no homology	4,20	PF13_0093	PC000722.00.0	EAA21884.1	conserved hypothetical protein	No	0
PB000249.03.0	no homology	4,18	PF13_0226	PC000096.00.0	EAA22618.1	conserved hypothetical protein	No	0
PB107847.00.0	no homology	4,17	-	PC401007.00.0	-	hypothetical protein	No	0
PB000464.00.0	kinesin-related protein, putative	4,17	MAL12P1.109	PC000617.04.0	EAA21870.1	kinesin-related protein, putative	No	0
PB000497.00.0	no homology	4,15	-	-	EAA21856.1	hypothetical protein	No	0
PB100146.00.1	no homology	4,15	PF10560c	-	EAA22256.1	conserved hypothetical protein	No	0
PB000176.03.0	no homology	4,10	PFD0910w	PC000077.02.0	EAA17643.1	conserved hypothetical protein	No	0
PB107172.00.0	no homology	4,10	-	PC000724.04.0	EAA17896.1	hypothetical protein	No	0
PB000028.02.0	no homology	4,09	MAL8P1.137	PC000644.04.0	EAA19577.1	conserved hypothetical protein	No	0
PB000201.02.0	transporter, putative	4,02	PFE0825w	PC001134.02.0	EAA22873.1	transporter, putative	No	10
PB000876.00.0	small heat shock protein, putative	4,02	MAL8P1.78	PC000572.02.0	EAA16935.1	small heat shock protein, putative	No	0
PB001642.02.0	no homology	4,00	PF10720w	PC000745.01.0	EAA22130.1	conserved hypothetical protein	No	10
PB000314.02.0	no homology	4,00	PF10460w	PC000742.03.0	EAA22662.1	conserved hypothetical protein	No	0
PB000195.02.0	no homology	3,98	PFE0480c	-	-	conserved hypothetical protein	No	0
PB106639.00.0	phosphatidylinositol synthase, putative	3,96	MAL13P1.82	PC106162.00.0	-	phosphatidylinositol synthase, putative	No	0
PB105848.00.0	no homology	3,95	-	-	EAA22916.1	hypothetical protein	No	0
PB001004.01.0	no homology	3,94	PF10_0290	PC000162.01.0	EAA21500.1	conserved hypothetical protein	No	0
PB000780.01.0	no homology	3,93	PF11230c	PC000461.00.0	EAA19354.1	conserved hypothetical protein	No	3
PB000212.01.0	chromatin assembly factor 1 p55 subunit,	3,90	PF14_0314	PC000775.03.0	EAA17737.1	chromatin assembly factor 1 p55 subunit,	No	0
PB108044.00.0	no homology	3,90	-	PC102010.00.0	EAA19722.1	hypothetical protein	No	0
PB000196.01.0	no homology	3,90	PF11_0481	PC000802.00.0	EAA21855.1	conserved hypothetical protein	No	0
PB001439.02.0	aspartyl (acid) protease, putative	3,88	PF13_0133	PC301713.00.0	EAA21507.1	aspartyl (acid) protease, putative	No	1
PB000362.03.0	no homology	3,85	MAL6P1.303	PC000622.02.0	-	conserved hypothetical protein	No	0
PB000230.01.0	no homology	3,84	MAL12P1.218	PC000002.02.0	EAA23022.1	conserved hypothetical protein	No	0
PB000875.03.0	no homology	3,84	PF14_0333	PC300665.00.0	EAA15476.1	conserved hypothetical protein	No	0
PB000207.03.0	no homology	3,82	PFE1285w	PC000819.02.0	EAA15402.1	conserved hypothetical protein	No	0
PB104395.00.0	no homology	3,81	-	PC000634.02.0	EAA22867.1	hypothetical protein	No	0
PB000717.02.0	no homology	3,81	PF10_0070	PC001375.02.0	EAA21272.1	conserved hypothetical protein	No	0
PB001068.03.0	no homology	3,80	MAL6P1.80	PC001208.02.0	-	conserved hypothetical protein	No	0
PB000814.03.0	no homology	3,80	PF10205w	PC301616.00.0	EAA20438.1	conserved hypothetical protein	Yes	0
PB000861.02.0	small GTPase Rab11, putative	3,79	PF13_0119	PC302555.00.0	EAA19507.1	small GTPase Rab11, putative	No	0
PB101358.00.0	no homology	3,75	-	-	EAA20110.1	hypothetical protein	No	2
PB000997.02.0	no homology	3,71	PFE1440c	PC000602.00.0	-	conserved hypothetical protein	No	0
PB000842.00.0	RNA binding protein, putative	3,69	PFE1310c	PC000091.01.0	EAA15644.1	RNA binding protein, putative	No	0
PB300957.00.0	no homology	3,68	PF14_0628	PC001247.02.0	-	conserved hypothetical protein	No	0

Gene ID	Homology	wild_type pdbozi	Pf orth	PC orth	PY orth	Annotation	SP	# tm domains
PB001222.02.0	aspartate carbamoyltransferase, putative	3,68	MAL13P1.221	PC000879.01.0	EAA18377.1	aspartate carbamoyltransferase, putative	No	0
PB000193.03.0	myosin A tail domain interacting protein MTIP,	3,67	MAL12P1.444	PC000024.03.0	EAA15642.1	myosin A tail domain interacting protein MTIP,	No	0
PB000538.01.0	no homology	3,64	MAL13P1.154	PC000400.04.0	EAA22153.1	conserved hypothetical protein	No	0
PB000365.00.0	no homology	3,62	PF08_0015	—	EAA18360.1	conserved hypothetical protein	No	0
PB000011.02.0	no homology	3,61	PF11095w	PC000867.03.0	—	conserved hypothetical protein	No	0
PB000594.02.0	no homology	3,61	PF14_0330	PC000210.03.0	EAA22020.1	conserved hypothetical protein	No	0
PB300803.00.0	no homology	3,59	PF14_0682	PC000148.02.0	—	conserved hypothetical protein	No	2
PB000530.01.0	no homology	3,55	PF0725c	PC000528.02.0	—	—	—	—
PB300207.00.0	no homology	3,54	PF11_0226	—	—	conserved hypothetical protein	No	0
PB000282.00.0	no homology	3,53	PF13_0123	PC301536.00.0	EAA20981.1	conserved hypothetical protein	No	0
PB000090.02.0	no homology	3,53	MAL6P1.274	—	—	conserved hypothetical protein	No	0
PB103795.00.0	Pb-fam-3 protein	3,52	—	PC106287.00.0	EAA18307.1	Pb-fam-3 protein	No	1
PB001092.00.0	no homology	3,52	MAL7P1.123	PC405116.00.0	EAA21471.1	conserved hypothetical protein	Yes	0
PB000989.01.0	multidrug resistance protein, putative	3,52	PF1150w	PC000730.02.0	EAA22011.1	multidrug resistance protein, putative	No	11
PB102411.00.0	no homology	3,51	—	PC107965.00.0	—	hypothetical protein	Yes	0
PB402680.00.0	no homology	3,50	MAL8P1.66	—	—	conserved hypothetical protein	No	0
PB000596.00.0	no homology	3,49	PF14_0499	—	EAA17165.1	conserved hypothetical protein	No	0
PB000952.03.0	no homology	3,48	—	PC000016.00.0	—	hypothetical protein	No	0
PB001036.00.0	Pb-fam-5 protein	3,48	PF1540w	PC000049.03.0	—	Pb-fam-5 protein	No	0
PB101408.00.0	no homology	3,46	—	PC300129.00.0	EAA19075.1	hypothetical protein	No	0
PB300696.00.0	no homology	3,44	MAL13P1.342	—	—	conserved hypothetical protein	No	0
PB000151.03.0	s-adenosylmethionine synthetase, putative	3,43	PF11090w	PC000567.01.0	EAA18424.1	s-adenosylmethionine synthetase, putative	No	0
PB001530.02.0	Pb-fam-5 protein	3,39	MAL13P1.195	PC000129.04.0	EAA17607.1	Pb-fam-5 protein	No	0
PB000986.00.0	no homology	3,38	PF11_0107	PC000701.03.0	EAA15463.1	conserved hypothetical protein	No	2
PB108086.00.0	no homology	3,38	—	PC400525.00.0	—	hypothetical protein	No	0
PB000000.01.0	no homology	3,37	PF10_0273	PC000736.02.0	—	conserved hypothetical protein	No	4
PB000365.03.0	ABC transporter, putative	3,35	PF14_0244	PC000011.02.0	EAA21447.1	ABC transporter, putative	No	2
PB000296.03.0	no homology	3,34	—	PC301422.00.0	—	hypothetical protein	No	0
PB000671.00.0	no homology	3,34	MAL7P1.87	PC301839.00.0	EAA22856.1	conserved hypothetical protein	No	0
PB300464.00.0	no homology	3,33	—	PC301291.00.0	—	hypothetical protein	No	0
PB000865.00.0	ookinete surface protein Pos28-1, putative	3,32	PF10_0302	—	EAA16664.1	ookinete surface protein Pos28-1, putative	Yes	0
PB105756.00.0	no homology	3,32	—	—	EAA16127.1	conserved hypothetical protein	No	0
PB001104.02.0	Ser/Thr protein kinase, putative	3,31	PF14_0392	PC000765.01.0	EAA17959.1	Ser/Thr protein kinase, putative	No	0
PB001170.01.0	no homology	3,31	—	PC000871.03.0	—	hypothetical protein	No	0
PB000278.01.0	no homology	3,30	PF11_0372	PC000033.04.0	EAA20517.1	conserved hypothetical protein	No	0
PB100346.00.0	no homology	3,27	—	PC402915.00.0	EAA21933.1	hypothetical protein	No	0
PB301410.00.0	no homology	3,25	PF11_0046	PC000049.00.0	—	conserved hypothetical protein	Yes	0
PB101372.00.0	no homology	3,24	—	PC105010.00.0	EAA22797.1	hypothetical protein	No	0
PB101089.00.0	no homology	3,22	—	PC107744.00.0	—	hypothetical protein	No	0
PB300241.00.0	no homology	3,22	PF0910w	PC000219.04.0	—	conserved hypothetical protein	No	0
PB000480.01.0	no homology	3,20	PF11_0393	PC000585.04.0	EAA18635.1	conserved hypothetical protein	No	0
PB001266.00.0	no homology	3,19	—	PC000524.02.0	—	hypothetical protein	No	0
PB000243.02.0	no homology	3,19	PF07_0056	PC000408.03.0	EAA22090.1	conserved hypothetical protein	No	0
PB001206.02.0	no homology	3,18	PF14_0216	PC301134.00.0	EAA21221.1	conserved hypothetical protein	No	0
PB001090.00.0	no homology	3,18	PF07_0094	PC000501.01.0	EAA21473.1	conserved hypothetical protein	No	0
PB000437.01.0	no homology	3,15	PF14_0625	PC000443.00.0	EAA21417.1	conserved hypothetical protein	No	0
PB001343.02.0	no homology	3,14	MAL6P1.225	PC000012.05.0	EAA22802.1	conserved hypothetical protein	No	0
PB000218.02.0	no homology	3,13	PF11_0172	—	EAA21582.1	conserved hypothetical protein	No	11
PB000633.00.0	no homology	3,12	PF14_0522	PC000965.01.0	EAA18976.1	conserved hypothetical protein	No	0
PB000084.01.0	no homology	3,11	PF08_0097	PC000564.00.0	EAA18262.1	conserved hypothetical protein	No	3

Gene ID	Homology	wild_type pabozi	Pf orth	PC orth	PY orth	Annotation	SP	# tm domains
PB405795.00.0	no homology	3,09	MAL12P1.203	PC100644.00.0	-	conserved hypothetical protein	No	0
PB405114.00.0	no homology	3,09	-	PC000636.01.0	-	hypothetical protein	No	1
PB000314.00.0	no homology	3,09	MAL12P1.206	PC000477.00.0	EAA20426.1	conserved hypothetical protein	No	0
PB102130.00.0	no homology	3,08	-	PC301339.00.0	-	hypothetical protein	No	0
PB000719.03.0	no homology	3,07	MAL13P1.130	PC001158.02.0	EAA16662.1	conserved hypothetical protein	No	5
PB101232.00.0	no homology	3,05	-	PC400585.00.0	EAA18509.1	conserved hypothetical protein	Yes	1
PB001577.02.0	serine/threonine protein phosphatase, putative	3,05	PFI1360c	PC000103.04.0	EAA16027.1	serine/threonine protein phosphatase, putative	No	0
PB001053.01.0	cop-coated vesicle membrane protein p24	3,02	PF13_0082	PC107876.00.0	EAA16232.1	cop-coated vesicle membrane protein p24	No	1
PB300561.00.0	no homology	3,01	MAL6P1.44	PC000219.00.0	-	conserved hypothetical protein	No	0
PB000710.03.0	probable glycine C-acetyltransferase, putative	2,99	PF14_0155	PC001399.02.0	EAA15427.1	probable glycine C-acetyltransferase, putative	No	1
PB000784.01.0	no homology	2,96	-	-	EAA15854.1	hypothetical protein	No	0
PB000297.00.0	no homology	2,96	PFI1210w	PC000171.01.0	EAA17387.1	conserved hypothetical protein	No	0
PB001148.00.0	pyridine nucleotide transhydrogenase, putative	2,95	PF14_0508	PC000209.01.0	EAA18012.1	pyridine nucleotide transhydrogenase, putative	No	11
PB000235.02.0	no homology	2,94	MAL13P1.138	PC000643.00.0	EAA21438.1	conserved hypothetical protein	No	0
PB403064.00.0	Pb-fam-1 protein	2,93	-	PC107269.00.0	-	Pb-fam-1 protein	No	0
PB001658.02.0	no homology	2,93	PF14_0143	-	-	conserved hypothetical protein	No	0
PB101863.00.0	no homology	2,93	-	PC000589.00.0	-	hypothetical protein	No	0
PB000147.02.0	no homology	2,92	PF10_0146	PC000270.01.0	EAA22614.1	conserved hypothetical protein	No	0
PB000500.01.0	no homology	2,92	-	PC000860.01.0	EAA20002.1	hypothetical protein	No	0
PB001077.01.0	no homology	2,90	PF11_0091	PC301184.00.0	EAA19384.1	conserved hypothetical protein	No	0
PB108046.00.0	no homology	2,89	-	PC405424.00.0	-	hypothetical protein	No	0
PB300014.00.0	no homology	2,89	-	PC300046.00.0	-	hypothetical protein	No	0
PB000033.03.0	no homology	2,88	-	PC000159.04.0	-	hypothetical protein	No	0
PB001164.00.0	no homology	2,88	PF14_0379	PC000376.03.0	-	conserved hypothetical protein	No	0
PB000313.00.0	no homology	2,88	-	PC000248.03.0	-	hypothetical protein	No	0
PB000164.03.0	no homology	2,86	MAL12P1.130	PC000377.05.0	EAA18035.1	conserved hypothetical protein	No	0
PB104452.00.0	no homology	2,86	-	PC403958.00.0	EAA22240.1	conserved hypothetical protein	No	0
PB000363.03.0	no homology	2,86	MAL13P1.215	PC301815.00.0	EAA17521.1	conserved hypothetical protein	No	0
PB300356.00.0	no homology	2,85	PF0580c	PC001211.02.0	-	conserved hypothetical protein	No	0
PB301467.00.0	Sec61-gamma subunit of protein translocation	2,85	PFB0450w	PC301670.00.0	-	Sec61-gamma subunit of protein translocation	No	1
PB106737.00.0	no homology	2,84	-	PC106110.00.0	EAA20408.1	conserved hypothetical protein	No	0
PB000936.01.0	no homology	2,83	PFI1145w	PC000487.04.0	EAA17167.1	conserved hypothetical protein	No	0
PB102525.00.0	no homology	2,81	-	PC108360.00.0	EAA15956.1	hypothetical protein	No	2
PB106659.00.0	no homology	2,81	-	PC405659.00.0	-	hypothetical protein	No	0
PB000457.03.0	no homology	2,81	PF10_0156	PC000113.01.0	EAA18893.1	conserved hypothetical protein	No	0
PB001235.00.0	no homology	2,80	MAL6P1.33	PC000506.04.0	EAA21531.1	conserved hypothetical protein	No	0
PB001449.02.0	no homology	2,77	-	PC000383.01.0	EAA19462.1	hypothetical protein	No	0
PB105259.00.0	no homology	2,77	-	PC102645.00.0	EAA22191.1	conserved hypothetical protein	No	0
PB000118.01.0	no homology	2,76	PFI1680w	PC300799.00.0	EAA17235.1	conserved hypothetical protein	No	0
PB000801.00.0	no homology	2,75	PF14_0487	PC301449.00.0	-	conserved hypothetical protein	No	0
PB001212.02.0	no homology	2,73	MAL6P1.101	PC000444.03.0	EAA21360.1	conserved hypothetical protein	No	4
PB105705.00.0	no homology	2,73	-	PC101262.00.0	EAA22192.1	hypothetical protein	No	0
PB000151.00.0	no homology	2,71	MAL13P1.236	PC000715.01.0	-	conserved hypothetical protein	No	0
PB100542.00.0	no homology	2,70	-	PC000195.00.0	-	hypothetical protein	No	0
PB001334.02.0	no homology	2,69	PF14_0040	PC000172.05.0	-	conserved hypothetical protein	No	0
PB000115.00.0	no homology	2,69	-	PC001012.02.0	-	hypothetical protein	No	0
PB000232.02.0	no homology	2,69	PFI0595c	PC109033.00.0	EAA22991.1	conserved hypothetical protein	No	5
PB000273.01.0	adenylate kinase 1, putative	2,69	PFD0755c	PC000249.02.0	EAA22307.1	adenylate kinase 1, putative	No	0
PB000407.02.0	no homology	2,68	PFI1465w	PC000412.02.0	-	conserved hypothetical protein	No	0
PB000251.03.0	no homology	2,68	PFB0690w	PC000131.01.0	EAA21342.1	conserved hypothetical protein	No	0

Gene ID	Homology	wild type <i>pabozi</i>	Pf orth	PC orth	PY orth	Annotation	SP	# tm domains
PB000924.03.0	QF122 antigen, putative	2,68	PF10_0115	PC000217.00.0	-	QF122 antigen, putative	No	0
PB105834.00.0	no homology	2,67	-	PC000590.04.0	-	hypothetical protein	No	0
PB000513.02.0	Ser/Thr protein kinase, putative	2,65	MAL13P1.278	PC000375.01.0	EAA21778.1	Ser/Thr protein kinase, putative	No	0
PB000416.01.0	no homology	2,65	PF08_0017	PC000418.04.0	EAA18358.1	conserved hypothetical protein	No	0
PB000396.00.0	calcium/calmodulin-dependent protein kinase 2,	2,64	MAL12P1.375	PC001103.02.0	EAA18495.1	calcium/calmodulin-dependent protein kinase 2,	No	0
PB000272.00.0	no homology	2,62	-	PC000296.00.0	EAA22310.1	hypothetical protein	No	0
PB000503.02.0	no homology	2,61	PFI0575c	PC000754.03.0	EAA22994.1	conserved hypothetical protein	No	0
PB000220.03.0	GDP-mannose 4,6-dehydratase, putative	2,60	PF08_0077	PC000770.03.0	EAA18685.1	GDP-mannose 4,6-dehydratase, putative	No	0
PB000331.02.0	no homology	2,60	PF11_0217	PC000689.00.0	EAA15311.1	conserved hypothetical protein	No	0
PB001293.00.0	Pb-fam-5 protein	2,58	MAL12P1.334	PC000267.04.0	EAA15233.1	Pb-fam-5 protein	No	0
PB000676.02.0	no homology	2,58	MAL6P1.128	PC301341.00.0	-	Pb-fam-5 protein	No	0
PB108683.00.0	no homology	2,56	-	PC106029.00.0	-	hypothetical protein	No	0
PB300360.00.0	Ribosomal protein S17 homologue, putative	2,56	MAL13P1.327	PC000668.02.0	-	Ribosomal protein S17 homologue, putative	No	0
PB000023.01.0	no homology	2,54	-	PC001015.02.0	-	hypothetical protein	No	0
PB000663.01.0	no homology	2,53	-	PC000798.04.0	-	hypothetical protein	No	0
PB100015.00.0	no homology	2,53	-	PC107103.00.0	-	hypothetical protein	No	0
PB000904.03.0	no homology	2,52	MAL12P1.235	PC000213.00.0	EAA17948.1	conserved hypothetical protein	No	0
PB401402.00.0	no homology	2,52	-	PC000220.03.0	-	hypothetical protein	Yes	0
PB000229.02.0	pre-mRNA splicing factor, putative	2,51	PF10_0217	PC000212.01.0	EAA21016.1	pre-mRNA splicing factor, putative	No	0
PB000411.00.0	no homology	2,51	-	PC300130.00.0	EAA16146.1	hypothetical protein, conserved	No	0
PB000591.01.0	no homology	2,51	PF10_0278	PC000234.03.0	EAA17544.1	conserved hypothetical protein	No	0
PB102403.00.0	no homology	2,50	-	PC107182.00.0	-	hypothetical protein	No	0
PB405932.00.0	no homology	2,49	-	PC404473.00.0	-	hypothetical protein	No	0
PB000974.00.0	no homology	2,48	PF13_0026	-	EAA21621.1	conserved hypothetical protein	No	0
PB000202.02.0	unknown protein, mb2, putative	2,48	PFE0830c	PC000384.03.0	-	unknown protein, mb2, putative	No	0
PB000070.00.0	Pb-fam-5 protein	2,48	PFC0895w	-	EAA17152.1	Pb-fam-5 protein	No	0
PB000621.01.0	no homology	2,46	PF11_0045	PC000340.02.0	-	conserved hypothetical protein	No	0
PB001169.02.0	dihydroorotate dehydrogenase, mitochondrial	2,45	MAL6P1.36	PC000269.03.0	EAA22047.1	dihydroorotate dehydrogenase, mitochondrial	No	0
PB300991.00.0	no homology	2,44	PF14_0252	-	-	conserved hypothetical protein	No	0
PB100016.00.0	no homology	2,44	-	PC404874.00.0	EAA22190.1	hypothetical protein	No	0
PB104794.00.0	no homology	2,44	-	-	EAA22890.1	hypothetical protein	No	0
PB001621.02.0	no homology	2,43	PF10_0047	PC000074.04.0	EAA17572.1	conserved hypothetical protein	No	0
PB000902.02.0	no homology	2,42	PF14_0471	PC000242.00.0	-	conserved hypothetical protein	No	0
PB000786.01.0	50S ribosomal subunit protein L17, putative	2,42	PFE1125w	PC000287.02.0	EAA22352.1	50S ribosomal subunit protein L17, putative	No	0
PB300220.00.0	no homology	2,41	-	PC301488.00.0	-	hypothetical protein	No	1
PB000892.00.0	pbs36 homologue, putative	2,40	PFD0210c	PC000354.01.0	EAA20644.1	pbs36 homologue, putative	No	0
PB106176.00.0	no homology	2,40	-	-	EAA21242.1	hypothetical protein	No	0
PB101260.00.0	no homology	2,39	-	PC109157.00.0	-	hypothetical protein	No	0
PB000591.02.0	cytochrome c oxidase assembly protein, putative	2,39	PF14_0331	PC000793.01.0	EAA22019.1	cytochrome c oxidase assembly protein, putative	No	0
PB000435.02.0	no homology	2,38	PF14_0065	PC000755.04.0	EAA18649.1	conserved hypothetical protein	No	6
PB103867.00.0	no homology	2,37	-	PC106050.00.0	-	hypothetical protein	No	0
PB001116.03.0	long-chain-fatty-acid--CoA ligase, putative	2,36	MAL12P1.374	PC000662.04.0	-	long-chain-fatty-acid--CoA ligase, putative	No	0
PB001160.00.0	no homology	2,35	PFI1155w	PC000700.00.0	EAA21836.1	conserved hypothetical protein	No	0
PB404039.00.0	no homology	2,35	-	PC108585.00.0	-	hypothetical protein	No	0
PB103706.00.0	no homology	2,34	PF11_0182	PC103008.00.0	-	hypothetical protein	No	0
PB000326.02.0	no homology	2,33	PF14_0452	PC000200.00.0	-	conserved hypothetical protein	No	0
PB001095.03.0	DNA helicase, putative	2,33	PFB0730w	PC000453.04.0	-	DNA helicase, putative	No	0
PB001520.02.0	glycerol kinase, putative	2,33	PF13_0269	PC001056.02.0	EAA19393.1	glycerol kinase, putative	No	0
PB000579.00.0	porphobilinogen deaminase, putative	2,33	MAL12P1.96	PC000109.00.0	EAA21199.1	porphobilinogen deaminase, putative	No	0
PB001505.02.0	no homology	2,32	PFB0625w	PC302425.00.0	EAA15722.1	conserved hypothetical protein	No	0

Gene ID	Homology	wild_type pdbozi	Pf orth	PC orth	PY orth	Annotation	SP	# tm domains
PB001101.02.0	no homology	2,32	PF14_0395	PC301918.00.0	EAA20592.1	conserved hypothetical protein	No	0
PB001197.02.0	no homology	2,31	MAL13P1.129	PC000259.02.0	EAA20932.1	conserved hypothetical protein	No	0
PB103734.00.0	no homology	2,31	_	PC403173.00.0	EAA17366.1	hypothetical protein	No	0
PB000548.00.0	no homology	2,29	PF07_0076	_	EAA17953.1	conserved hypothetical protein	No	0
PB000545.01.0	no homology	2,29	PFE0650c	PC302649.00.0	EAA16516.1	conserved hypothetical protein	No	0
PB301540.00.0	no homology	2,29	MAL6P1.20	PC001234.02.0	_	conserved hypothetical protein	No	0
PB000826.03.0	no homology	2,28	PFB0395w	PC106046.00.0	EAA16684.1	conserved hypothetical protein	No	0
PB000642.00.0	no homology	2,26	PFE0890c	PC000444.04.0	EAA20968.1	conserved hypothetical protein	No	0
PB000701.02.0	serine protease belonging to subtilisin family	2,24	PFE0355c	PC000265.03.0	_	serine protease belonging to subtilisin family,	No	0
PB001245.00.0	no homology	2,24	PF13_0169	PC001215.02.0	EAA21607.1	conserved hypothetical protein	No	0
PB300486.00.0	co-chaperone GrpE, putative	2,24	PF11_0258	PC000284.03.0	_	co-chaperone GrpE, putative	No	0
PB000463.00.0	no homology	2,22	PF14_0266	PC001271.02.0	EAA16337.1	conserved hypothetical protein	No	0
PB401854.00.0	no homology	2,22	_	PC100938.00.0	_	hypothetical protein	No	0
PB000104.00.0	no homology	2,21	MAL6P1.126	PC001148.02.0	_	conserved hypothetical protein	No	1
PB000239.00.0	no homology	2,21	MAL7P1.31	PC000924.01.0	EAA16586.1	conserved hypothetical protein	No	0
PB000877.01.0	ATP-dependent DNA helicase, putative	2,20	PF13_0330	PC000286.04.0	_	ATP-dependent DNA helicase, putative	No	0
PB000864.03.0	aspartyl protease, putative	2,20	PFC0495w	PC000527.00.0	EAA21080.1	aspartyl protease, putative	No	0
PB300800.00.0	no homology	2,20	PF07_0122	PC000794.04.0	_	conserved hypothetical protein	No	0
PB000990.00.0	no homology	2,20	PF07_0124	PC000155.01.0	EAA20807.1	conserved hypothetical protein	No	0
PB106971.00.0	no homology	2,20	_	PC401233.00.0	_	hypothetical protein	No	0
PB104900.00.0	no homology	2,19	_	_	EAA17797.1	hypothetical protein	No	0
PB107195.00.0	no homology	2,19	_	PC105035.00.0	_	hypothetical protein	No	0
PB001070.00.0	phosphoenolpyruvate carboxykinase, putative	2,19	PF13_0234	PC001162.02.0	EAA20524.1	phosphoenolpyruvate carboxykinase, putative	No	0
PB000018.03.0	Pb-fam-5 protein	2,19	PF14_0272	PC000688.04.0	_	Pb-fam-5 protein	No	0
PB200000.00.0	BIR protein, putative	2,19	_	_	EAA19381.1	BIR protein, putative	No	0
PB000041.03.0	mitochondrial carrier protein, putative	2,19	MAL12P1.398	PC000732.01.0	EAA20831.1	mitochondrial carrier protein, putative	No	0
PB000109.00.0	no homology	2,18	PF11_0210	PC000694.01.0	EAA19157.1	conserved hypothetical protein	No	2
PB001528.02.0	no homology	2,17	PF13_0246	PC000225.02.0	EAA20630.1	conserved hypothetical protein	No	0
PB001551.02.0	no homology	2,16	MAL6P1.202	PC001310.02.0	EAA21647.1	conserved hypothetical protein	No	1
PB000655.02.0	no homology	2,16	MAL8P1.138	_	_	conserved hypothetical protein	No	0
PB000033.01.0	transporter protein, putative	2,15	MAL6P1.94	PC000437.02.0	EAA20622.1	transporter protein, putative	Yes	8
PB000774.00.0	no homology	2,14	PF14_0525	PC001199.02.0	EAA18979.1	conserved hypothetical protein	No	0
PB000001.01.0	methyltransferase, putative	2,13	PF13_0286	PC000023.01.0	EAA21785.1	methyltransferase, putative	No	0
PB103620.00.0	no homology	2,13	_	PC104232.00.0	_	hypothetical protein	No	0
PB000873.01.0	cyclic nucleotide phosphodiesterase, putative	2,13	PF14_0672	PC000253.02.0	EAA15368.1	cyclic nucleotide phosphodiesterase, putative	No	5
PB001277.00.0	RNA helicase , putative	2,12	PF11_0077	PC000086.00.0	_	RNA helicase , putative	No	0
PB000068.01.0	no homology	2,12	MAL6P1.157	PC000487.03.0	EAA16533.1	conserved hypothetical protein	No	0
PB000225.02.0	no homology	2,12	MAL12P1.371	PC000588.02.0	EAA19214.1	conserved hypothetical protein	No	0
PB001104.03.0	RNA binding protein, putative	2,12	PF11_0471	PC000286.03.0	_	RNA binding protein, putative	No	0
PB001277.02.0	no homology	2,11	MAL13P1.88	PC000381.02.0	EAA22977.1	conserved hypothetical protein	No	0
PB001281.02.0	RNA-binding protein mei2 homologue, putative	2,11	MAL6P1.195	PC000385.01.0	_	RNA-binding protein mei2 homologue, putative	No	0
PB106201.00.0	no homology	2,11	_	PC105230.00.0	_	hypothetical protein	Yes	0
PB000789.02.0	signal peptidase, putative	2,11	PFI0215c	PC000121.03.0	EAA16738.1	signal peptidase, putative	No	0
PB001074.02.0	no homology	2,10	PF13_0289	PC000661.00.0	EAA16829.1	conserved hypothetical protein	No	0
PB000717.03.0	no homology	2,10	_	PC301102.00.0	_	hypothetical protein	No	0
PB000929.00.0	no homology	2,10	PFE0510c	PC001193.02.0	_	conserved hypothetical protein	No	0
PB102255.00.0	no homology	2,10	_	PC101656.00.0	EAA22739.1	hypothetical protein	No	0
PB000685.00.0	no homology	2,09	MAL7P1.105	_	_	conserved hypothetical protein	No	0
PB001553.02.0	no homology	2,09	MAL13P1.50	PC000016.03.0	_	conserved hypothetical protein	No	0
PB000029.03.0	no homology	2,08	PF14_0355	PC107330.00.0	EAA19966.1	conserved hypothetical protein	No	0

Gene ID	Homology	wild type pabozi	Pf orth	PC orth	PY orth	Annotation	SP	# tm domains
PB104901.00.0	no homology	2,08	-	PC103209.00.0	-	hypothetical protein	No	0
PB000475.03.0	no homology	2,08	PF10_0060	PC001364.02.0	EAA18845.1	conserved hypothetical protein	No	0
PB000163.03.0	no homology	2,07	MAL13P1.139	PC103989.00.0	EAA21439.1	conserved hypothetical protein	No	0
PB000747.03.0	no homology	2,07	PF10_0076	PC000935.01.0	EAA19572.1	conserved hypothetical protein	No	0
PB001322.02.0	no homology	2,07	PF14_0578	PC302580.00.0	EAA21118.1	conserved hypothetical protein	No	0
PB000050.01.0	no homology	2,07	MAL12P1.198	PC301490.00.0	-	conserved hypothetical protein	No	2
PB104419.00.0	no homology	2,07	PF14_0033	PC103440.00.0	EAA18922.1	conserved hypothetical protein	No	0
PB001275.02.0	adrenodoxin-type ferredoxin, putative	2,06	MAL12P1.141	PC000097.05.0	-	adrenodoxin-type ferredoxin, putative	No	0
PB106915.00.0	no homology	2,06	-	PC102319.00.0	-	hypothetical protein	No	0
PB000083.02.0	no homology	2,05	PF14_0447	PC000109.02.0	-	conserved hypothetical protein	No	1
PB000186.00.0	u6 snRNA-associated sm-like protein, putative	2,05	PF13_0142	PC000526.03.0	EAA19315.1	u6 snRNA-associated sm-like protein, putative	No	0
PB000268.02.0	no homology	2,04	PFC0060c	-	EAA22891.1	Serine/threonine protein kinase, putative	No	0
PB000558.01.0	no homology	2,04	PFB0490c	PC302508.00.0	-	conserved hypothetical protein	No	0
PB000866.00.0	calmodulin, putative	2,04	PF10_0301	PC000895.02.0	EAA16665.1	calmodulin, putative	No	0
PB001639.02.0	no homology	2,04	PFE1145w	-	-	conserved hypothetical protein	No	0
PB300331.00.0	no homology	2,04	PF11_0459	PC300337.00.0	-	conserved hypothetical protein	No	2
PB000813.01.0	no homology	2,04	PF11_0420	PC000291.03.0	EAA17566.1	conserved hypothetical protein	No	1
PB001100.01.0	Pb-fam-1 protein	2,03	-	PC000503.02.0	EAA22754.1	Pb-fam-1 protein	No	0
PB000764.02.0	no homology	2,03	-	PC106759.00.0	EAA20595.1	hypothetical protein	No	0
PB107328.00.0	no homology	2,02	-	PC102559.00.0	-	hypothetical protein	No	0
PB000775.00.0	no homology	2,02	PF14_0524	PC001200.02.0	-	conserved hypothetical protein	No	0
PB104303.00.0	no homology	2,01	MAL12P1.487	PC000334.03.0	-	conserved hypothetical protein	Yes	0
PB000315.00.0	no homology	2,01	MAL12P1.205	PC108819.00.0	EAA20427.1	conserved hypothetical protein	No	0
PB000709.00.0	calmodulin, putative	2,01	PF14_0181	PC000857.00.0	EAA17188.1	calmodulin, putative	No	0
PB001041.03.0	no homology	2,01	PFI0195c	PC000908.02.0	-	conserved hypothetical protein	No	0
PB001151.02.0	no homology	2,00	PF08_0043	PC000887.01.0	EAA15442.1	conserved hypothetical protein	No	0
PB000975.01.0	no homology	2,00	-	PC000952.01.0	-	hypothetical protein	No	0
PB001353.02.0	methionine aminopeptidase, putative	2,00	PFE1360c	PC000923.02.0	EAA21010.1	methionine aminopeptidase, putative	No	0
PB108385.00.0	no homology	1,99	-	-	EAA21077.1	hypothetical protein	No	0
PB000006.03.0	no homology	1,99	-	PC108581.00.0	-	hypothetical protein	No	0
PB000573.02.0	no homology	1,98	MAL13P1.262	PC000267.00.0	EAA19545.1	conserved hypothetical protein	No	0
PB000794.03.0	no homology	0,51	-	PC301938.00.0	EAA20979.1	hypothetical protein	No	0
PB300744.00.0	no homology	0,51	PFC0730w	PC000063.01.0	-	conserved protein, putative	No	2
PB000254.00.0	no homology	0,50	PFB0765w	PC000598.00.0	-	conserved hypothetical protein	No	0
PB000485.03.2	metacaspase-like protein, putative	0,50	PF14_0363	PC000188.05.0	EAA17714.1	metacaspase-like protein, putative	No	0
PB000878.02.0	no homology	0,50	PF10_0063	PC000206.00.0	-	conserved hypothetical protein	No	0
PB001284.02.0	Nucleosome assembly protein, putative	0,50	PFI0930c	PC000173.00.0	EAA20954.1	Nucleosome assembly protein, putative	No	0
PB001110.01.0	60S ribosomal protein L12, putative	0,50	PFE0850c	PC000160.01.0	EAA16192.1	60S ribosomal protein L12, putative	No	0
PB000845.02.0	malonyl coa-acyl carrier, putative	0,50	PF13_0066	PC001144.02.0	EAA17522.1	malonyl coa-acyl carrier, putative	No	0
PB000600.01.0	no homology	0,50	-	PC000413.00.0	-	hypothetical protein	No	0
PB301288.00.0	pyruvate dehydrogenase E1 component, alpha	0,50	PF11_0256	PC000512.00.0	-	pyruvate dehydrogenase E1 component, alpha	No	1
PB001106.00.0	pyruvate kinase, putative	0,50	PF10_0363	PC000332.00.0	EAA15660.1	pyruvate kinase, putative	No	0
PB000835.00.0	no homology	0,50	PFI0990c	PC001171.02.0	EAA17249.1	conserved hypothetical protein	No	0
PB000071.00.0	no homology	0,50	PFC0905c	PC000544.00.0	EAA17150.1	conserved hypothetical protein	No	0
PB000891.00.0	pf52 protein, putative	0,49	PFD0215c	PC000135.01.0	EAA20643.1	pf52 protein, putative	Yes	1
PB000315.01.0	no homology	0,49	MAL12P1.56	PC000002.01.0	EAA21130.1	conserved hypothetical protein	No	0
PB000124.03.0	no homology	0,49	PF07_0020	PC301475.00.0	-	conserved hypothetical protein	No	0
PB000108.00.0	beta-ketoacyl-acyl carrier prot. synthase III	0,49	PFB0505c	PC001097.02.0	-	beta-ketoacyl-acyl carrier prot. synthase III	No	1
PB000879.01.0	hsp60, putative	0,49	PF10_0153	PC000876.03.0	EAA16505.1	hsp60, putative	No	0

Gene ID	Homology	wild type <i>pdbozi</i>	Pf orth	PC orth	PY orth	Annotation	SP	# tm domains
PB000164.02.0	tubulin beta chain, putative	0,49	PF10_0084	PC300626.00.0	EAA17778.1	tubulin beta chain, putative	No	0
PB000932.00.0	cold-shock protein, putative	0,49	PFA0470c	PC106733.00.0	-	cold-shock protein, putative	No	0
PB000642.01.0	no homology	0,49	MAL13P1.237	PC000017.00.0	EAA19945.1	conserved hypothetical protein	No	0
PB001177.00.0	Heat shock protein, putative	0,49	PF10875w	PC300345.00.0	EAA16958.1	Heat shock protein, putative	Yes	0
PB000730.01.0	no homology	0,49	MAL12P1.208	PC300426.00.0	EAA20424.1	conserved hypothetical protein	No	0
PB001286.00.0	polyadenylate-binding protein, putative	0,49	MAL12P1.234	PC000304.00.0	EAA17420.1	polyadenylate-binding protein, putative	No	0
PB000039.03.0	DEAD/DEAH box helicase, putative	0,49	MAL12P1.400	PC001092.02.0	EAA20829.1	DEAD/DEAH box helicase, putative	No	0
PB001186.02.0	Ser/Thr protein kinase, putative	0,48	PFB0605w	PC000854.03.0	EAA22506.1	Ser/Thr protein kinase, putative	No	0
PB001564.02.0	GDP-fucose synthase, putative	0,48	PF10_0137	-	EAA21323.1	GDP-fucose synthase, putative	No	0
PB000456.03.0	enolase, putative	0,48	PF10_0155	PC000044.01.0	EAA18892.1	enolase, putative	No	0
PB000142.00.0	no homology	0,48	-	PC001170.02.0	-	hypothetical protein	No	0
PB000251.02.0	ran binding protein 1, putative	0,48	PFD0950w	PC301359.00.0	EAA19671.1	ran binding protein 1, putative	No	0
PB000237.02.0	no homology	0,48	PFC1045c	PC000071.00.0	EAA15267.1	conserved hypothetical protein	No	0
PB001128.00.0	phosphatidylethanolamine-binding protein,	0,48	MAL12P1.191	PC000489.00.0	EAA21244.1	phosphatidylethanolamine-binding protein,	No	0
PB000533.00.0	ubiquitin/ribosomal fusion protein uba52	0,48	PF13_0346	PC101345.00.0	-	ubiquitin/ribosomal fusion protein uba52	No	0
PB001231.02.0	no homology	0,48	PF10_0333	PC001323.02.0	EAA17494.1	conserved hypothetical protein	No	0
PB000714.01.0	glutamate dehydrogenase, putative	0,47	PF14_0286	PC000581.03.0	EAA20557.1	glutamate dehydrogenase, putative	No	0
PB300693.00.0	no homology	0,47	-	PC000039.01.0	-	hypothetical protein	No	0
PB000465.03.0	eukaryotic initiation factor 5a, putative	0,47	MAL12P1.42	PC000848.00.0	EAA19701.1	eukaryotic initiation factor 5a, putative	No	0
PB000781.01.0	PF16 protein, putative	0,47	PF11_0318	PC000547.01.0	EAA20453.1	PF16 protein, putative	No	0
PB000371.03.0	ubiquitin-conjugating enzyme E2, putative	0,47	MAL6P1.65	PC000265.05.0	EAA22697.1	ubiquitin-conjugating enzyme E2, putative	No	0
PB001096.02.0	histone H2A variant, putative	0,47	PFC0920w	PC000384.04.0	EAA15833.1	histone H2A variant, putative	No	0
PB000178.02.0	40S ribosomal subunit protein S14, putative	0,46	PFE0810c	PC000743.01.0	EAA20993.1	40S ribosomal subunit protein S14, putative	No	0
PB000019.02.0	no homology	0,46	PF08_0086	PC000612.03.0	EAA21161.1	conserved hypothetical protein	No	0
PB000882.01.0	DEAD box helicase, putative	0,46	PF13_0077	PC301153.00.0	EAA15534.1	DEAD box helicase, putative	No	0
PB000674.02.0	DNA ligase 1, putative	0,46	MAL13P1.22	PC000404.01.0	EAA20874.1	DNA ligase 1, putative	Yes	0
PB000427.02.0	no homology	0,46	PF11_0360	-	EAA16206.1	conserved hypothetical protein	No	0
PB000132.02.0	pyruvate dehydrogenase E1 beta subunit,	0,46	PF14_0441	-	EAA19412.1	pyruvate dehydrogenase E1 beta subunit,	No	0
PB000050.00.0	no homology	0,46	PF13_0154	PC300945.00.0	EAA19689.1	conserved hypothetical protein	No	2
PB001333.02.0	guanylyl cyclase, putative	0,46	PF14_0043	PC000277.03.0	EAA19773.1	guanylyl cyclase, putative	No	0
PB000446.01.0	ADP-ribosylation factor, putative	0,46	PF10_0203	PC300214.00.0	EAA16453.1	ADP-ribosylation factor, putative	No	0
PB000252.01.0	no homology	0,45	PFD0430c	PC000807.02.0	EAA16049.1	conserved hypothetical protein	No	0
PB000777.02.0	no homology	0,44	PF14_0376	PC001098.02.0	EAA18792.1	conserved hypothetical protein	No	1
PB001003.01.0	endoplasmic reticulum-resident calcium binding	0,44	PF11_0098	PC000827.03.0	EAA17014.1	endoplasmic reticulum-resident calcium binding	No	0
PB000757.02.0	fructose-bisphosphate aldolase, putative	0,44	PF14_0425	PC001392.02.0	EAA15467.1	fructose-bisphosphate aldolase, putative	No	0
PB000523.01.0	60S Acidic ribosomal protein P2, putative	0,44	PFC0400w	PC000329.02.0	EAA22419.1	60S Acidic ribosomal protein P2, putative	No	0
PB001589.02.0	no homology	0,44	PFD0200c	PC000942.02.0	EAA20648.1	conserved hypothetical protein	No	0
PB000649.01.0	cysteine protease, putative	0,44	PFB0325c	PC301037.00.0	EAA21481.1	cysteine protease, putative	No	0
PB000225.03.0	DNA replication licensing factor mcm7 homology	0,43	PF07_0023	PC000280.04.0	EAA22793.1	DNA replication licensing factor mcm7 homology	No	0
PB000927.03.0	heat shock protein DNAJ homologue Pjf4,	0,43	MAL12P1.113	PC000656.04.0	EAA20515.1	heat shock protein DNAJ homologue Pjf4,	No	0
PB000830.02.0	no homology	0,43	PF10_0027	PC000055.02.0	EAA23036.1	conserved hypothetical protein	No	0
PB300359.00.0	no homology	0,43	PF10_0194	PC301314.00.0	-	conserved hypothetical protein	No	0
PB000489.00.0	no homology	0,43	PF13_0116	PC000040.04.0	EAA15654.1	conserved hypothetical protein	No	0
PB000780.02.0	peptide chain release factor subunit 1,	0,42	PFB0550w	PC000129.05.0	EAA15287.1	peptide chain release factor subunit 1,	No	0
PB000900.00.0	14-3-3 protein homologue, putative	0,42	MAL8P1.69	PC000316.05.0	EAA21233.1	14-3-3 protein homologue, putative	No	0
PB000304.03.0	no homology	0,42	MAL12P1.406	-	-	conserved hypothetical protein	No	0
PB300259.00.0	peptide release factor, putative	0,42	PF11575c	PC000172.02.0	-	peptide release factor, putative	No	0
PB000246.02.0	adenosylhomocysteinase	0,41	PFE1050w	PC000295.02.0	EAA22407.1	adenosylhomocysteinase	No	0
PB001001.03.0	ribosomal protein L21e, putative	0,41	PF14_0240	-	-	ribosomal protein L21e, putative	No	0
PB000111.01.0	no homology	0,41	MAL12P1.63	PC000707.03.0	EAA21236.1	conserved hypothetical protein	No	0

Gene ID	Homology	<u>wild type</u> <u>pdbози</u>	Pf orth	PC orth	PY orth	Annotation	SP	# tm domains
PB001080.03.0	translation elongation factor EF-1, subunit	0,41	PF11_0245	PC000020.00.0	_	translation elongation factor EF-1, subunit	No	0
PB000337.00.0	histidine kinase DhkG, putative	0,40	PF13_0237	PC000346.00.0	EAA17519.1	histidine kinase DhkG, putative	No	0
PB000184.00.0	oxidoreductase, putative	0,40	PF13_0144	PC001369.02.0	_	oxidoreductase, putative	No	0
PB000556.00.0	no homology	0,40	PF11_0342	PC000797.01.0	EAA19908.1	conserved hypothetical protein	No	0
PB000224.00.0	ATP-dependent transporter, putative	0,40	PF14_0133	PC000002.05.0	_	ATP-dependent transporter, putative	No	0
PB001026.00.0	circumsporozoite (CS) protein, putative	0,40	PFC0210c	PC000415.02.0	_	circumsporozoite (CS) protein, putative	No	0
PB000371.02.0	exonuclease, putative	0,39	MAL8P1.35	PC000215.00.0	EAA21504.1	exonuclease, putative	No	0
PB001532.02.0	no homology	0,39	_	_	EAA20722.1	hypothetical protein	No	0
PB405568.00.0	no homology	0,39	_	PC104760.00.0	_	hypothetical protein	No	1
PB001057.02.0	GcpE protein, putative	0,38	PF10_0221	PC000451.04.0	EAA21021.1	GcpE protein, putative	No	0
PB000251.01.0	no homology	0,38	PFD0425w	PC000627.01.0	EAA16050.1	conserved hypothetical protein	No	0
PB000521.01.0	asparagine synthetase, putative	0,38	PFC0395w	PC000916.01.0	EAA22420.1	asparagine synthetase, putative	No	0
PB300429.00.0	no homology	0,38	PF11525w	PC000471.04.0	_	conserved hypothetical protein	No	0
PB000378.03.0	no homology	0,37	PF11_0347	PC000485.00.0	EAA21960.1	conserved hypothetical protein	No	0
PB106189.00.0	no homology	0,37	_	PC400135.00.0	_	hypothetical protein	No	0
PB000084.03.0	glyceraldehyde-3-phosphate dehydrogenase,	0,36	PF14_0598	PC000143.05.0	EAA22840.1	glyceraldehyde-3-phosphate dehydrogenase,	No	0
PB000163.02.0	dihydrolipoamide acetyltransferase, putative	0,34	PF10_0407	PC000373.02.0	EAA16454.1	dihydrolipoamide acetyltransferase, putative	No	0
PB000991.02.0	no homology	0,33	_	PC000102.04.0	EAA19835.1	hypothetical protein	No	0
PB000862.00.0	no homology	0,32	PF08_0074	PC000311.03.0	EAA20360.1	conserved hypothetical protein	No	0
PB000817.02.0	heat shock 70 kDa protein, putative	0,32	PF08_0054	PC000186.01.0	EAA18319.1	heat shock 70 kDa protein, putative	No	0
PB000270.03.0	no homology	0,31	_	_	EAA20721.1	hypothetical protein	No	0
PB000158.02.0	cysteine repeat modular protein 2 homologue,	0,30	MAL7P1.92	PC000204.00.0	EAA17272.1	cysteine repeat modular protein 2 homologue,	No	5
PB200061.00.0	BIR protein, putative	0,30	_	_	EAA21967.1	BIR protein, putative	No	1
PB000500.03.0	uridine phosphorylase, putative	0,27	PFE0660c	PC000630.00.0	EAA16513.1	uridine phosphorylase, putative	No	0

Supporting Online Material

Supplementary Table S2 Transcripts associated with ookinete motility and potentially invasion are less abundant in *pbdozi* mutants

Name	Gene ID	<u>wild type</u> <i>pdbozi</i>
ankyrin 3, putative	<i>pb000272.02.0</i>	6.98
25 kDa ookinete surface antigen precursor	<i>pb000266.01.0</i>	6.26
kinesin-like protein, putative	<i>pb000817.01.0</i>	5.41
myosin a, putative	<i>pb001069.00.0</i>	5.39
von willebrand factor a-domain-related protein	<i>pb000020.03.0</i>	5.12
membrane skeletal protein, putative	<i>pb000461.01.0</i>	4.72
kinesin-related protein, putative	<i>pb000464.00.0</i>	4.17
aspartyl (acid) protease, putative	<i>pb001439.02.0</i>	3.88
myosin A tail domain interacting protein MTIP	<i>pb000193.03.0</i>	3.67
ookinete surface protein Pos28-1, putative	<i>pb000865.00.0</i>	3.32
serine protease belonging to subtilisin family	<i>pb000701.02.0</i>	2.24

Supplementary Table S3 Correlation of microarray gene expression with real time quantitative RT-PCR expression (wild type/*pbdozi* null mutant). The fold difference in mRNA abundance is presented; positive values indicate elevated abundance in the wild type parasite and negative values indicate elevated abundance in the DOZI mutant parasite.

gene ID	microarray	RTQ	RT-PCR
<i>pb000245.02.0</i>	5.70	13.18	▶
<i>pb000245.02.0</i>	4.29	5.06	
<i>pb000245.02.0</i>	3.71	4.99	
<i>pb000249.03.0</i>	5.90	9.51	▶
<i>pb000249.03.0</i>	5.78	8.40	
<i>pb000249.03.0</i>	3.73	6.63	
<i>pb000266.01.0</i>	17.27	7.11	<i>p25</i>
<i>pb000266.01.0</i>	4.99	2.64	
<i>pb000266.01.0</i>	4.56	2.28	
<i>pb000633.00.0</i>	11.31	5.39	▶
<i>pb000633.00.0</i>	4.26	2.91	
<i>pb000633.00.0</i>	2.10	2.60	
<i>pb000864.03.0</i>	3.23	6.36	▶
<i>pb000864.03.0</i>	3.10	2.38	
<i>pb000864.03.0</i>	2.69	2.16	
<i>pb000865.00.0</i>	11.39	9.25	<i>p28</i>
<i>pb000865.00.0</i>	3.39	7.11	
<i>pb001059.00.0</i>	9.92	15.89	▶
<i>pb001059.00.0</i>	9.00	3.73	
<i>pb001059.00.0</i>	6.77	2.38	
<i>pb001148.00.0</i>	8.75	6.41	▶
<i>pb001148.00.0</i>	4.72	2.01	
<i>pb300813.00.0</i>	10.48	33.59	
<i>pb300813.00.0</i>	9.99	8.94	
<i>pb300813.00.0</i>	7.16	3.71	

<i>pb300567.00.0</i>	8.63	29.04
<i>pb300567.00.0</i>	7.16	21.26
<i>pb300567.00.0</i>	5.94	15.35
<i>pb000163.02.0</i>	-6.73	-9.99
<i>pb000163.02.0</i>	-3.16	-5.66
<i>pb000371.02.0</i>	-3.68	-5.54
<i>pb000371.02.0</i>	-2.75	-5.06
<i>pb000371.02.0</i>	-2.35	-2.75

p25 and *p28* are known to be translationally repressed.

◀ These transcripts were predicted to be translationally repressed on the basis of their transcription and translation profiles published in Hall *et al.* (11).

Supplementary Table S4**A Primers used in the generation of targeting vectors, and diagnostic PCRs to validate correct integration of these plasmids**

name oligonucleotide sequence

537 TGCTCTAGAATGAATTTTAAATACAGTTTTATT
 538 TGCTCTAGATTACATTACTATCACGTAAATAAC
 695 AATATTCATAACACACTTTTAAGC
 1214 CGGGATCCCTTTGTATAGTTCATCCATGCC
 1419 AGATCTGGGACAAGAATGTG
 1420 ACTAGTGACAGCCATCTCCATCTG
 1662 GATTCATAAATAGTTGGACTTG
 1753 CCGTATGTTGCATCACCTTCACCC
 2238 AAAGAATTCAATATACAGAAAATATATCCTG
 2239 AAAAAGCTTACAAGTGTTATATATATTACAG
 2240 AAACCCGGGACCACATCATAATGATGAACAC
 2241 AAAGGATCCTAATGCGAACAAAACGTGATCC
 2290 ATAAAACCTCGTTATGCATG
 2339 TAATTGTGTCGCTTCAAATG
 2340 TAATTCTTTTATCATAGCAG
 2381 CATGGATCCAGTATAGATGGATGGATC
 2387 TATTATATCATGCTTGCG

B Primers used to generate templates for northern and FISH analysis

name	oligonucleotide sequence	gene
535	TGCTCTAGAATGAATACTTATTACAGTGTT	<i>p25</i> northern
536	TGCTCTAGATTAAATGATATTTGAAAATATTAG	<i>p25</i> northern
1462	CATGCCATGGATGAATACTTATTACAGTG	<i>p25</i> FISH
1463	CCGGAATTCTTAAATGATATTTGAAAATATTAG	<i>p25</i> FISH
537	TGCTCTAGAATGAATTTTAAATACAGTTTTATT	<i>p28</i> northern an FISH

538	TGCTCTAGATTACACTACTATCACGTAAATAAC	<i>p28</i> northern and FISH
1677	GCCTTTATTAGCATTTGTATGTG	<i>warp</i>
1678	GGATTAAGCTTTTGTGGCTCC	<i>warp</i>
1484	AGCCGATAAAGTTGACTTTAT	<i>eef1a</i>
1486	ATAACCCAATTTATTGCACAC	<i>eef1a</i>
1525	TTGGCGCGCTATAATATTCCTTCAATAAGGAAATAC	<i>p47</i>
1526	TCCCCGCGGGTACCTTTTCCATATGCTCATAGTCG	<i>p47</i>
2339	TAATTGTGTCGCTTCAAATG	<i>pbdozi</i>
2381	CATGGATCCAGTATAGATGGATGGATC	<i>pbdozi</i>
2398	CATATGGCCCTAGAAATTGGCAATC	<i>pb000245.02.0</i>
2399	GGATTCATGGATCTATGAAATGTATCCAC	<i>pb000245.02.0</i>
2375	CATATGCAACCTTTCAATTTCTTAAC	<i>pb000633.00.0</i>
2376	GGATCCATATTTTATTCTTACAATGGTTG	<i>pb000633.00.0</i>
2330	TTATGGAAATACAATACTCG	<i>nima kinase</i>
2372	GTTTGTTTTTCTTTTAAAGG	<i>nima kinase</i>
1797	AAGAATTCGGAATTACTGAATAATGAAAAGC	<i>pbccp1/2</i>
1798	TACTAGTTAATGTATGTATATTATACGC	<i>pbccp1/2</i>
196	CGGGATCCATGAGTAAAGGAGAAGAAC	<i>gfp</i>
197	CGGGATCCATTTATTTGTATAGTTCATCC	<i>gfp</i>

C Oligonucleotides used in northern analysis

name	oligonucleotide sequence	gene
87R	ATATGCTTAAATTCGGCGG	<i>5'end LSU</i>
L647	GGAAACAGTCCATCTATAATTG	<i>A/B LSU 5'</i>

D Primers used in RT-PCR of *pbdozi* and *pbdozi::gfp*

name	oligonucleotide sequence	gene
2339	TAATTGTGTCGCTTCAAATG	<i>pbdozi</i>
2560	ATTTTGCATTGAAGCCGC	<i>pbdozi</i>
2339	TAATTGTGTCGCTTCAAATG	<i>pbdozi::gfp</i>
1214	CGGGATCCCTTTGTATAGTTCATCCATGCC	<i>pbdozi::gfp</i>

E Primers used in RTQ-PCR

name	oligonucleotide sequence	gene
PB 163 F	GGAACCCCTTTCAGCAAATGA	<i>pb000163.02.0</i>
PB 163 R	AATGTAGCAGCATGGGAACC	<i>pb000163.02.0</i>
PB 245 F	AAAGCAATGATGCCACATA	<i>pb000245.02.0</i>
PB 245 R	CCACTGCTTCTCATTTGCAC	<i>pb000245.02.0</i>
PB 249 F	TCCCAAATGGGAAATACATGA	<i>pb000249.03.0</i>
PB 249 R	GCGGGCATTCCAAATTTAC	<i>pb000249.03.0</i>
PB 266 F	AATGGATTTTTGGCTCAAATG	<i>pb000266.01.0 = p25</i>
PB 266 R	TGGCAAGTTTCTGGAACACA	<i>pb000266.01.0</i>
PB 371 F	GGGCAAGCAATTTTATGGAA	<i>pb000371.02.0</i>
PB 371 R	CCCTAAATTCGTCCCTTCAA	<i>pb000371.02.0</i>
PB 567 F	GGCATGGAAAAAGGACAAGA	<i>pb300567.00.0</i>
PB 567 R	TTTTCTAATTCCTCTGTCATATATTTT	<i>pb300567.00.0</i>
PB 633 F	ATTTAGCCCCAATGCAACAA	<i>pb000633.00.0</i>
PB 633 R	TGCAGGTTGGGACTCATTTT	<i>pb000633.00.0</i>
PB 813 F	CAAAGTCCCAGAAAATGCAA	<i>pb000813.01.0</i>
PB 813 R	TGGTTTTTGGGCATTTCTTT	<i>pb000813.01.0</i>
PB 864 F	GTTTGATAACTGGGCCATCG	<i>pb000864.03.0</i>
PB 864 R	TGTCAAGGGACATCAGTCCA	<i>pb000864.03.0</i>
PB 865 F	GAAAACACTTGCAGCAAT	<i>pb000865.00.0 = p28</i>
PB 865 R	CATTTTCCAGCGTTACATTCA	<i>pb000865.00.0</i>
PB1059 F	AAGGAGGGATCGAAACTGGT	<i>pb001059.00.0</i>
PB1059 R	TCTGTTGGGTGCAAACCATA	<i>pb001059.00.0</i>
PB1087 F	TGGAAGATAATTACACAAACACCA	<i>pb301087.00.0</i>
PB1087 R	TGCATATCCTTTTATAAATCCAGTTC	<i>pb301087.00.0</i>
PB1148 F	GGGATGAAAACCAAACATGG	<i>pb001148.00.0</i>
PB1148 R	GGATTCGCGTGAAACAATTT	<i>pb001148.00.0</i>
P47F	AATGCCGCAGGATTAATG	<i>p47</i>
P47R	CATCGTACGGTTTCGGATTT	<i>p47</i>

F Primers used in RT-PCR of immunoprecipitation eluates

name	oligonucleotide sequence	gene
1462	CATGCCATGGATGAATACTTATTACAGTG	<i>p25</i>
1463	CCGGAATTCTTAAATGATATTTGAAAATATTAG	<i>P25</i>
537	TGCTCTAGAATGAATTTTAAATACAGTTTTATT	<i>p28</i>
538	TGCTCTAGATTACACTATCACGTAAATAAC	<i>p28</i>
2398	CATATGGCCCTAGAAATTGGCAATTC	<i>pb000245.02.0</i>
2399	GGATTCATGGATCTATGAAATGTATCCAC	<i>pb000245.02.0</i>
2878	TGAATATAATGAAGAACC	<i>pb000249.03.0</i>
249R	GCGGGCATTCCAAATTTAC	<i>pb000249.03.0</i>

2375	CATATGCAACCTTTCAATTTCTTAAC	<i>pb000633.00.0</i>
2376	GGATCCATATTTTATTCTTACAATGGTTG	<i>pb000633.00.0</i>
864F	GTTTGATAACTGGGCCATCG	<i>pb000864.03.0</i>
864R	TGTCAAGGGACATCAGTCCA	<i>pb000864.03.0</i>
2880	AAGAAGATAACAATTTTCGC	<i>pb001059.00.0</i>
1059R	TCTGTTGGGTGCAAACCATA	<i>pb001059.00.0</i>
2881	TATCCAAGTTAATTTTGC	<i>pb001148.00.0</i>
1148R	GGATTCGCGTGAAACAATTT	<i>pb001148.00.0</i>
567F	GGCATGGAAAAAGGACAAGA	<i>pb300567.00.0</i>
567R	TTTTCTAATTCCTCTGTCATATATTTT	<i>pb300567.00.0</i>
813F	CAAAGTCCCAGAAAATGCAA	<i>pb300813.00.0</i>
813R	TGGTTTTTGGGCATTTCTTT	<i>pb300813.00.0</i>
1525	TTGGCGCGCTATAATATTCCTTCAATAAGGAAATAC	<i>p47</i>
1526	TCCCCGCGGGTACCTTTTCCATATGCTCATAGTCG	<i>p47</i>

Supporting Online Material

References

1. S. M. Khan et al., *Cell* **121**, 675-687 (2005).
2. C. J. Janse et al., *Mol.Biochem.Parasitol.* (2005).
3. M. R. van Dijk et al., *Cell* **104**, 153-164 (2001).
4. G. Mair, E. Ullu, C. Tschudi, *J.Biol.Chem.* **275**, 28994-28999 (2000).
5. C. J. Janse et al., *Parasitology* **91 (Pt 1)**, 19-29 (1985).
6. A. L. Beetsma, T. J. van de Wiel, R. W. Sauerwein, W. M. Eling, *Exp.Parasitol.* **88**, 69-72 (1998).
7. R. A. Vervenne, R. W. Dirks, J. Ramesar, A. P. Waters, C. J. Janse, *Mol.Biochem.Parasitol.* **68**, 259-266 (1994).
8. M. R. van Dijk, personal communication
9. C. J. Janse et al., *Mol.Microbiol.* **50**, 1539-1551 (2003).
10. H. H. Chou, A. P. Hsia, D. L. Mooney, P. S. Schnable, *Bioinformatics.* **20**, 2893-2902 (2004).
11. N. Hall et al., *Science* **307**, 82 (2005)