



HAL
open science

Etude de la régulation transcriptionnelle des gènes lors du cycle érythrocytaire de *Plasmodium falciparum*

Mathieu Gissot

► **To cite this version:**

Mathieu Gissot. Etude de la régulation transcriptionnelle des gènes lors du cycle érythrocytaire de *Plasmodium falciparum*. Biochimie [q-bio.BM]. Université Pierre et Marie Curie - Paris VI, 2005. Français. NNT: . tel-00008492

HAL Id: tel-00008492

<https://theses.hal.science/tel-00008492>

Submitted on 14 Feb 2005

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.

THESE DE DOCTORAT DE L'UNIVERSITE PIERRE ET MARIE CURIE (PARIS 6)

Spécialité Parasitologie

Présentée par

Mathieu GISSOT

Pour obtenir le grade de
DOCTEUR DE L'UNIVERSITE PARIS 6

**Etude de la régulation transcriptionnelle des gènes lors du
cycle érythrocytaire de *Plasmodium falciparum*.**

*Implication des acteurs de la régulation transcriptionnelle dans les évènements
clefs du cycle érythrocytaire.*

Soutenue le 07 février 2005

Devant le jury composé de

Pr. Vincent MARECHAL
Pr. Christian VIVARES
Dr. Jean-François DUBREMETZ
Dr. Artur SCHERF
Dr. Isabelle ROSINSKI-CHUPIN
Dr. Catherine VAQUERO

Président du jury
Rapporteur
Rapporteur
Examineur
Examineur
Directeur de thèse

*Homme libre, toujours tu chériras la mer !
La mer est ton miroir ; tu contemples ton âme
Dans le déroulement infini de sa lame,
Et ton esprit n' est pas un gouffre moins amer.*

Baudelaire.

Remerciements

Je tiens tout d'abord à remercier chaleureusement Dominique Mazier pour l'accueil qu'elle m'a réservé au sein de son unité de recherche et pour son œil bienveillant sur mes travaux.

Merci aux membres de mon jury d'avoir bien voulu se pencher quelques moments sur mon travail afin de l'enrichir de leurs critiques et commentaires. Je tiens plus particulièrement à remercier le Pr. Vivares et le Dr. Dubremetz d'avoir accepté avec enthousiasme d'être rapporteur de ma thèse.

Merci à Catherine Vaquero de m'avoir donné la chance, que d'autres m'avaient refusée, de me lancer dans le grand bain de la Science. Merci de m'avoir guidé, accompagné et dirigé sur le chemin sinueux de ma thèse et pendant plus de 3 années. Si la connaissance est le savoir qui fait grandir, c'est cet inestimable trésor que m'a transmis Catherine. En me permettant librement de me confronter à la réalité de la paillasse, d'appréhender les difficultés de la réflexion scientifique et de m'impliquer dans la rédaction des articles, elle m'a offert la précieuse opportunité de grandir. Sans ces discussions passionnées, ses conseils avisés et cette liberté encadrée, mon travail n'aurait pu être identique. Pour tout ça et pour ton impressionnante capacité de travail, merci !

Merci à Sylvie Briquet de m'avoir offert son savoir et sa science sans contrepartie, d'avoir su m'aider à partager un peu plus. Merci pour sa chaleureuse présence, sa compétence et son professionnalisme.

Merci à Philippe Refour d'avoir, avec abnégation et au prix de longs mois de sa thèse, construit, réalisé et produit la biopuce à ADN qui a été l'outil de base de mon travail. Merci pour son amitié et son inflexible goût pour les choses bizarres qui peuvent (éventuellement) se manger. Non, je ne ferais pas d'import de Root beer !!

Merci à l'ensemble des personnels de l'unité 511 qui ont fait de mon séjour un moment forcément inoubliable. Plus particulièrement et dans le désordre merci à Charlotte, Carine, Anaïs, Julie, Virginie, Saddi, Maryse, Yvette, Nassira, Liliane, Monique, Olivier, Anthony, Jean-François, Maurel, Mr. Farahti, Mr. Sobolovic et Nicolas. Merci aux habitants de la plateforme P3S et notamment à Abiba et Laurent, pour leurs conseils et pour leur gentillesse. Merci aux stagiaires que j'ai torturées.

Merci aux responsables du Magistère de génétique de l'Université Paris 7, de m'avoir offert, tout au long de ma thèse, d'ouvrir une fenêtre sur le monde de l'enseignement.

Merci à mes parasites, morts dans d'affreuses souffrances.

Pour finir, merci à mes parents et à mon frère d'avoir toujours cru en moi, de m'avoir donné cette envie et cette curiosité qui me poussent maintenant à aller au-delà de nos frontières. Merci pour votre présence rassurante et vos mots, prononcés ou pas, qui forment ce pont entre ce que je suis et mon devenir. Merci à mon frère d'avoir ouvert la voie de la Science et de me permettre de suivre sa piste. Merci aux yeux d'Emilie, à leur force et à leur lumière.

SOMMAIRE

Liste des abréviations	1
INTRODUCTION	3
1. Le Paludisme	4
1.1 Aspects historiques	4
1.2 Aspects géographiques et démographiques	5
1.3 Aspects économiques	7
1.4 Aspects cliniques	9
2. Les cycles de vie de Plasmodium	11
2.1 Cycle parasitaire des plasmodies infectant l'homme	11
2.1.1 Le cycle exo-érythrocytaire	12
2.1.2 Le cycle érythrocytaire	12
2.1.3 Gamétocytogénèse	15
2.1.4 Cycle sexué chez le moustique	17
2.2 Cycle cellulaire	18
3. Le génome de Plasmodium falciparum	21
3.1 Composition et organisation du génome	21
3.2 La structure des chromosomes	22
3.3 Vue générale de la composition protéique	23
3.4 La composition en protéines assignées à la régulation génétique.	26
3.4.1 La machinerie générale de transcription	27
3.4.2 Régulateurs de la structure de la chromatine	28
3.4.2.a L'organisation de la chromatine	28
3.4.2.b Modification et remodelage de la chromatine	29
3.4.3 Facteurs de transcription spécifiques de séquence	31
3.4.3.a La superclasse des Domaines Basiques	32
3.4.3.b La superclasse des coordinateurs en doigt de Zinc	32
3.4.3.c La superclasse des Hélice-Tour-Hélice	34
3.4.3.d La superclasse des facteurs à architecture bêta	35
3.4.3.e Les FT non classées	37
3.4.4 Coactivateurs et corépresseurs transcriptionnels	37
3.4.4.a Le complexe «Mediator»	37
3.4.4.b Autres co-facteurs	38
3.4.5 Facteurs impliqués dans la régulation post transcriptionnelle	39
3.4.5.a Au niveau de l'ARNm	39

3.4.5.b Au niveau de la traduction	40
3.4.5.c Au niveau post-traductionnel	41
4. La régulation de l'expression des gènes chez <i>P. falciparum</i>	44
4.1 Etude à grande échelle de l'expression des gènes	44
4.1.1 Les études pré-génome	44
4.1.2 Les études post-génome	45
4.2 La régulation transcriptionnelle	48
4.2.1 La structure des promoteurs	48
4.2.2 Les éléments de régulation	50
4.2.3 Les particularités	53
4.2.3.a Site d'initiation de la transcription	53
4.2.3.b Les régions homopolymériques	54
4.2.4 Les facteurs de transcription	55
4.3 Régulation post-transcriptionnelle	55
4.3.1 Régulation de l'épissage	55
4.3.2 Régulation de la stabilité des ARNm et de la traduction	56
4.3.3 ARN anti-sens	57
5. Présentation des travaux de thèse	60
MATERIEL & METHODES	62
1. Matériels	63
1.1 Clones de <i>P. falciparum</i>	63
1.2 Souches bactériennes et plasmides	63
1.3 Anticorps	63
2. Méthodes	64
2.1 Puce à ADN ciblée	64
2.1.1 Les gènes sélectionnés	64
2.1.2 Production des ADNc marqués et hybridation	65
2.1.3 Acquisition et analyse des données	65
2.2 RT-PCR quantitative en temps réel	65
2.3 Immuno-précipitation de la chromatine	68
RESULTATS	70
1. Recherche de facteurs impliqués dans la transition de la prolifération asexuée à la différenciation sexuée.	71
1.1. Résumé	71

Article 1.	72
<i>Transcriptome of 3D7 and its gametocyte-less derivative F12 Plasmodium falciparum clones during erythrocytic development using a gene-specific microarray assigned to gene regulation, cell cycle and transcription factors.</i>	
2. Caractérisation fonctionnelle et rôle de PfMyb1, un facteur de transcription spécifique de séquence, dans la régulation des gènes au cours du cycle érythrocytaire.	84
2.1. Caractérisation fonctionnelle de PfMyb1	84
2.1.1 Résumé	84
Article 2.	86
<i>Characterization of PfMyb1 transcription factor during erythrocytic development of 3D7 and F12 Plasmodium falciparum clones.</i>	
2.1.2 Résultats complémentaires	92
2.1.2.a Expression de la protéine PfMyb1 recombinante	92
2.1.2.b Etude de la cinétique d'expression du transcrit <i>pfmyb1</i>	93
2.1.2.c Etude de la cinétique d'expression de la protéine PfMyb1	95
2.2. Rôle de PfMyb1 lors du cycle érythrocytaire et caractérisation de son réseau de régulation.	97
2.2.1 Résumé	97
Article 3.	98
<i>PfMyb1, a Plasmodium falciparum transcription factor, is required for intra-erythrocytic growth and controls key genes for cell cycle regulation.</i>	
DISCUSSION	137
BIBLIOGRAPHIE	152
ANNEXES	167
1. Numéros d'accèsion des protéines liées à la régulation de l'expression des gènes chez <i>P. falciparum</i>.	168
1.1. Protéines annotées par nos soins.	168
1.2. Protéines annotées par Coulson et collaborateurs [58].	171
2. Article 4 : Vers une utilisation des bio-puces à ADN pour l'étude d'isolats plasmodiaux de terrain.	172
3. Données et matériels supplémentaires de l'article 1.	180
4. Données et matériels supplémentaires de l'article 3.	192

Liste des abréviations :

A: adénine ;
aa: acide aminé ;
ADN: acide désoxyribonucléique ;
ADNc: ADN complémentaire à l'ARN ;
ARN: acide ribonucléique ;
ARNdb: ARN double brin ;
ARNm: ARN messenger ;
ARNr: ARN ribosomaux ;
ARNt: ARN de transfert ;
ChIP: immuno-precipitation de la chromatine ;
CO: Constans ;
C_T: cycle threshold ;
DNase: désoxyribonucléase ;
dNTP: désoxyribonucléotides tri-phosphate ;
DPE: downstream promoter element ;
DTT: dichlorodiphényltrichloroéthane ;
FT: facteur de transcription ;
HTH: Hélice-Tour-Hélice ;
Inr: Initiator ;
IRP: iron regulatory protein ;
kb: kilobase ;
Mb: mégabase ;
MRE: myb regulatory element ;
NAP: nucleosome assembly protein ;
NLS: site de localisation nucléaire ;
p.: page ;
p.i: post-invasion ;
pb: paire de bases ;
PCR: réaction de polymérisation en chaîne ;
PIB: produit intérieur brut ;
qPCR: PCR quantitative en temps réel ;

Liste des abréviations

RNase: ribonucléase ;

RPMI: milieu "*Roswell Park Memorial Institute*" ;

RRM: *RNA Recognition Motif* ;

RT: transcription reverse ;

T: thymine ;

TAF: facteurs associés à la TBP ;

TBP: *TATA-binding-protein* ;

INTRODUCTION

1. Le Paludisme

Le paludisme est causé par un parasite hématozoaire de la famille des Apicomplexes et du genre *Plasmodium* et est transmis à l'homme par des moustiques femelles du genre *Anopheles*. Chez l'homme, quatre espèces de plasmodies sont infectieuses : *P. falciparum*, *P. vivax* et *P. ovale* et *P. malaria*. Mais l'essentiel de la mortalité est due à *P. falciparum*. On retrouve trace de la description des symptômes de cette maladie depuis l'Antiquité. Cependant le chemin a été long avant d'aboutir à l'identification du vecteur et de l'agent causal de la principale maladie parasitaire humaine.

1.1 Aspects historiques

L'homme et le parasite responsable du paludisme ont une longue histoire de co-évolution. Cette relation hôte-parasite marque leur évolution commune et c'est ainsi qu'il y a 10 000 ans, la rapide expansion de *Plasmodium falciparum* coïncide avec la croissance de la population humaine [1]. Il n'est donc pas étonnant que depuis l'Antiquité, on connaisse l'existence de fièvres intermittentes et parfois mortelles. On trouve ainsi dans les anciens manuscrits chinois et égyptiens, de même que dans la littérature grecque et romaine, des témoignages relatifs aux fièvres épidémiques dont les signes ressemblent beaucoup à ceux du paludisme. Mais il faut attendre 1717 pour que Giovanni Lancisi attribue cette maladie aux émanations nocives des marais, d'où le nom *malaria*, de l'italien « mal aria » ou air vicié ; puis enfin 1880 pour que le Docteur Laveran, un médecin de l'armée française, soit témoin de l'exflagellation d'un gamétocyte mâle lors de l'observation au microscope d'échantillons de sang de patients impaludés. Il permit ainsi la découverte d'un hématozoaire parasite, responsable du paludisme.

Dès 1882, on émit l'hypothèse de la transmission de la maladie par un moustique. Cette hypothèse fut confirmée quelques temps après, en 1897, quand Ronald Ross fit la découverte de l'importance du moustique dans le cycle de vie du paludisme aviaire et décrivit des kystes paludéens dans les parois de l'estomac d'*Anophèles* [2]. En 1898, Grassi décrivit le cycle de transmission complet du parasite et prouva que le paludisme observé chez l'homme est transmis par des moustiques du genre *Anophèles* [3].

Cependant, malgré un enrichissement progressif des connaissances mondiales concernant le cycle du parasite et la transmission de la maladie à l'homme, ce fléau touche encore la quasi-majorité de la population mondiale.

1.2 Aspects géographiques et démographiques

L'aire de répartition géographique du paludisme est limitée par l'aire d'extension de ses vecteurs, les moustiques du genre *Anophèles*, et par les conditions climatiques permettant au parasite d'effectuer son développement chez le moustique. Vers 1900 se situe la date à laquelle on considère que l'homme n'est pas intervenu massivement sur la distribution géographique et démographique du paludisme (Figure 1). Cette répartition prend en compte toutes les espèces de *Plasmodium* infectant l'homme. En tenant compte de ces critères, vers 1900, 77 % de la population mondiale est touchée par le paludisme dans plus de 140 pays touchés sur 53 % de la surface mondiale [4].

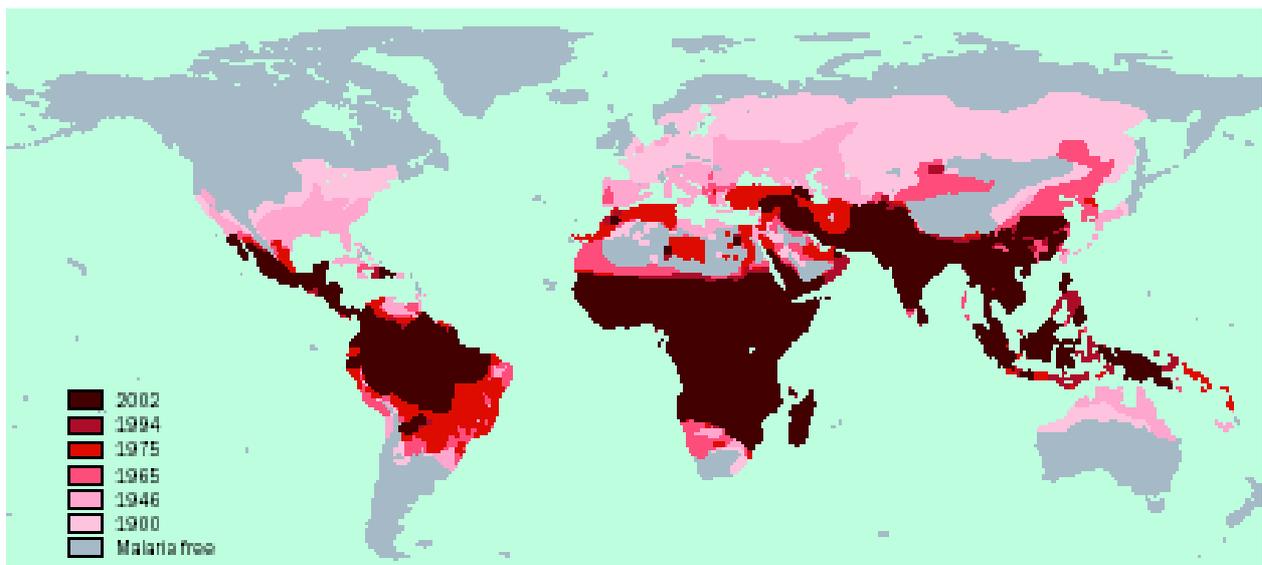


Figure 1. Distribution des zones de risque de paludisme de 1900 à 2002 d'après [4].

Les efforts afin de contrôler le paludisme ont fortement réduit sa distribution au cours du 20^{ème} siècle. En effet, la figure 1 permet de se rendre compte de la diminution de près de la moitié de la zone de risque pour le paludisme. Si ces efforts ont été largement récompensés (Figure 1), on peut identifier deux périodes: de 1900 à 1994, où l'on constate que la surface mondiale de risque et la population exposée diminuent (Tableau 1), puis de 1994 à 2002, voire 2010 (projection), la quantité de population exposée augmente à surface constante (Tableau 1).

Tableau 1. Population en zone de risque de paludisme d'après [4].

Table 1. Global population at risk from malaria from preintervention to 2010 (-1900-2010)							
Time	Global population	Land area malarious		Countries at risk	Population exposed		
Years	<i>n</i>	<i>km²</i>	%	<i>n</i>	<i>n</i>	%	
1900	1 158 409 472	77 594 480	53-16	140	892 373 056	77-03	
1946	2 391 400 960	58 565 752	40-12	130	1 635 815 808	68-40	
1965	3 363 417 344	53 492 988	36-65	103	1 924 360 320	57-21	
1975	4 085 759 488	48 075 780	32-93	91	2 121 086 592	51-91	
1992	5 419 255 808	43 650 812	29-90	88	2 565 702 144	47-34	
1994	5 582 432 256	39 537 020	27-08	87	2 570 555 136	46-05	
2002	6 204 095 488	39 758 172	27-24	88	2 996 419 584	48-30	
2010	6 807 085 056	39 758 172	27-24	88	3 410 862 080	50-11	

Cette stagnation de la capacité des hommes à faire reculer les zones de transmission du paludisme est certainement à mettre en corrélation avec la classe d'endémicité dans laquelle se trouvent les zones où les efforts de contrôle et le développement ont porté leurs fruits. Par exemple entre 1900 et 2002, les zones en classe épidémique ont été totalement éradiquées tandis que les zones holoendémiques n'ont pas changé de surface [4]. Ces informations reflètent la difficulté croissante dans la maîtrise de la maladie à mesure que l'intensité de la transmission augmente. Elles soulignent aussi que les méthodes d'éradication de masse des vecteurs moustiques avec notamment l'utilisation du dichlorodiphényltrichloroethane (DDT) accompagnées des évolutions positives du développement économique et sanitaire dans ces régions ont permis un contrôle efficace de la maladie. Ce premier programme d'éradication globale du paludisme (1955-1969) initié par l'organisation mondiale de la santé introduisit donc l'usage massif du DDT ainsi que l'utilisation d'une chemoprophylaxie dans un second temps. L'essentiel des résultats connus aujourd'hui est le résultat de cet effort global, tandis que pendant la période 1965-1992, les résultats obtenus sont le fait des efforts particuliers de chaque pays. Cependant, ces derniers efforts ont eu des résultats limités sur la période 1992-2002 (Tableau 1). C'est pourquoi en 1998 a été lancé le programme « Roll back Malaria » dont les objectifs ambitieux sont de réduire de moitié le risque de transmission, la morbidité et la mortalité due au paludisme d'ici 2010 [5]. Afin d'accéder à ces résultats, ce programme prévoit d'atteindre 60 % de la population d'enfants et de femmes enceintes protégés par une moustiquaire traitée aux insecticides, 60 % des cas de paludisme traités sous 24 h après la déclaration des symptômes, 60 % des femmes enceintes recevant une thérapie présomptive intermittente, et 60 % des épidémies détectées dans les deux semaines de leur apparition et traitées correctement pendant les deux semaines suivantes [5].

La géographie du paludisme s'est maintenant concentrée à des régions où le taux de transmission semble être l'ultime rempart devant lequel les méthodes actuelles restent inopérantes. Par exemple, le caractère grandement anthropophile des moustiques sévissant en

Afrique Sub-Saharienne résulte à un taux de 121 piqûres infectieuses par personne et par an [6]. Cette donnée concomitante avec la dominance de *P. falciparum* dans ces régions, fait de l'Afrique Sub-Saharienne la zone dans laquelle se concentre plus de 85 % des cas de décès causés par le paludisme. En 2002, le paludisme était responsable de plus d'un million de morts, en majorité des enfants de moins de 5 ans et des femmes enceintes, et près de 400 millions de personnes étaient infectées [7]. Ce fléau cause donc la mort d'un enfant africain toutes les 40 secondes, plaçant le paludisme dans les 3 maladies infectieuses causant mondialement le plus de morts.

1.3 Aspects économiques

La co-évolution de l'homme et *Plasmodium* a conduit à imposer des contraintes aux deux partenaires de l'interaction hôte-parasite. Certains polymorphismes génétiques, comme l'anémie falciforme, ont été ainsi modifiés dans certaines régions sous cette pression de sélection [8]. Cette donnée souligne l'influence du paludisme dans les régions endémiques. Celle-ci se traduit d'ailleurs en termes économiques, la relation entre paludisme et pauvreté étant depuis longtemps établie. Pour s'en convaincre, il suffit de comparer la carte de la distribution du risque palustre (Figure 1) avec la carte de distribution du produit intérieur brut (PIB) (Figure 2). La pauvreté est concentrée dans les régions tropicales et sub-tropicales dans les mêmes limites que celles décrites pour la transmission du paludisme.

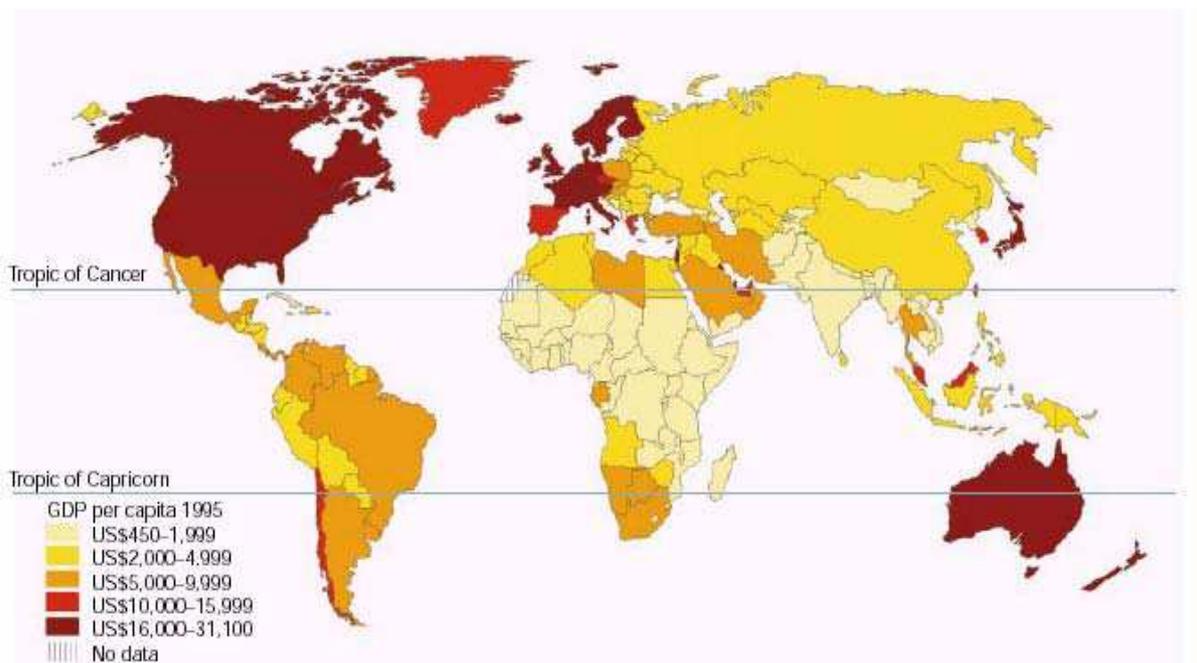


Figure 2. Distribution globale du PIB d'après [9]

Evidemment, la relation paludisme-pauvreté est à double sens: le paludisme conduit à plus de pauvreté et cette dernière est responsable de l'augmentation du paludisme. Cette maladie n'est pas la seule cause du retard de développement des pays qu'elle touche mais le fardeau économique qu'elle représente participe certainement au ralentissement de la croissance de ces pays. Ce fardeau se manifeste par exemple par des pertes de près de 20 % de revenu pour certains pays africains en zone d'endémie [9] et près de 25 % du revenu d'une famille africaine [10]. Le coût engendré par cette maladie est direct pour une famille quand il est lié aux dépendances médicales (déplacement vers le centre médicalisé, diagnostic, traitement) ou de prévention (moustiquaire, insecticide....) ou simplement au coût de la perte des revenus (présents et futurs) causée par la diminution du temps travaillé (absentéisme) ou mort d'un membre, et pour l'Etat concernant les dépendances dues au contrôle du vecteur, aux structures de santé, à la prévention et à la recherche et à la diminution de la production du pays.

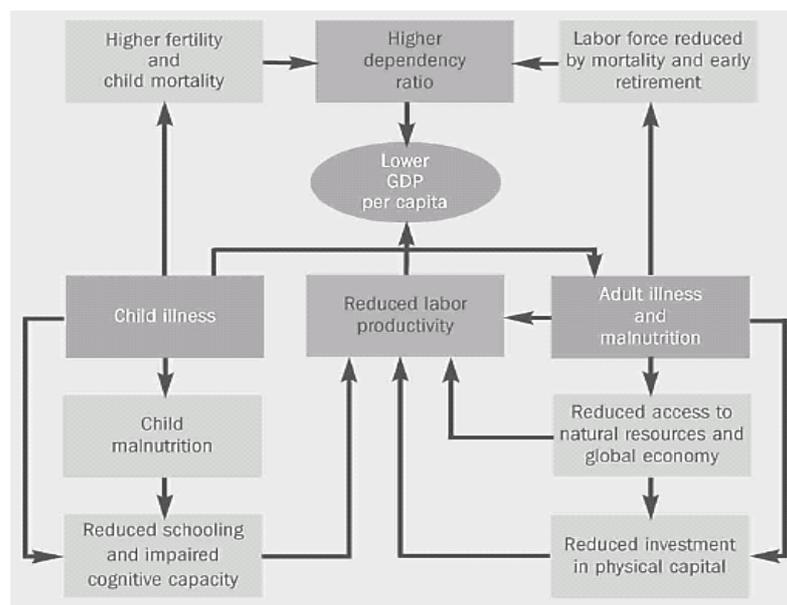


Figure 3. Impact d'une maladie sur le PIB d'après [10]

Cette maladie influe donc sur l'organisation de la famille et de l'Etat en zone d'endémie. Les effets indirects sur la famille et l'Etat sont dépendants de ces changements de comportement. Cette maladie influe sur l'éducation des enfants, la démographie, les migrations et les économies de la famille. De plus, elle joue un rôle négatif sur le commerce, le tourisme et les investissements [9]. Par exemple, en zone d'holoendémie, le paludisme compte pour 25 % de toute la mortalité des enfants entre 0 et 4 ans [11]. Cette dramatique mortalité influe sur les décisions concernant la fertilité des familles. Il s'ensuit une augmentation de la fertilité et une

baisse de l'investissement pour l'éducation de chaque enfant, une perte de capital due à l'augmentation du nombre d'enfants « non productifs » et une influence négative sur la capacité à économiser [9]. Le paludisme a aussi un effet sur l'état de nutrition des enfants, l'absentéisme scolaire et le développement des capacités motrices et du cerveau. Plus généralement, le paludisme a une influence directe sur le capital humain menant par bien des chemins à une perte de PIB (Figure 3) pour le pays touché [10].

Une dernière donnée souligne la place de cette maladie dans la société africaine : près de 42% des tentatives de suicides chez les adolescents sénégalais se font par absorption de chloroquine [12].

1.4 Aspects cliniques

L'accès palustre consiste généralement en des accès de fièvres accompagnés de multiples symptômes comme des malaises, nausées, maux de tête, douleurs musculaires et diarrhées modérées. La faible spécificité de ces symptômes est un des facteurs aggravant de l'endémie palustre, car ils sont souvent attribués à des infections intestinales. Ces manifestations sont le fruit de l'infection par le parasite des érythrocytes humains (voir cycle section 2.1) et de sa rapidité de développement. Ceci implique que la fréquence des épisodes de fièvres est dépendante de l'espèce de *Plasmodium* : trois jours pour *P. falciparum*, *P. vivax* et *P. ovale*, et quatre jours pour *P. malariae*. Les complications pouvant survenir sont dépendantes de plusieurs facteurs, dont notamment l'immunité éventuellement acquise par l'hôte (son histoire immunologique), le fond génétique de l'hôte mais aussi du degré de chimiorésistance et de virulence de la population de *Plasmodium* impliquée dans l'infection. Dans le cas de *P. falciparum*, il existe trois complications majeures qui peuvent être concomitantes et ont un pronostic mortel: l'anémie grave, le syndrome de détresse respiratoire et le neuropaludisme.

L'anémie grave est définie par un hémocrite inférieure à 0,15 % associée à la présence de *P. falciparum*. Cette complication est la cause principale de mortalité palustre en région de transmission constante même si cette donnée peut être surestimée car l'anémie peut avoir d'autres causes (malnutrition). Elle est causée par la corrélation de deux évènements: d'une part la destruction des érythrocytes par le parasite et, d'autre part, la réduction de la production d'érythrocytes durant la maladie [13]. La détresse respiratoire est souvent la cause de la survenue d'un œdème respiratoire chez l'adulte mais est dépendante d'une acidose métabolique chez l'enfant. Cette dernière est sans doute causée par le parasite lui-même et la

réponse des cellules humaines (expression de cytokines) à l'infection [13]. Elle représente près de 15 % de la mortalité palustre chez l'enfant [14]. Enfin, le neuropaludisme est une complication qui ne survient que lors de l'infection par *P. falciparum*. Elle consiste en une encéphalopathie diffuse se traduisant par un syndrome neurologique qui conduit généralement à un coma qui peut être suivi par la mort. La séquestration des globules rouges parasités, ainsi que la réponse de l'épithélium et du système immunitaire à cette obstruction des vaisseaux sanguins, semblent avoir un rôle important dans cette complication [13]. La susceptibilité à cette dernière complication dépend fortement du statut immun de l'hôte et touche donc principalement les enfants en Afrique et les adultes en Asie. La mortalité consécutive est d'environ 20 % chez l'adulte et 15 % chez l'enfant. De plus, 5 % des adultes et 10 % des enfants rescapés de cette complication ont des séquelles neurologiques persistantes (surdité, cécité, hémiplégie, psychoses, retard mental, troubles du comportement) [15].

Ces complications sont liées aux différents évènements du cycle de vie de *Plasmodium*, dont nous allons découvrir la complexité.

2. Les cycles de vie de *Plasmodium*

2.1 Cycle parasitaire des plasmodies infectant l'homme

Le cycle biologique de *Plasmodium* peut être caractérisé par deux phases :

- une phase de prolifération chez l'homme (hôte intermédiaire) par multiplication asexuée, appelée schizogonie. Cette phase se tient dans deux sites différents chez l'homme et de manière chronologique: d'abord au sein des hépatocytes dans le foie (cycle exo-érythrocytaire, Figure 4A), puis dans les érythrocytes circulants (cycle érythrocytaire, Figure 4B). Elle se tient aussi chez le moustique à la suite de la phase sexuée.
- une phase de différenciation sexuée suivie d'une multiplication asexuée, appelée sporogonie (Figure 4C), qui commence chez l'homme par la production de gamétocytes mâles et femelles (gamétocytogénèse) et qui continue chez le moustique par la maturation de ceux-ci en gamètes mâles et femelles et la fin du cycle sexué (Figure 4).

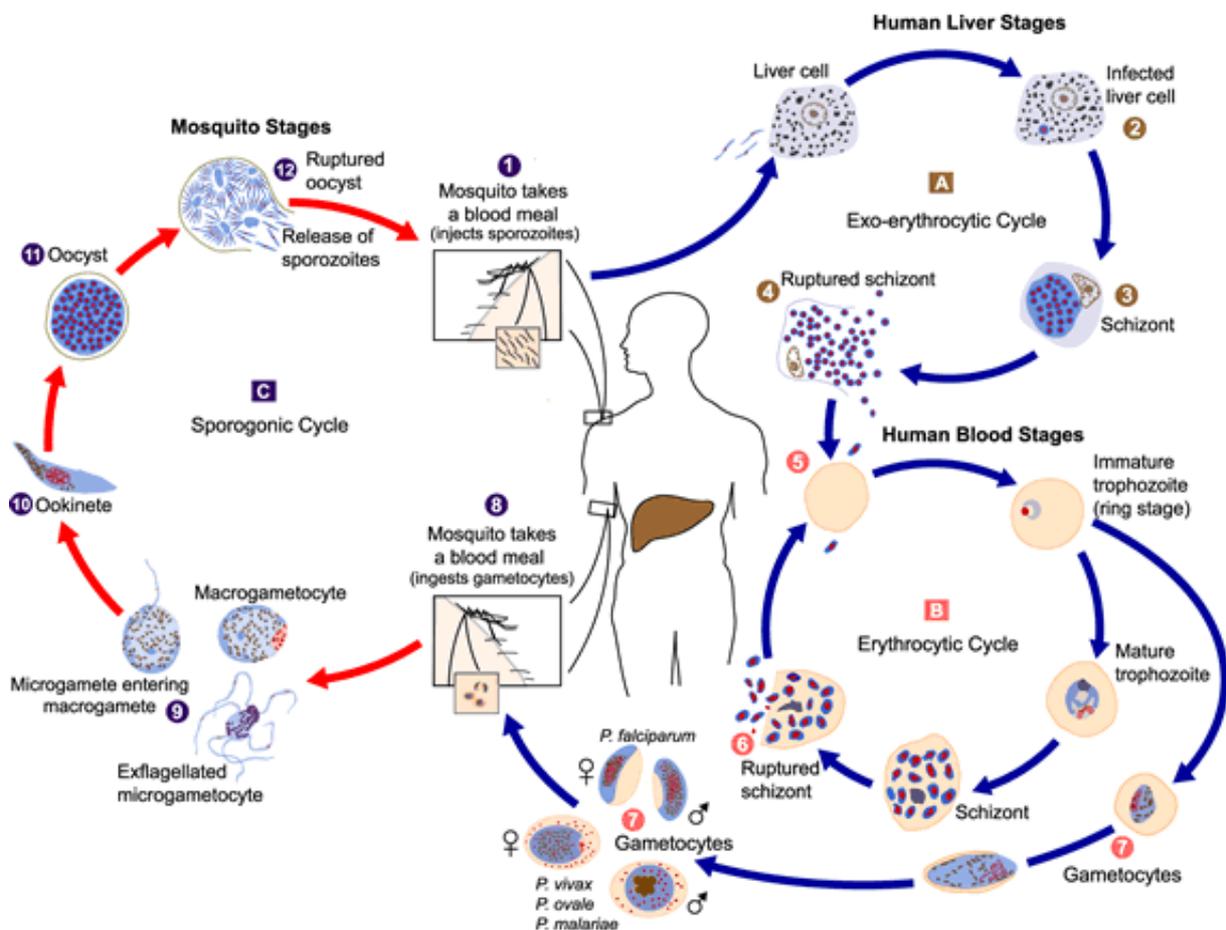


Figure 4 . Cycle de développement des plasmodies infectant l'homme d'après [16].

2.1.1 Le cycle exo-érythrocytaire

La schizogonie commence chez l'homme par l'injection sous-cutanée, lors du repas sanguin du moustique, de sporozoïtes contenus dans sa salive.

Même si l'on peut dénombrer un grand nombre de sporozoïtes dans les glandes salivaires, seul un petit nombre semble être inoculé. Le destin des sporozoïtes dépend ensuite de leur capacité à migrer sur des distances de plusieurs micromètres en quelques minutes. Les parasites entrent alors dans la circulation sanguine et voyagent rapidement vers le foie. L'entrée du sporozoïte dans le foie est un processus dépendant de plusieurs migrations successives à travers deux types cellulaires [17]. Après avoir été arrêté par l'épithélium, le sporozoïte migre le long de l'endothélium pour atteindre les cellules macrophagiques de Kupffer. Il envahit activement ces cellules et semble capable de les traverser sans dommages [17]. Une fois en contact avec les hépatocytes, le sporozoïte envahit rapidement une cellule mais ne semble pas se développer dans cette première cellule. Le sporozoïte serait capable de migrer à travers plusieurs hépatocytes avant de finalement commencer son développement dans un nouvel hépatocyte [18]. Au début de la cascade de différenciation, le sporozoïte nouvellement entré dans l'hépatocyte est situé à l'intérieur d'une membrane parasitophore près du noyau de la cellule et ses organelles caractéristiques sont désassemblées sous 24 heures [17] afin de former le trophozoïte hépatique. Cette forme est uninucléée. Lors de l'infection par *P. ovale* et *P. vivax*, une partie des trophozoïtes hépatiques reste sous forme quiescente uninucléée dans les hépatocytes. Cette forme, appelée hypnozoïte pourrait être à l'origine des rechutes observées lors d'infections par ces deux plasmodies.

Le parasite commence ensuite une division cellulaire qui conduit, après six à sept jours, à la production d'un schizonte hépatique dépassant la taille originelle de l'hépatocyte et composé de milliers de mérozoïtes, forme infectieuse pour les érythrocytes (Figure 4 A3). Après rupture de la membrane parasitophore du schizonte et de la membrane plasmique de l'hépatocyte (Figure 4 A4), les mérozoïtes sont libérés dans la circulation sanguine, où ils débent le cycle érythrocytaire.

2.1.2 Le cycle érythrocytaire

Ce cycle dure environ 48 h chez *P. falciparum*. Il commence par l'invasion d'un érythrocyte sain par un mérozoïte. Celui-ci est morphologiquement et fonctionnellement adapté pour l'invasion de l'érythrocyte et la formation de la vacuole parasitophore. Le pôle apical du mérozoïte contient des vésicules de sécrétion, de nombreux ribosomes ainsi qu'une

mitochondrie et le plastide. Au pôle basal, se trouvent le noyau. Un cytosquelette sous pelliculaire donne au mérozoïte une forme ovoïde [19]. Les rhoptries, les micronèmes et les granules denses sont ces vésicules de sécrétion libérant leur contenu durant le processus d'invasion, étape essentielle pour la formation de la membrane parasitophore, le remodelage et le changement de composition de la membrane de l'érythrocyte. La sécrétion du contenu de chaque vésicule est un processus séquentiel au cours de l'invasion de l'érythrocyte [20]. L'étape initiale de l'invasion est constituée par une interaction à faible affinité avec la cellule hôte par le biais de protéines de surface du mérozoïte comme MSP1 [20]. Une seconde étape conduit à la réorientation du parasite sous la dépendance d'une interaction plus forte avec la membrane de la cellule. A la suite de la réorientation du domaine apical en direction de la membrane de la cellule, un attachement et une jonction serrée avec cette dernière est la conséquence de l'exocytose du contenu micronemal. La décharge du contenu des rhoptries permet le mouvement du parasite vers l'intérieur de l'érythrocyte. Puis, la libération du contenu des rhoptries suivie par celle des granules denses sont responsables de la formation de la membrane de la vacuole parasitophore.

Après l'invasion, le parasite commence à modifier le contenu et la structure de l'érythrocyte. Le premier stade de développement est l'anneau (de 1 h post-invasion à 18 h p.i.), qui tient son nom de sa morphologie sur frottis (Figure 5). Le parasite commence à se nourrir des constituants de l'érythrocyte, l'hémoglobine. L'hème résultant de ce catabolisme va être transformée tout au long du cycle érythrocytaire en cristaux d'hemozoin, formant un pigment brun.

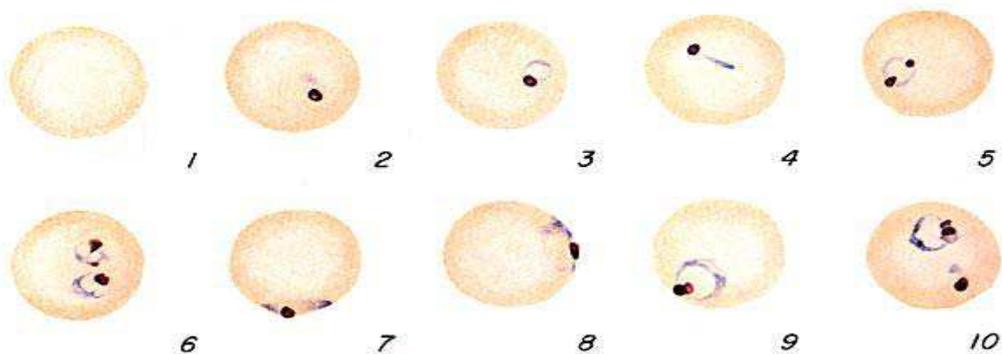


Figure 5. Développement du stade anneau, d'après [16]. 1 : érythrocyte non infecté, 2-10 : stades anneau.

Le parasite continue son développement : son volume augmente pour se transformer en trophozoïte (de 19h p.i. à 30h p.i.). Cette forme du cycle érythrocytaire est caractérisée par

une augmentation de l'export des protéines parasites à la surface de l'érythrocyte. Un réseau de membranes se met en place dans le cytosol de l'érythrocyte assurant les connections entre la vacuole parasitophore et la membrane de l'érythrocyte [21]. Des structures membranaires se forment sous la membrane de l'érythrocyte, dont la fonction pourrait être la maturation de certaines protéines avant leur export à la membrane [22]. Certaines de ces protéines parasites exportées à la membrane de l'érythrocyte, forment des agrégats qui produisent de petites protubérances appelées « knobs ». Ces changements à la surface de l'érythrocyte sont impliqués dans la capacité des érythrocytes parasités à adhérer aux cellules épithéliales. L'ingestion du cytosol de l'érythrocyte continue et conduit à la digestion de l'hémoglobine et à la formation d'hémozoïne. L'espace occupé par le trophozoïte dans le cytosol augmente à mesure que la surface de la vacuole digestive s'accroît (Figure 6).

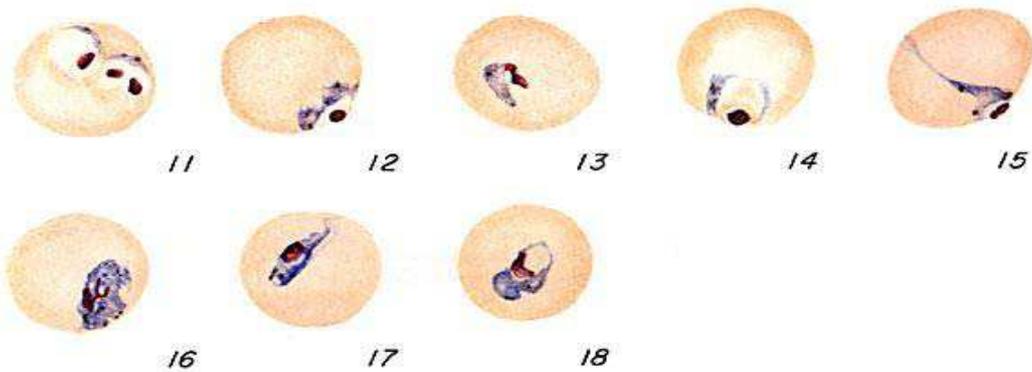


Figure 6. Développement du stade trophozoïte (11-18), d'après [16].

Le schizonte (de 31 h à 48 h p.i.) est l'ultime phase de développement du parasite au sein de l'érythrocyte, qui va aboutir à la formation de 16 à 32 mérozoïtes pour *P. falciparum* (Figure 7). La déplétion de l'hémoglobine du cytosol continue et entraîne la formation d'un cristal d'hémozoïne dense dans la vacuole digestive. La dernière division nucléaire est accompagnée de la constitution de centres de différenciation des mérozoïtes dans lesquels sont synthétisés et se mettent en place les organites apicaux du mérozoïte. Avant l'individualisation de chaque mérozoïte, chaque mitochondrie et le plastide font mouvement vers le centre du cytoplasme du schizonte. La séparation physique de chaque mérozoïte par une membrane s'opère et marque la dernière étape de la schizogonie. Finalement, la membrane de la vacuole parasitophore et la membrane de l'érythrocyte sont brisées (Figure 4B 6) par le biais d'un mécanisme actif impliquant plusieurs protéases [23] libérant les mérozoïtes et permettant d'amorcer un nouveau cycle érythrocytaire asexué.

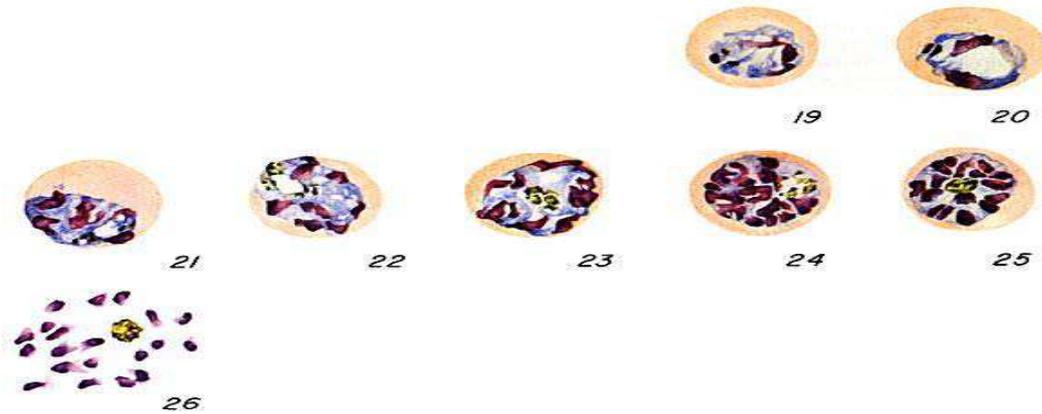


Figure 7. Développement du stade schizonte (19-26) d'après [16].

2.1.3 Gamétocytogénèse

La gamétocytogénèse est un processus qui conduit une partie des parasites à arrêter leur prolifération asexuée et à initier une différenciation sexuée. Pendant le cycle érythrocytaire, une partie des parasites sort du cycle érythrocytaire asexué pour devenir des gamétocytes, seule forme infectieuse pour le moustique (Figure 4B 7). Lors de l'infection par *P. falciparum*, l'observation de gamétocytes ne se fait que 7 à 15 jours après l'identification de stades asexués [24]. Cette transformation conduit à des changements du métabolisme, de l'expression protéique et morphologique. Les gamétocytes mâles et femelles (Figure 8) subissent un développement différent qui les conduit à la maturité au bout de cinq stades caractéristiques et de dix jours d'évolution.

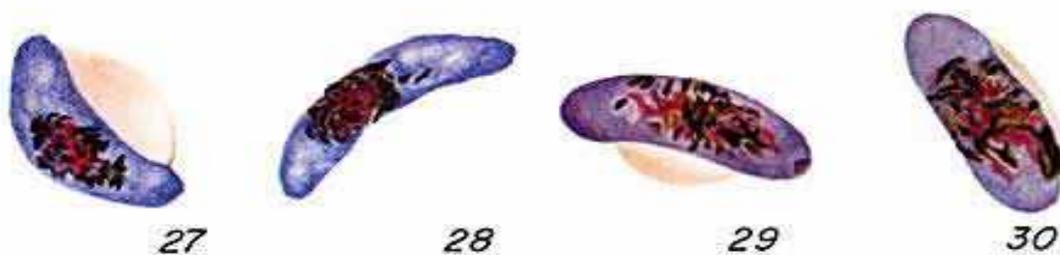


Figure 8. Parasite au stade gamétocytes d'après [16]. 27-28 : gamétocytes femelles, 29-30 : gamétocytes mâles.

Cette maturation permet la préparation du cycle sexué chez le moustique et la mise en place de processus permettant l'adaptation au changement d'environnement drastique que subit le parasite lors de son arrivée dans le moustique (température et pH). De nombreuses variations

dans le contenu en ARN et protéique des gamétocytes ont été identifiées en comparaison aux autres stades du cycle érythrocytaire. Par exemple, l'expression d'un type différent d'ARN ribosomal dans les gamétocytes et lors du cycle sexué chez le moustique souligne la spécialisation des gamétocytes.

Le processus permettant la différenciation du mérozoïte en gamétocyte n'est pour l'instant que partiellement connu. Il est établi que tous les mérozoïtes émergeant d'un même schizonte partagent tous le même destin vers la prolifération asexuée ou la différenciation sexuée [25]. Ceci impliquerait que le choix de ce destin est fait avant la formation des mérozoïtes. De plus, dans un même schizonte, chaque mérozoïte devient gamétocyte soit mâle soit femelle [26], impliquant une spécification de la voie de différenciation lors du choix vers le développement en gamétocyte.

L'influence de nombreux facteurs environnementaux a été décrite dans le choix de l'initiation du cycle sexué [27]. L'immunité de l'hôte figure parmi ces facteurs. Certaines hormones de l'hôte et l'activité de drogues anti-malariques montrent l'importance des facteurs de « stress » pour le parasite dans le choix de la voie de différenciation. Le rôle des voies de transduction du signal dans le mécanisme d'acheminement du message de « stress » est donc probable. Cette hypothèse est étayée par la possible implication des protéines G trimériques [28] et de l'AMP cyclique [29] dans l'orientation vers la différenciation sexuée. De plus, l'âge des érythrocytes et l'état de leur environnement intracellulaire sont des paramètres qui pourraient avoir un rôle dans ce processus [30].

La régulation de ce mécanisme est aussi intimement liée à l'état de la population de parasites, une relation inverse ayant été établie entre le nombre de parasites asexués et la quantité de conversions en gamétocytes [27]. En outre, cette relation semble dépendante d'une régulation autocrine par la sécrétion de facteurs diffusibles [31] ou par la libération du contenu des érythrocytes parasités lors de la libération des mérozoïtes [32].

Les effecteurs dans le parasite du message de différenciation délivrés par l'environnement extérieur sont encore inconnus. Cependant, certains gènes ont été identifiés comme des marqueurs précoces de la différenciation en gamétocyte. C'est le cas notamment de *pfg27* et *pfs16*, dont l'expression transcriptionnelle commence à l'initiation de la gamétocytogénèse [33]. Ces deux gènes ont été clairement impliqués dans le processus de gamétocytogénèse, car l'abolition de leur expression conduit à une incapacité totale de produire des gamétocytes [34] ou à une production amoindrie [35]. L'expression de ces gènes est d'ailleurs altérée, voire abolie, dans certains clones de *P. falciparum* qui sont incapables à produire des gamétocytes [36]. Des clones ayant ce type de phénotype ont été identifiés, mais ils comportaient

généralement une délétion partielle ou totale du chromosome 9 [37]. Cependant, certains clones dérivant d'une culture continue du clone 3D7, producteur de gamétocytes et possédant une incapacité partielle ou totale à produire des gamétocytes, portent un chromosome 9 entier et un caryotype apparemment inchangé [36]. L'ensemble de ces données implique qu'il existe de nombreux facteurs génétiques régulant ce processus mais aucun acteur de la régulation de celui-ci n'est aujourd'hui connu.

2.1.4 Cycle sexué chez le moustique

Lors d'un repas sanguin chez un sujet infecté, le moustique procède à l'ingestion d'érythrocytes parasités, parmi lesquels seuls les gamétocytes sont infectieux pour le moustique. A l'arrivée dans l'intestin du moustique, les gamétocytes se différencient rapidement en gamètes mâles et femelles. Ce processus est appelé gamétogenèse. Dans la circulation sanguine, les gamétocytes ont arrêté leur différenciation depuis plusieurs jours. Dès leur entrée dans l'intestin du moustique, ils se différencient en quelques secondes. Les gamétocytes mâles et femelles lysent l'érythrocyte dans lequel ils résident. Une fois hors de la cellule hôte, les macrogamètes (femelles) sont prêtes pour être fertilisées, tandis que les microgamétocytes (mâles) entrent en division cellulaire, répliquent leur génome trois fois et subissent trois divisions endomitotiques. Dans le même temps, huit axonèmes sont assemblés, dont le rôle est de permettre la motilité de chacun des huit microgamètes. L'exflagellation est le mécanisme permettant aux microgamètes de sortir du corps résiduel du gamétocyte mâle. La gamétogenèse est initiée grâce à une petite molécule produite par le moustique, l'acide xanthurique, et sous condition d'une température inférieure de 5°C à celle de l'hôte vertébré. Lors de la microgamétogenèse, cette molécule semble déclencher un flux calcique qui agit comme régulateur de la progression du cycle cellulaire par l'intermédiaire d'une kinase dépendante du calcium [38]. La fertilisation du macrogamète par un microgamète est réalisée dans les quinze minutes suivant l'exflagellation. Le zygote ainsi formé est capable de se différencier, après une étape de méiose, en ookinète, forme invasive pour l'intestin du moustique. A ce stade de développement, l'ookinète traverse l'épithélium intestinal du moustique et se différencie alors en oocyste. Le génome est alors répliqué un grand nombre de fois. Tous les éléments sont alors produits et assemblés afin de former des sporozoïtes. Ces formes sont spécialisées dans la migration et l'invasion, comme en témoigne la présence d'organelles telles que les rhoptries. Les sporozoïtes sont libérés de l'oocyste dans la cavité de l'hémocoèle. Ils migrent ensuite le long de la partie basale de l'épithélium intestinal ou sont transportés par le flux de l'hémolymphe jusqu'aux glandes salivaires. Une fois l'épithélium des

glandes salivaires reconnu et après réorientation, le sporozoïte franchit la membrane basale des cellules sécrétoires et envahit une cellule en formant une vacuole parasitophore. Les parasites migrent alors jusqu'à la cavité sécrétoire et, finalement, au conduit salivaire [39]. L'ensemble de ces événements, depuis la production du sporozoïte dans l'oocyste jusqu'à son arrivée dans le conduit salivaire, entraîne une maturation qui permet de le rendre infectieux. Cette maturation se traduit par une expression différenciée de certains gènes, dont la fonction semble essentielle au caractère infectieux pour l'homme des sporozoïtes des glandes salivaires [40]. Le cycle est bouclé lors du repas sanguin du moustique, lequel effectue alors une injection sous-cutanée de sporozoïtes.

2.2 Cycle cellulaire

Le cycle de *P. falciparum* est caractérisé par une succession de prolifération et de différenciation, dont la régulation est cruciale pour la complétion du cycle. Pourtant, l'étude de cette orchestration et de sa régulation n'a pu être accomplie en raison des difficultés techniques liées à la manipulation des stades en dehors de ceux érythrocytaires. C'est pourquoi les seules données présentées ici ont pour objet cette phase du cycle, responsable de la pathologie chez l'homme et de la transmission chez le moustique.

Le stade anneau n'est pas le lieu d'une activité cellulaire très intense, la production d'ARN et de protéines se mettant en place et devenant réellement significative au stade trophozoïte [41]. La synthèse d'ADN débute à ce stade du cycle érythrocytaire (Figure 9), préparant ainsi le stade schizonte où la réplication de l'ADN semble concomitante avec la division cellulaire.

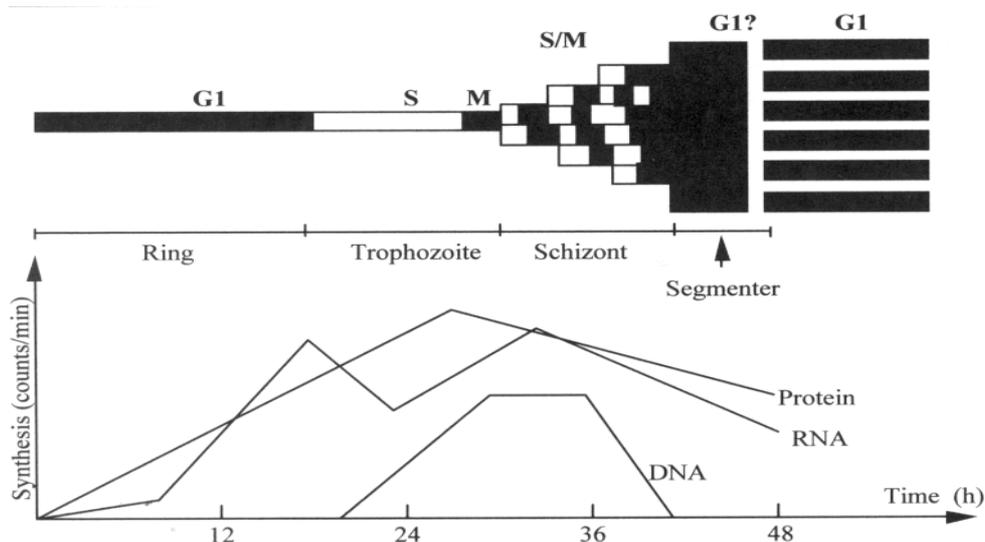


Figure 9. Cycle cellulaire pendant le cycle érythrocytaire de *P. falciparum* d'après [42].

Ce chevauchement contribue à la difficulté à assigner à chaque stade du cycle érythrocytaire une correspondance avec les stades du cycle cellulaire tel qu'il est défini chez les eucaryotes [43, 44]. Le modèle actuel, qui est encore grandement à confirmer expérimentalement, décrit le cycle cellulaire comme suit. Le mérozoïte serait en phase G0 avec une chromatine condensée, puis suit le stade anneau et jeune trophozoïte en phase G1 où la chromatine est décondensée mais la synthèse d'ADN n'a pas encore débutée. Lors de l'entrée en phase S, le trophozoïte commence la réplication de l'ADN qui continue jusqu'à la segmentation des schizontes de concert avec la division nucléaire (phase M) (Figure 9) [42-44].

Pendant la réplication du génome, les chromosomes ne sont pas condensés et la membrane nucléaire persiste : ce processus est décrit comme la cryptomitose. Chaque division est initiée au niveau d'une plaque centriolaire associée à la membrane nucléaire qui ancre le fuseau mitotique en développement [45]. La plaque centriolaire se divise en deux centrioles qui migrent vers deux côtés opposés du noyau résultant à la production d'un fuseau mitotique complet [44]. Une fois la réplication du matériel nucléaire en deux corps nucléaires distincts effectuée, le fuseau se désassemble et une plaque centriolaire est alors associée à chaque corps nucléaire [46]. La mitose dans le schizonte semble asynchrone, impliquant que chaque noyau dans un schizonte ne se divise pas le même nombre de fois [46]. L'indépendance de chaque noyau à l'intérieur d'un même schizonte pose la question de la régulation de chaque phase du cycle cellulaire. Une hypothèse a été avancée présentant chaque plaque centriolaire comme une unité de régulation du cycle cellulaire indépendante de la régulation des autres unités [43]. De manière surprenante, l'utilisation de plusieurs substances connues pour arrêter le cycle cellulaire à certains points de contrôle a montré que *P. falciparum* ne semblait pas posséder de point de contrôle au niveau des phases S et M [42]. Cependant, la régulation du cycle cellulaire reste une évidence lors du cycle érythrocytaire. Les acteurs de cette régulation chez les eucaryotes, les kinases et les cyclines dépendantes des kinases, ont été découverts chez *P. falciparum* [47, 48]. Si leurs rôles respectifs dans la régulation du cycle cellulaire restent largement inconnus, l'activité d'une kinase homologue de cdc2, nommée PfPK5, semble impliquée dans la régulation de la division nucléaire [41, 49].

L'ensemble de ces données permet une description partielle de la régulation du cycle cellulaire chez *P. falciparum* mais souligne le caractère limité des connaissances sur la biologie de ce parasite.

L'échec des entreprises d'éradication de la malaria à grande échelle, l'augmentation des résistances des moustiques aux insecticides et des parasites aux chimiothérapies actuellement

utilisées, soulèvent la question de la capacité des chercheurs à fournir de nouvelles armes contre ce fléau en expansion. Cette recherche semble difficile sans l'apport de nouvelles connaissances sur la régulation des étapes clés de la vie de *P. falciparum*. Le séquençage du génome récemment achevé offre une opportunité sans précédent pour la compréhension et l'étude du parasite causant la mort de plus d'un million de personnes par an.

3. Le génome de *Plasmodium falciparum*

En 1996, après un effort de la communauté scientifique internationale pour rassembler les quinze millions de dollars nécessaires [50], le séquençage et l'annotation du génome du clone 3D7 de *Plasmodium falciparum* ont été lancés. Dès 1998, le chromosome 2 a été entièrement séquencé et publié [51], suivi du chromosome 3 en 1999 [52]. L'ensemble de la séquence des douze chromosomes restants a été rendu disponible dès fin 2002 [53] ouvrant à la communauté scientifique un champ d'investigation à la taille de l'enjeu : trouver de nouveaux moyens afin de combattre le paludisme.

3.1 Composition et organisation du génome

Le génome nucléaire du clone 3D7 de *Plasmodium falciparum* est composé de 22,8 mega bases (Mb) réparties sur quatorze chromosomes dont la taille varie de 0,643 Mb à 5,29 Mb. La composition en nucléotides est largement dominée par les adénines (A) et les thymines (T) qui représentent en moyenne 80,6 % du génome (Tableau 2), cette proportion pouvant atteindre plus de 90 % dans certaines parties du génome, notamment intergéniques. Cette composition nucléotidique a été la source de nombreux problèmes techniques lors du séquençage, notamment à cause de l'instabilité des fragments d'ADN, et lors de l'ordonnement des différentes parties du génome.

Environ 5300 phases ouvertes de lecture ont été identifiées avec une taille moyenne de 4300 paires de bases (pb), laissant donc plus de la moitié du génome non codant. La taille moyenne des phases ouvertes de lecture sans leurs introns (2300 pb) est supérieure à celle constatée chez quatre autres eucaryotes unicellulaires dont le génome a été séquencé (Tableau 2). Cette donnée souligne la dimension exceptionnelle des gènes de *P. falciparum*, d'autant plus qu'aucun élément transposable n'a été identifié dans le génome. Les gènes sont organisés en mono cistrons, espacés par des régions intergéniques dont la taille moyenne est de 1700 bp.

Le génome code pour 45 ARN de transfert (ARNt) représentant l'ensemble des anticodons. Au contraire de la majorité des eucaryotes séquencés à l'heure actuelle, le génome de *P. falciparum* ne semble pas comporter de répétition des gènes codant pour les ARNt. Cette organisation semble cependant conservée chez certains protozoaires parasites comme un autre apicomplexe, *Cryptosporidium parvum*, et la microsporidie *Encephalitozoon cuniculli* (Tableau 2).

Tableau 2. Données générales sur les génomes de cinq eucaryotes unicellulaires d'après [54]

Feature	<i>C. parvum</i>	<i>P. falciparum</i>	<i>S. pombe</i>	<i>S. cerevisiae</i>	<i>E. cuniculi</i>
Size (Mbp)	9.1	22.9	12.5	12.5	2.5
(G+C) content (%)	30	19.4	36	38.3	47
No. of genes	3807	5268	4929	5770	1997
Mean gene length (bp) excluding introns	1795	2283	1426	1424	ND
Gene density (bp per gene)	2382	4338	2528	2088	1256
Percent coding	75.3	52.6	57.5	70.5	90
Genes with introns (%)	5	53.9	43	5	ND
Intergenic regions					
(G+C) content %	23.9	13.6	32.4	35.1	45
Mean length (bp)	566	1694	952	515	129
RNAs					
No. of tRNA genes	45	43	174	299	44
No. of 5S rRNA genes	6	3	30	100-200	3
No. of 5.8S, 18S, and 28S rRNA units	5	7	200-400	100-200	22

De manière similaire, les gènes codant pour les ARN ribosomiaux (ARNr) sont en petit nombre et ne sont pas organisés en répétition en tandem, à l'instar de *C. parvum* (Tableau 2) [54] mais contrairement à de nombreux eucaryotes. En effet, les unités codant pour les ARNr 18S-5,8S-28S sont distribuées sur plusieurs chromosomes. De plus, la séquence de chaque unité semble différer des autres unités. Une classification en deux types d'ARNr a été créée en fonction de leur expression au cours du développement de *P. falciparum*: les ARNr de type A sont exprimés préférentiellement chez l'hôte humain tandis que le type S est exprimé chez le moustique [55]. Le génome de *P. falciparum* contient sept loci codant pour des ARNr, ces unités pouvant être incomplètes ou composées de séquences divergentes du type A ou S. Il est intéressant de noter que les différences de séquence entre les deux types peuvent conduire à de grandes variations fonctionnelles, comme il est apparu lors d'expériences montrant la létalité du remplacement du gène codant pour l'ARNr de la levure *S. cerevisiae* par l'ARNr de type S de *P. falciparum*. La même expérience réalisée avec l'ARNr de type A ne semble provoquer aucun retard de croissance chez la levure *S. cerevisiae* [56].

3.2 La structure des chromosomes

Les chromosomes de *P. falciparum* semblent avoir une organisation caractéristique avec deux régions se détachant nettement.

D'une part, les chromosomes semblent structurés autour de centromères composés par des petites répétitions disposées en tandem sur une longueur de 2 à 3 kilobases (kb) et dont le contenu en bases A et T excède les 97 %.

D'autre part, les régions sub-téломériques semblent aussi avoir une organisation particulière (Figure 10). Elles sont agencées en cinq sous régions conservées composées de répétitions en tandem. Une de ces cinq sous régions contient des répétitions de 21 pb formant l'élément

Rep20. Cet élément a été impliqué dans la co-localisation des télomères de différents chromosomes à la périphérie nucléaire [57]. Ce mécanisme permet sans doute les échanges inter-chromosomiques dans les régions sub-télomériques impliquées dans la variation antigénique chez *P. falciparum*.

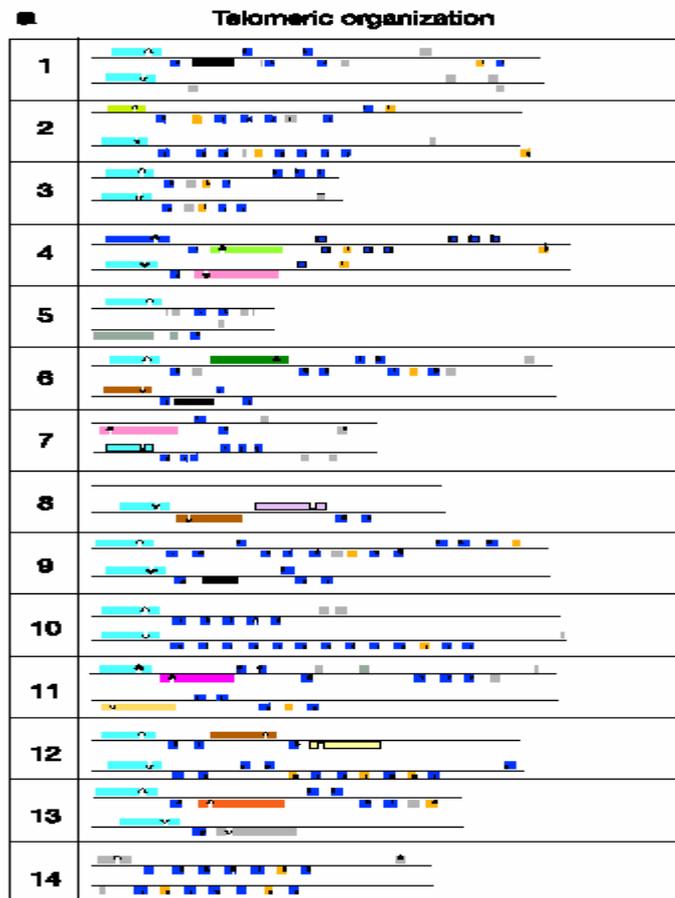


Figure 10. Organisation des régions subtélomériques des chromosomes de *P. falciparum*, d'après [53]. Gènes *rifin* en bleu foncé, *stevor* en orange et *var* en autres couleurs.

Ce phénomène de variation antigénique est essentiellement dû à des familles multigéniques (*var*, *rifin* et *stevor*) situées dans les régions sub-télomériques (Figure 10). L'organisation génomique de ces régions, présentée dans cette figure, démontre la présence d'au moins un gène *var* par télomère.

3.3 Vue générale de la composition protéique

Sur les 5268 protéines prédites dans le génome de *P. falciparum*, plus de 60 % n'ont pas de similarité suffisante dans leur séquence primaire avec des protéines d'autres organismes pour

pouvoir leur assigner une fonction potentielle. Cette proportion est bien plus importante que pour la plupart des génomes déjà séquencés.

A l'opposé, l'annotation de 40 % des protéines a permis leur réunion dans des groupes fonctionnels. En comparaison avec *S. cerevisiae* (dont le génome possède un nombre de gènes équivalent, voir tableau 2), et dans l'état actuel de l'annotation du génome de *Plasmodium*, il apparaît que tous les groupes fonctionnels répertoriés possèdent un plus petit nombre de protéines impliquées, à l'exception notable des groupes «adhésion de cellule à cellule» et «échappement au système immunitaire», dont les surreprésentations peuvent être expliquées par les contraintes engendrées par le cycle de vie de *Plasmodium*. De manière plutôt surprenante compte tenu du cycle complexe de *P. falciparum*, très peu de protéines ont pu être identifiées comme appartenant aux groupes «cycle cellulaire» et «facteur de transcription». En effet, on compte environ 1,3 % de protéines du génome assignées à la régulation de la transcription contre 4 % pour les levures [58] (Tableau 3).

Tableau 3. Pourcentage de protéines assignées à la régulation de la transcription pour 8 génomes d'eucaryotes d'après [58].

Species	Size ^a	Matches	% Genome
<i>P. falciparum</i>	5273	69	1.3
<i>S. cerevisiae</i>	6292	249	4.0
<i>S. pombe</i>	4882	205	4.2
<i>C. elegans</i>	19,099	936	4.9
<i>D. melanogaster</i>	13,707	865	6.3
<i>A. thaliana</i>	27,387	1807	6.6
<i>H. sapiens</i>	29,303	1987	6.8
<i>A. gambiae</i>	15,101	1045	6.9

Size : nombre de protéines dans le génome ; Matches : nombre de protéines assignées à la régulation de la transcription ; % Genome : pourcentage des protéines assignées à la régulation de la transcription par rapport à la totalité des protéines du génome.

Sur l'ensemble des protéines, seulement 14 % ont pu être identifiés comme des enzymes, ce qui est notablement faible en comparaison des génomes de bactéries, dont un tiers des protéines sont considérées comme des enzymes. Malgré cette donnée, de nombreuses voies métaboliques ont pu être partiellement reconstituées permettant d'ouvrir de nouvelles perspectives thérapeutiques, à l'instar de la découverte de protéines canal (Figure 11).

Quelques manques importants sont tout de même à noter:

- d'une part, l'apparente absence de deux sous unités de l'ATP synthétase, une enzyme clef pour la production d'ATP par les mitochondries,

- d'autre part, l'absence de certaines protéines composant les canaux permettant le transit d'ion sodium ou potassium, ceci malgré de nombreuses preuves expérimentales de l'export de ce type de canaux à la membrane des érythrocytes [59].

Par ailleurs, *Plasmodium* possède des familles protéiques qui lui sont uniques. C'est le cas notamment des gènes *var*, dont les domaines caractéristiques sont un particularisme de *Plasmodium* (voir dans la revue [60]). Ces gènes sont impliqués dans la variation antigénique et le phénomène de cytoadhérence, deux mécanismes ayant un rôle prépondérant dans la physiopathologie du paludisme.

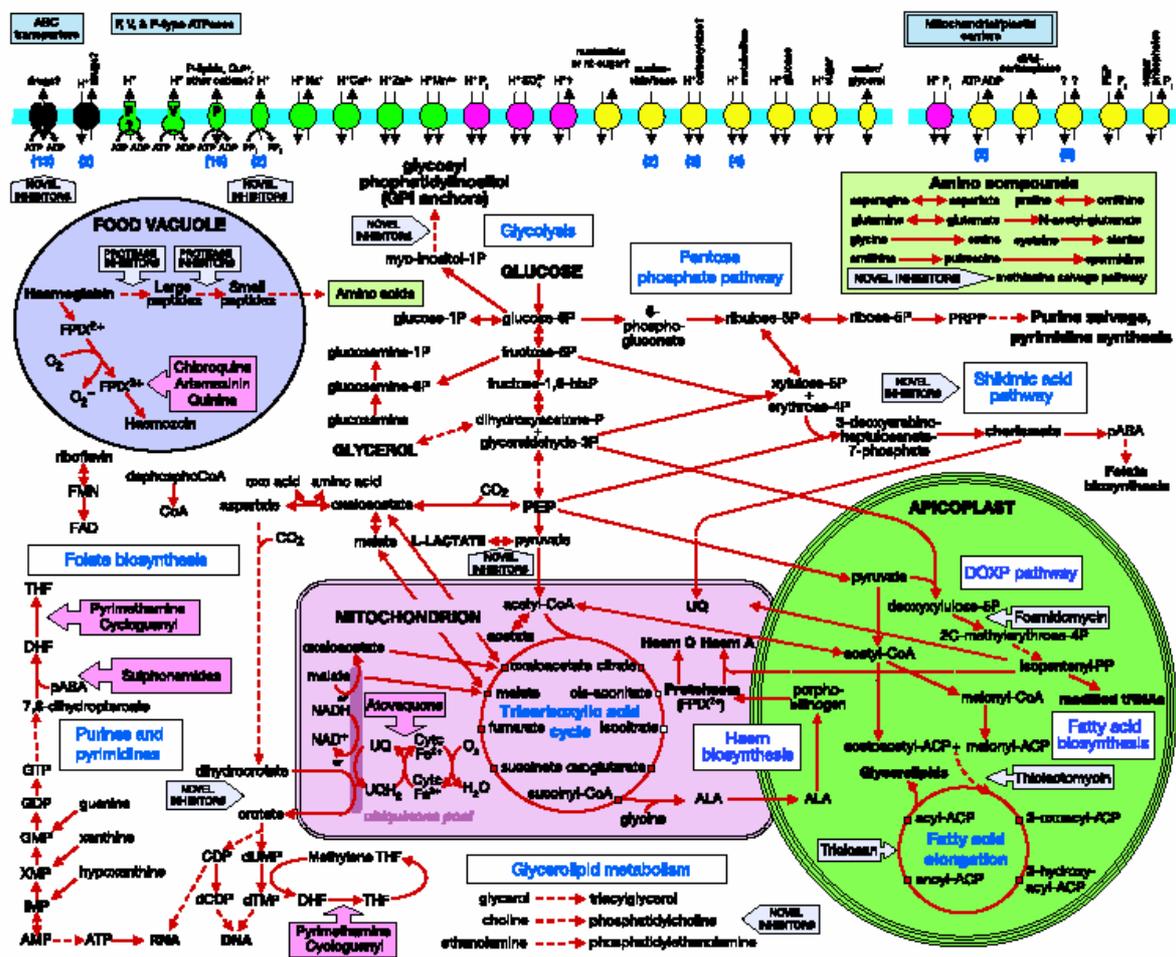


Figure 11. Vue d'ensemble du métabolisme et du transport de *P. falciparum*, d'après [53].

L'ensemble des données exposées ci-dessus et ci-dessous est à relativiser à la lumière de plusieurs faits qui nécessitent d'être rappelés. Les recherches par similarité de séquences sont la base fondamentale de la caractérisation du génome tel qu'il a été décrit dans plusieurs publications [53, 58, 61] et tel que nous avons tenté de le caractériser dans les lignes qui

suivent. Ces techniques ont déjà largement prouvé leurs limites, d'autant plus lorsque l'on compare des organismes éloignés (c'est le cas de la plupart des organismes dont le génome est disponible sur les bases de données en comparaison à *P. falciparum*). Ce fait est d'ailleurs amplifié par la richesse en AT du génome de *P. falciparum* qui impose un biais de codons. Si l'on n'y prenait pas garde, ces écueils mèneraient à apparenter *P. falciparum* à *A. thaliana* [62] au mépris des limites mentionnées. Les algorithmes d'alignement de séquence qui déterminent la similarité entre deux séquences sont d'ailleurs très peu efficaces en dessous de 30 % de similarité [63]. Cela souligne l'importance de la recherche de nouveaux outils pour la comparaison et l'alignement de séquences. L'utilisation de ces programmes a d'ailleurs conduit à la mésestimation du nombre de transporteurs (S. Krishna, communication personnelle) lors de la publication du génome [53]. Par ailleurs, il ne faut pas sous-estimer la capacité de deux séquences primaires différentes à former des structures secondaires-tertiaires proches pouvant ainsi accomplir des fonctions identiques. On ne peut trouver par exemple, par les moyens de recherche classiques, de domaine SH3 chez *P. falciparum* [61]. Un repliement similaire à celui de ce domaine est pourtant retrouvé dans la structure d'une protéine de *P. falciparum* [64], démontrant les limites de l'analyse simple de la séquence en acides aminés. Pour finir, comme nous l'avons vu plus tôt, *P. falciparum* s'est adapté en créant des familles de protéines qui lui sont propres. Ce même schéma aurait pu être répété au cours de l'adaptation de fonctions ancestrales au cycle de vie de *P. falciparum*. L'étude du génome, dans ces limites et avec les outils aujourd'hui disponibles, reste cependant un instrument puissant afin de mieux comprendre les mécanismes biologiques utilisés par *P. falciparum* pour s'adapter à la complexité des environnements rencontrés chez les différents hôtes. Cette capacité implique d'ailleurs une aptitude à réguler finement l'expression de ses gènes. L'analyse du contenu protéique du génome pourrait ainsi révéler un éclairage nouveau sur la régulation génétique chez *P. falciparum*.

3.4 La composition en protéines assignées à la régulation génétique.

L'ensemble des données présentées ci-dessous sur la présence de certains gènes dans le génome de *P. falciparum* provient en plus des quelques séquences que nous avons annotées:

- soit d'articles décrivant le contenu protéique du génome de *P. falciparum* [53, 58, 61],
- soit de l'utilisation de l'outil de recherche des domaines de la banque Pfam [65] dans PlasmoDB [66] et après vérification avec l'outil MotifScan [67],

- soit de recherches grâce à l'outil BLAST de PlasmoDB,
- soit finalement d'articles décrivant la découverte de certaines protéines de *P. falciparum* avant la publication du génome.

L'ensemble des numéros d'accèsion est disponible en Annexe 1.1 et 1.2. Ces résultats sont sous le coup des restrictions que l'on vient de citer et sont soumis aux modifications des bases de données utilisées.

3.4.1 La machinerie générale de transcription

Le génome code pour l'ensemble des douze sous-unités de l'ARN polymérase II connue pour réaliser la transcription des ARN messagers (ARNm). La plus grande sous-unité de l'ARN polymérase II contient, chez un grand nombre d'eucaryotes, une extension de son domaine C-terminal comprenant une répétition en tandem d'un heptapeptide [68]. Cette extension joue un rôle important dans un grand nombre de processus vitaux, comme l'élongation, l'épissage et la régulation de l'expression des gènes [69]. La présence de cette extension semble corrélée avec une capacité de contrôle accrue sur l'ensemble de ces processus et notamment sur la complexité des modes de régulation de l'expression des gènes [70], comme en témoigne son absence chez les trypanosomidés [71]. Dans ce cadre, la présence de cette extension (avec des répétitions imparfaites) dans la grande sous-unité de l'ARN polymérase II de *P. falciparum* [70] est une donnée importante dans la compréhension de la régulation des gènes chez cet apicomplexe.

La liaison de l'ARN polymérase II à l'ADN au niveau du site d'initiation de la transcription est dépendante de la présence de plusieurs complexes protéiques (TFIIB, TFIID, TFIIF et TFIIIE) où la TATA-binding-protein (TBP) assure la liaison à l'ADN au niveau du promoteur proximal. Contrairement aux levures, le génome de *P. falciparum* code pour une TBP [72] et un homologue de la TBP [73]. La présence de plusieurs TBP est une caractéristique des eucaryotes supérieurs, où ces protéines sont associées à une liaison à une séquence alternative à la boîte TATA. Cette diversification suggère une complexité de la régulation de l'activité transcriptionnelle: plusieurs complexes formés avec les deux différentes TBP pourraient avoir des activités spécifiques en fonction de leur capacité à reconnaître différents types de promoteurs basaux [74]. La séquence de ces deux protéines semble avoir conservé les caractéristiques fonctionnelles des TBP en terme de liaison à l'ADN et de dimérisation [72, 73]. Cependant, étant donné l'extrême richesse en A+T des séquences intergéniques, les moyens par lesquels la TBP lie l'ADN au niveau du site d'initiation de la transcription restent encore à découvrir. Par ailleurs, aucun homologue des facteurs associés à la TBP (TAF) n'a

pu être découvert. Au vu de la faible conservation de ces TAF chez les métazoaires, cette donnée n'apparaît pas étonnante [75].

Les protéines permettant l'assemblage autour du promoteur et de la RNA polymérase II du complexe de préinitiation de la transcription et composant les complexes TFIIB, TFIIF et TFIIE ont été aussi identifiées dans le génome. Pour finir, l'ensemble des sous-unités du complexe TFIIH, assurant l'initiation de la transcription par la RNA polymérase II ont été identifiées. Les protéines permettant l'élongation des transcrits sont aussi présentes dans le génome.

La découverte de l'ensemble de ces protéines suggère que *P. falciparum* possède l'ensemble de la machinerie basale conservée chez les eucaryotes, lui permettant d'assurer la synthèse des ARN messagers ainsi que la régulation transcriptionnelle de ses gènes.

3.4.2 Régulateurs de la structure de la chromatine

3.4.2.a L'organisation de la chromatine

L'organisation de l'ADN à l'intérieur de la chromatine et des chromosomes joue un rôle central dans la régulation de processus cellulaires cruciaux. La chromatine est l'organisation structurale basale du matériel génétique chez les eucaryotes. L'unité fondamentale de la chromatine est le nucléosome. Ce dernier est composé de la particule cœur et d'une région de liaison. Le cœur, dont la structure est très conservée parmi les espèces, est composée d'environ 146 pb d'ADN enroulées selon environ 1,7 tours autour d'un octamère protéique comprenant deux exemplaires de chacune des histones H3, H4, H2A et H2B. Chez *P. falciparum*, il est intéressant de noter que l'organisation dans le génome des gènes codant pour ces quatre histones est particulière. En effet, les gènes codant pour les protéines H2A et H3 ont un sens de transcription opposé et partagent le même promoteur sur le chromosome 6. Il en est de même pour les protéines H2B et H4 sur le chromosome 11. Cette structuration pourrait supposer une régulation commune de ces deux couples d'histones.

La compaction de l'ADN dans le nucléosome est le premier niveau de condensation de la chromatine. Le second niveau est composé par une répétition de nucléosomes entre lesquels l'ADN lien est nu ou associé à une histone de liaison comme l'histone H1. En dépit de leur nom, ces protéines ne sont pas apparentées aux histones formant la particule coeur, mais permettent le passage vers un état plus condensé de la chromatine en se liant à l'ADN qui entre ou sort du nucléosome [76]. Si l'on en croit l'analyse actuelle du génome, *P. falciparum* ne posséderait pas d'histone de liaison. Même si ces histones de liaison semblent impliquées dans la régulation transcriptionnelle de certains gènes, elles ne semblent pas responsables de

la régulation globale du niveau de transcription [77]. De plus, elles n'apparaissent pas comme indispensables à la survie des eucaryotes unicellulaires comme les levures [78]. Cet ensemble de données pourrait indiquer que l'apparente absence des histones de liaison chez *P. falciparum* ne perturberait pas fondamentalement la régulation transcriptionnelle.

Par ailleurs, le génome semble aussi coder pour des variants des histones composant la particule cœur du nucléosome. Ces variants ont été notamment impliqués dans la spécification d'états chromatiniens particuliers: dans les régions centromériques, afin de délimiter l'euchromatine de l'hétérochromatine (le variant H3.3 remplaçant l'histone H3 méthylé en lysine 9 dans l'euchromatine) ou dans des processus d'extinction de l'activation de la chromatine (comme pour l'inactivation du chromosome X) [79].

3.4.2.b Modification et remodelage de la chromatine

De nombreuses protéines sont capables d'organiser l'état de compaction de la chromatine, régulant ainsi l'accès des protéines à l'ADN. Deux types d'activités enzymatiques permettent de réaliser cette régulation et semblent présentes chez *P. falciparum*. Elles interviennent lors de la modification de la chromatine d'une part et lors du remodelage de la chromatine d'autre part.

3.4.2.b.1 Complexes de modification de la chromatine

Sous le terme «modification de la chromatine» se cache un nombre important de modifications post-traductionnelles des histones composant la particule cœur du nucléosome. La plupart des modifications ont lieu au niveau de l'extrémité N-terminale des histones et sont généralement l'acétylation et la méthylation, même si d'autres modifications peuvent aussi intervenir chez les eucaryotes comme la phosphorylation, l'ubiquitylation ou l'ADP-ribosylation [80]. Les modifications des histones peuvent affecter directement la structure de la chromatine, en la condensant ou la relâchant, et contrôlant ainsi son accès à d'autres protéines. Elles pourraient aussi être impliquées dans la formation d'une surface d'interaction sur la chromatine pour d'autres protéines responsables de la traduction du «code des histones» en un pattern d'expression transcriptionnelle [81].

Le génome de *P. falciparum* semble coder pour plusieurs histone acétyltransférases notamment de la famille de GCN5 : un homologue de la protéine Elp3p et de la protéine GCN5 déjà étudiée chez *P. falciparum* [82], ou de la famille MYST avec des homologues des protéines MOZ. Chez les eucaryotes, les histones acétyltransférases fonctionnent en larges complexes multiprotéiques [83]. Par exemple, le complexe d'activation transcriptionnelle

SAGA a pour sous-unité catalytique la protéine GCN5 et pour composant la protéine Ada2, dont un homologue a été caractérisé chez *P. falciparum* [84].

L'acétylation des histones est un processus réversible et les histones déacétylases sont connues pour catalyser cette réaction. *P. falciparum* semble posséder au moins trois histones déacétylases conventionnelles putatives en plus d'au moins deux membres de la famille des histones déacétylases SIRTUIN nécessitant l'apport du NAD⁺ comme co-facteur. Un des membres les plus étudiés de cette dernière famille est Sir2, une protéine essentielle à *S. cerevisiae* permettant l'extinction de la transcription de certains loci particuliers comme les télomères ou les unités codant pour les ARNr [85].

Contrairement à l'acétylation, la méthylation des histones semble plus stable, voire irréversible. Cette modification pourrait disparaître passivement par dilution au cours de la réplication, par changement de l'histone méthylée (voir plus haut) ou par clivage protéique de la partie portant le groupe méthyle. Au moins quatre histone-méthylases ont pu être identifiées par l'analyse du génome de *P. falciparum*, toutes appartenant à la famille SET. De manière intéressante, une de ces protéines associée à un domaine MYND (voir la section 3.4.3) connue pour assurer la répression transcriptionnelle de certains gènes chez l'homme [86] est bien présente chez *P. falciparum*.

3.4.2.b.2 Complexes de remodelage de la chromatine

Les protéines appartenant aux complexes de remodelage de la chromatine sont décrites pour leur capacité à modifier la disposition spatiale des nucléosomes sur l'ADN ou à créer un état remodelé du nucléosome, caractérisé par la modification des interactions histone-ADN. Ces modifications ne semblent pas transformer la structure des nucléosomes. Dans ce type de réactions, les protéines responsables du remodelage utilisent l'hydrolyse de l'ATP afin de libérer l'énergie nécessaire au déplacement des nucléosomes sur l'ADN. A la base des complexes de remodelage de la chromatine se trouvent donc les sous-unités capables d'hydrolyser l'ATP: les ATPases. Les différents complexes sont donc classés en fonction de la séquence de l'ATPase. Selon le type de complexe protéique de remodelage, le type d'histone cible et l'environnement de la chromatine, la nouvelle position du nucléosome peut être plus permissive ou au contraire inhibe l'accès d'autres protéines à leurs sites de liaison [87]. Le génome de *P. falciparum* comprend au moins neuf ATPases identifiées par homologie avec les ATPases de type SNF2/SWI2 dont l'activité principale est d'altérer les liaisons histone-ADN. De plus, chez *P. falciparum*, une ATPase de type ISWI a été identifiée et appartient à la famille des protéines de remodelage de la chromatine capables de modifier la disposition spatiale des nucléosomes sur l'ADN. Pour finir, une ATPase homologue de la

protéine CHD a été identifiée chez *P. falciparum*. Ces protéines sont connues pour faire partie de complexes associant remodelage de la chromatine et déacétylation des histones. Elles donnent un exemple de la combinaison de ces deux activités dans la régulation transcriptionnelle.

L'ensemble de ces données suggère que *P. falciparum* possède une machinerie complète capable d'engendrer, comme chez d'autres eucaryotes, les modifications et le remodelage de la chromatine afin de réguler le niveau de transcription de ces gènes.

En plus de cette machinerie basale de remodelage de la chromatine, le génome de *P. falciparum* semble posséder plusieurs gènes codant pour des protéines à domaines PHD ou Bromodomaine connus pour être impliqués dans la régulation du processus de remodelage de la chromatine. Cependant, elles ne possèdent pas d'homologie suffisante à d'autres protéines eucaryotes pour être associées de manière plus formelle à ces phénomènes.

Pour finir, il est intéressant de remarquer que ces mêmes complexes sont aussi impliqués dans des processus tels que la réplication ou la réparation de l'ADN. Certaines de ces protéines ont des formes atypiques, comme c'est le cas pour l'homologue de l'ATPase ISWI, où plusieurs domaines en doigt de zinc de type PHD ont été ajoutés au domaine ATPase. Cette diversification pourrait impliquer une adaptation de ces protéines à la régulation transcriptionnelle chez *Plasmodium*.

3.4.3 Facteurs de transcription spécifiques de séquence

L'analyse du génome de *P. falciparum* semble dévoiler une apparente pauvreté en gène codant pour des protéines pouvant être prédites comme des facteurs de transcription (FT) putatifs et spécifiques de séquence. En effet, si l'on considère la classification de ces protéines dans la base de données TRANSFAC [88], bon nombre de classes n'a pas actuellement de représentant dans le génome. En prenant en compte les quatre superclasses englobant toutes les classes de FT connues à ce jour (les facteurs à domaines basiques, à coordinateurs à doigt de zinc, à Hélice-Tour-Hélice (HTH), à architecture alpha), deux de ces superclasses, (domaines basiques et HTH) ne sont représentées que par une famille de FT. Dans ce cadre, des familles de FT très importantes chez les eucaryotes supérieurs, comme les Homéodomains, domaine Pou, Ets et à crochet-AT, sont orphelines de représentants dans le génome. Cette vision est néanmoins tempérée par le fait que certaines de ces familles ne sont pas représentées chez les levures. C'est notamment le cas pour les familles Ets et domaine Pou [89]. Cependant, l'analyse actuelle du génome permet d'identifier *a minima* des protéines appartenant à une classe de FT dans chaque superclasse. Nous avons donc utilisé les outils de

recherche par homologie cités plus haut, dans le but d'identifier la présence de FT dans génome de *P. falciparum*.

3.4.3.a La superclasse des Domaines Basiques

Une seule classe figure dans le génome: les FT à domaine bZIP. Le nombre de ses possibles représentants est aussi le plus élevé de toutes les familles: près de trente protéines possèdent un domaine de type bZIP chez *P. falciparum*. Ces protéines sont connues chez les eucaryotes pour lier de façon spécifique une séquence d'ADN palindromique sur les promoteurs et pour modifier l'activité transcriptionnelle des gènes correspondants [90]. Cette famille comporte notamment la protéine GCN4. Le domaine bZIP est composé d'un domaine «leucine zipper» et d'un domaine basique chevauchant. Le domaine leucine zipper pourrait permettre une homodimérisation ou hétérodimérisation nécessaire à une liaison stable du complexe protéique à l'ADN. La dimérisation pourrait rendre accessibles les deux régions basiques à l'ADN du petit sillon [91]. Ces protéines sont impliquées dans l'activation ou la répression transcriptionnelle de nombreux gènes essentiels au développement chez les eucaryotes.

3.4.3.b La superclasse des coordinateurs en doigt de Zinc

Cette superclasse est sans doute la mieux représentée dans le génome de *P. falciparum*. Les motifs en doigt de zinc forment de petits domaines repliés indépendamment du reste de la structure protéique autour d'un ion zinc. Ce dernier permettant de stabiliser la formation d'un feuillet bêta contre une hélice alpha. La liaison à l'ADN de ces motifs passe par l'interaction de trois acides aminés (aa) de l'hélice alpha et de trois bases sur l'ADN. La reconnaissance de la séquence d'ADN est aussi dépendante d'interactions secondaires. Les domaines étant capables de fonctionner de manière indépendante, la combinaison de plusieurs domaines en doigt de zinc ayant des affinités pour différents triplets permettrait une reconnaissance spécifique de longues séquences [92]. Cette propriété rend donc virtuellement possible la reconnaissance de nombreuses séquences d'ADN, donnant une flexibilité aux organismes possédant un grand nombre de différents domaines en doigt de zinc. Cette hypothèse pourrait permettre d'expliquer la forte représentation de cette superclasse chez *P. falciparum* et chez nombre d'eucaryotes.

- La classe des domaines C₂H₂ est la plus représentée avec près de vingt protéines chez *P. falciparum*. Ce domaine est formé de trente aa repliés sur un ion Zinc central qui est lui-même chélaté par deux cystéines et deux histidines [93]. Ce domaine en doigt de zinc est très commun chez les eucaryotes, ces protéines ayant été estimées comme comptant pour près de 1% des protéines totales chez les mammifères [92]. La plupart des protéines de cette classe capables de lier de manière spécifique l'ADN possède au minimum deux domaines.

Cependant, il a été montré qu'un seul domaine pouvait être suffisant pour assurer la liaison spécifique à une séquence d'ADN, cette dernière étant formée d'une succession de doublets GA [94]. Le premier rôle probable de ce domaine est la liaison spécifique à l'ADN et la régulation de la transcription. Cependant, certaines protéines, lorsqu'elles ont un grand nombre de répétitions de ces domaines, sont capables de lier l'ARN sous forme double ou simple brin et des hétéroduplexes ADN-ARN par un mécanisme inconnu à ce jour [95]. Chez *P. falciparum*, la plupart des protéines identifiées comportent un à deux motifs. Elles pourraient donc avoir un rôle dans la régulation de la transcription. Cette donnée est d'ailleurs plus intéressante à la lumière de la simplicité des séquences reconnues par ces facteurs de transcription (une succession de G ou de doublet GA comme on l'a vu plus haut) qui pourrait corrélérer avec la richesse en AT des promoteurs de *P. falciparum*. Cependant, un gène semble coder pour une protéine comportant une succession de onze domaines. Cette dernière structure est généralement associée à une capacité de reconnaissance multiple : ADN, ARN et protéines [95].

- La classe possédant un domaine en doigt de zinc B-box a aussi été impliqué dans la régulation transcriptionnelle. Ce domaine est présent dans un grand nombre de protéines, souvent en association avec un domaine en doigt de zinc de type RING et un domaine coil-coil, afin de former un motif tripartite. Les mutations dans ce domaine tripartite sont impliquées dans des maladies génétiques humaines et le cancer [96]. Les domaines B-box sont aussi décrits chez la protéine Constans [97]. Chez les plantes, cette protéine appartient à une famille de facteurs de transcription caractérisés par deux domaines B-box en tandem dont la fonction serait de participer aux interactions protéine-protéine [97]. Constans est impliquée dans la régulation transcriptionnelle de la floraison chez *Arabidopsis thaliana* et participe à l'activation transcriptionnelle de nombreux gènes [98]. Chez *Plasmodium*, au moins une protéine possède 2 motifs B-box en tandem et trois autres protéines un domaine B-box.

- Enfin, chez *P. falciparum*, quatre protéines semblent contenir des domaines en doigt de zinc de type MYND. Ce domaine a été associé à la répression transcriptionnelle de certains gènes chez l'homme [86]. Aucune de ces protéines ne semble contenir un domaine de liaison à l'ADN décrit à ce jour, mais pourrait être impliquée dans la régulation transcriptionnelle en coopération avec d'autres protéines capables de lier l'ADN ou de modifier la chromatine.

- Un autre domaine en doigt de zinc, DHHC, est retrouvé chez *P. falciparum* dans douze protéines. Ce domaine semble très conservé au cours de l'évolution des eucaryotes, mais son éventuelle participation dans la liaison de l'ADN reste putative [99]. Ce motif a été associé aux protéines ayant une activité palmitoyltransferase. Cette dernière n'est pas décrite chez

Plasmodium. On peut donc supposer que certaines de ces protéines sont réellement impliquées dans la régulation transcriptionnelle.

3.4.3.c La superclasse des Hélice-Tour-Hélice

Le motif HTH est un des domaines de liaison à l'ADN les plus répandus au cours de l'évolution, des procaryotes aux eucaryotes [100]. Le motif HTH est généralement composé de vingt aa formant deux hélices alpha connectées par quatre aa. La seconde hélice permet la reconnaissance séquence spécifique de l'ADN du grand sillon. Sur les cinq classes de FT, seule une classe est représentée dans le génome de *P. falciparum* et une seule famille dans cette classe : les facteurs de la classe des domaines à tryptophane et de la famille Myb.

Le domaine Myb a été identifié initialement à partir d'une protéine virale, v-Myb, impliquée dans la transformation cancéreuse de cellules hématopoïétiques [101]. Le domaine de liaison à l'ADN caractérise cette famille de FT et se compose usuellement de trois répétitions de cinquante aa, appelées R1, R2 et R3. Chaque répétition est formée de trois tryptophanes espacés régulièrement par dix-huit à dix-neuf aa. De nombreuses études ont montré que le domaine R1 n'est pas requis pour la liaison à l'ADN [102] mais permettrait une stabilisation de la conformation des domaines R2 et R3 lors de leur interaction avec l'ADN [103]. Les régions R2 et R3 ont été prédites pour former deux domaines HTH avec une connexion non conventionnelle. Elles sont d'ailleurs responsables de la liaison à l'ADN et notamment l'hélice alpha C-terminale des régions R2 et R3 qui seraient responsables de la reconnaissance de la séquence d'ADN [104]. Les régions R2 et R3 forment chez certaines protéines Myb un domaine suffisant pour donner une protéine fonctionnelle [105].

A l'intérieur de chaque région, les tryptophanes seraient responsables du maintien de l'assemblage HTH en formant une structure hydrophobe. Cependant, plusieurs protéines dont la fonction dans la transcription a été démontrée possèdent des répétitions imparfaites où les tryptophanes peuvent être remplacés par des résidus phénylalanines ou tyrosines, de la même manière qu'une insertion de quelques aa dans l'une des répétitions peut être observée [106, 107]. De plus, une cystéine aussi conservée que les résidus tryptophanes semble être importante pour la fonctionnalité de la protéine et son interaction à l'ADN [108]. Les protéines appartenant à cette famille sont connues pour interagir avec une séquence cible caractéristique de l'ADN. La spécificité de séquence semble être relativement bien conservée: une protéine Myb de *Dictyostelium* est capable de lier le même motif que des protéines Myb de vertébré ou de drosophile [109]. Cependant, une protéine Myb de levure Bas1p semble ne pas lier ce consensus MRE [107]. Ces données indiquent que la liaison à l'ADN des facteurs Myb n'est pas restreinte aux MRE, même si cette caractéristique semble conservée chez une

grande partie des protéines Myb. De plus, la liaison à l'ADN pourrait être favorisée par la présence de séquences riches en AT à la périphérie de la séquence consensus [110].

Cette donnée est particulièrement intéressante dans le cadre de l'extrême richesse en AT des séquences promotrices chez *P. falciparum*. Ces séquences riches en AT à la périphérie de la séquence consensus pourraient permettre une plus grande flexibilité de la molécule d'ADN et favoriser la liaison de la protéine Myb à l'ADN. De plus, ces séquences riches en AT semblent être indispensables à la transactivation, soit directement par la protéine Myb soit par le biais d'un partenaire [110]. Parmi les facteurs de transcription connus pour interagir avec les protéines Myb chez les eucaryotes, seule la famille des bZip [111] est représentée chez *P. falciparum*.

Chez *P. falciparum*, au moins neuf protéines possèdent au minimum les deux régions nécessaires à la liaison à l'ADN. Pour la plupart, ces protéines ont des domaines fonctionnels dans lesquels certains tryptophanes sont remplacés par des phénylalanines ou tyrosines.

3.4.3.d La superclasse des facteurs à architecture bêta

Cette grande superclasse de facteurs de transcription regroupe onze classes, dont cinq sont représentées chez *P. falciparum*. Nous ne reparlerons pas ici des facteurs de la classe des TBP (voir section 3.4.1) qui appartient aux facteurs généraux de transcription.

- Les facteurs de la classe des boîtes MADS comptent au moins un représentant dans le génome. Ces protéines sont connues chez les eucaryotes pour lier spécifiquement des séquences d'ADN riches en AT après avoir formé un dimère [112], induisant ainsi une courbure de l'ADN [113]. Leur action semble être dépendante de la présence de partenaires influençant le choix du gène régulé et même le type de régulation [114]. Parmi les nombreux partenaires identifiés chez d'autres eucaryotes, il est intéressant de noter que certains facteurs possédant une boîte MADS interagissent avec des coordinateurs en doigt de zinc de type C2H2, B-box ou bien des facteurs de type bZip; tous sont potentiellement présents dans le génome (voir plus haut) [113].

- Les facteurs de la classe des hétérodimères CCAAT sont représentés dans le génome grâce à l'existence des trois gènes codant pour les trois sous-unités du facteur CBF/NF-Y. Cet hétérodimère est connu pour lier une séquence d'ADN spécifique de motif CCAAT. L'activité de CBF/NF-Y requiert l'hétérodimérisation des sous-unités A et C par le biais d'un motif de repliement de type histone, puis l'interaction du complexe avec la sous-unité B [115]. CBF/NF-Y semble agir en synergie avec d'autres facteurs de transcription, cette coopération étant dépendante de la localisation précise du motif de liaison CCAAT et du motif de liaison du facteur coopérant. CBF/NF-Y semble être capable de stabiliser la liaison à

l'ADN d'autres facteurs. Ceci expliquerait la contrainte de localisation et de distance pour les deux séquences de liaison sur l'ADN [115]. De plus, CBF/NF-Y serait capable d'interagir avec la protéine TBP ainsi qu'avec des protéines dont la fonction est d'acétyler les histones. Ces dernières peuvent d'ailleurs acétyler le facteur CBF/NF-Y, même si la fonction de cette modification reste inconnue [116].

- Les facteurs de la classe des domaines Cold Shock ont été identifiés chez *Echerichia coli* en réponse à une baisse brutale de la température. Ces protéines sont capables de se lier spécifiquement à l'ADN [117] ou à l'ARN chez les eucaryotes, impliquant une capacité à agir aussi bien dans la régulation transcriptionnelle que post-transcriptionnelle [118]. Le mécanisme potentiel de répression de la transcription impliquerait une liaison à l'ADN sur une séquence spécifique qui permettrait la séparation des brins d'ADN et ainsi le déplacement du facteur responsable de l'activation transcriptionnelle [118]. Cette classe de facteur de transcription peut aussi avoir un effet activateur sur certains gènes possédant une séquence spécifique [119]. Il est intéressant de noter que chez certaines plantes ces facteurs ont gardé leur capacité de réponse à une baisse brutale de température [120]. Cette variation en température est une caractéristique du cycle de *Plasmodium* lors du passage de l'hôte vertébré (37 °C) au moustique (25 °C) et pourrait donner une signification fonctionnelle et physiologique aux 2 protéines possédant ce domaine chez *P. falciparum*.

- Les facteurs de la classe des HMG sont une famille de protéines non histones abondantes associées à la chromatine. Ce sont des facteurs architecturaux facilitant différents processus comme la transcription et/ou la recombinaison. Ces facteurs sont classés en trois familles : les protéines de type HMGA contiennent un domaine de liaison à l'ADN de type crochet-AT, les protéines de type HMGB contiennent un domaine boîte HMG et les protéines de type HMGN contiennent un domaine de liaison au nucléosome [121]. Chez *P. falciparum*, seules les protéines de type HMGB ont été trouvées, dont trois représentants de petite taille possédant des domaines de liaison non spécifiques de séquences mais spécifiques de structures. Cette liaison semble se produire préférentiellement sur les régions promotrices riches en AT chez les eucaryotes [122]. Cette famille de protéines est connue pour sa capacité à courber l'ADN. La capacité des protéines HMGB à promouvoir la transcription de certains gènes est rendue possible grâce à son aptitude à interagir avec des facteurs de transcription comme ceux de la famille des bZip [123]. De plus, ces protéines semblent intervenir dans la structure de la chromatine soit en influençant le mouvement et la dynamique des nucléosomes soit en interagissant avec les complexes formant la machinerie de modification de la chromatine [124].

3.4.3.e Les FT non classées

Le domaine «jumonji» a été identifié pour la première fois dans une protéine de souris impliquée dans le développement précoce [125], mais semble conservé dans un grand nombre d'eucaryote. Il est souvent associé à un domaine en doigt de zinc dont l'implication dans la liaison à l'ADN de ces protéines n'est pas encore clairement établie [126]. Une protéine humaine appartenant à cette famille a été décrite comme un FT capable d'induire une réduction de la transcription lors de la liaison à une séquence d'ADN riche en AT [127]. Une protéine de *P. falciparum* présente toutes les caractéristiques de cette famille de FT et possède les mêmes caractéristiques que la protéine humaine Jumonji.

3.4.4 Coactivateurs et corépresseurs transcriptionnels

Lors de la liaison des FT sur l'ADN, le message d'activation ou de répression de la transcription doit être, d'une manière ou d'une autre, transmis au complexe «ARN polymérase II /facteurs généraux de la transcription» afin qu'il soit traduit en terme de régulation de l'expression des gènes. Même si, chez les eucaryotes, les FT spécifiques de séquence semblent être responsables de la spécification de l'expression spatiale et temporelle des gènes, les coactivateurs et corépresseurs pourraient aussi participer à la détermination de la cinétique et l'efficacité de la transcription. Ces protéines pourraient avoir la capacité de recruter et ou de connecter la machinerie générale de transcription aux promoteurs où les FT sont déjà fixés sans eux-mêmes posséder la capacité de se lier à l'ADN [128].

3.4.4.a Le complexe «Mediator»

En effet, un complexe «Mediator» a été identifié après des études génétiques mettant en jeu la suppression d'une mutation de l'extension C-terminale de l'ARN polymérase II chez la levure, puis chez de nombreux eucaryotes supérieurs [129]. Ce complexe protéique ne semble pas forcément indispensable à la transmission de l'information portée par la liaison du FT sur le promoteur, mais pourrait être le vecteur d'une forte activation de la transcription en coordination avec TFIID [130]. Le recrutement du complexe Mediator semble être dépendant de la structure du promoteur, c'est-à-dire de la composition et de la disposition des sites d'interaction des FT et de la machinerie basale de transcription sur la séquence. Ce complexe pourrait ainsi fournir une architecture permettant à l'ARN polymérase II d'effectuer plusieurs tours de transcription [131] (Figure 12).

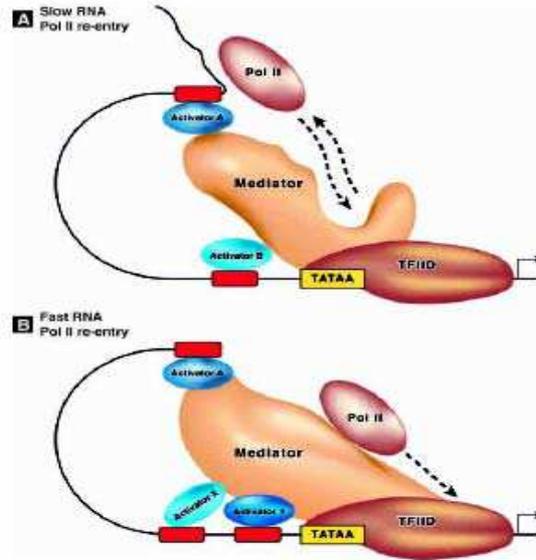


Figure 12. Schéma représentant les modes d'actions possibles du Mediator, d'après [129].

Ce complexe est composé d'une vingtaine de protéines dont tous les membres ne sont pas conservés des levures à l'homme [132]. Le génome de *P. falciparum* semble coder pour au moins trois protéines, conservées des levures à l'homme, contenues dans ce complexe : MED7, MED10 et MED31 [133] mais dont les fonctions exactes ne sont pas définies. Un tel complexe pourrait donc exister chez *P. falciparum* permettant de transmettre l'information provenant de la liaison de ces FT vers la machinerie basale de transcription.

3.4.4.b Autres co-facteurs

Le cofacteur MBF1 a été initialement identifié comme une protéine ne se liant pas à l'ADN mais nécessaire à l'activation transcriptionnelle par l'intermédiaire d'un facteur de transcription à domaine bZip chez le vers à soie *Bombyx mori* [134]. Cette protéine est très conservée au cours de l'évolution des eucaryotes. Sa fonction réside dans sa capacité à réaliser la connexion entre un facteur de transcription spécifique de séquence (notamment à domaine bZip) et la protéine TBP, comme observé chez la levure [135], l'homme, la drosophile, la plante *A. thaliana* [136] et plus récemment chez *C. parvum* [137]. Dans le génome de *P. falciparum*, un homologue de la protéine MBF1 semble être présent et possède une grande similarité de séquence avec les protéines MBF1 des eucaryotes.

Le complexe Ccr4-NOT est un régulateur global de l'expression des gènes, conservé des levures à l'homme. Chez *P. falciparum*, l'ensemble des 9 sous-unités de ce complexe sont présentes dans le génome et semblent être co-exprimées au stade érythrocytaire [138]. Même si ce complexe a été souvent associé à une régulation transcriptionnelle en influençant le positionnement de la TBP sur les boîtes TATA [139], la découverte de son rôle dans la

dégradation de l'ARN pourrait indiquer que l'activité de ce complexe ne réside pas uniquement dans l'initiation de la transcription [140].

En conclusion, l'ensemble de ces données sur la présence, dans le génome de *P. falciparum*, de nombreux acteurs de la régulation transcriptionnelle souligne la complexité des modes de régulation de l'activité transcriptionnelle disponibles au parasite pour répondre à la diversité des environnements qu'il traverse au cours de son cycle de vie. Cette étude indique ensuite que derrière une pauvreté présumée en nombre de FT se cache une diversification qui met à mal la possibilité d'une faible propension de *P. falciparum* à assurer la régulation transcriptionnelle de ces gènes par les acteurs communs de celle-ci chez les eucaryotes. De plus, l'utilisation potentielle de co-facteurs ouvre un champ de régulation supplémentaire et souligne la nécessité de la présence de FT spécifiques de séquence. Pour finir, ces données tendent à prouver que *P. falciparum* est capable d'assurer une régulation transcriptionnelle fine de ses gènes, au niveau de l'initiation de la transcription. Cependant, d'autres types de régulation pourraient être impliqués dans le développement, notamment par de nombreux mécanismes post-transcriptionnels.

3.4.5 Facteurs impliqués dans la régulation post transcriptionnelle

3.4.5.a Au niveau de l'ARNm

Comme nous venons de le voir, *P. falciparum* possède un grand nombre de protéines potentiellement impliquées dans la régulation de l'initiation et du niveau de transcription des gènes. Cependant, le niveau d'ARNm dans la cellule à un moment donné du développement est dépendant de l'équilibre entre la synthèse d'ARNm dans le noyau, sa maturation, son export et le niveau de sa dégradation dans le cytoplasme.

Le génome semble coder pour l'ensemble des éléments formant les complexes permettant la maturation et l'épissage des ARNm, en correspondance avec la relative richesse en intron de ses gènes. Il en est de même pour l'export des ARNm vers le cytoplasme, notamment grâce à la présence dans le génome de protéines impliquées dans la composition des pores nucléaires similaires à ceux identifiés chez les autres eucaryotes [61]. La présence de l'ensemble de ces composants pourrait d'ailleurs expliquer l'apparente multiplication chez *P. falciparum*, en comparaison avec *S. cerevisiae* qui a perdu une grande partie de sa machinerie d'épissage, du domaine RRM (RNA Recognition Motif) de liaison à l'ARN. Potentiellement, *P. falciparum* semble avoir les ressources protéiques nécessaires à la régulation de l'ensemble de ces processus.

En ce qui concerne la dégradation des ARNm dans le cytoplasme, il existe plusieurs voies décrites chez les eucaryotes impliquant la coiffe 5', la queue polyA et les régions 5' et 3' non traduites [141]. Les ARNm de *P. falciparum* ont été décrits comme étant polyadénylés, possédant une coiffe et des extrémités non traduites remarquablement longues, comme en atteste la caractérisation de plusieurs transcrits [142] et la présence dans le génome des enzymes responsables de ces activités. La première voie décrite chez les eucaryotes dépend de la dégradation partielle de la queue polyA. Cette activité est assurée par une polyA ribonucléase [143] dont la présence peut être détectée dans le génome de *P. falciparum*, en plus de l'existence de deux protéines du complexe Ccr4-NOT possédant une activité déadénylase chez la levure [140]. De même, des protéines possédant un domaine PUF capable de lier l'ARN, déjà décrites chez *P. falciparum* [144], ont été impliquées dans la dégradation des transcrits [145]. Ce dommage partiel induit une perte de la coiffe et une initiation de la dégradation de l'ARNm par une activité exoribonucléique 5'→3' ou 3'→5', ces 2 activités pouvant être assurées par des protéines codées dans le génome. Pour finir, quelques protéines de *P. falciparum* semblent comporter un domaine en doigt de zinc de type CCCH qui a été impliqué dans le processus de déadénylation et de dégradation de l'ARN [146].

La seconde voie est indépendante de la queue polyA et liée à une activité endonucléolytique suivie d'une dégradation par une activité exoribonucléique 5'→3' ou 3'→5'. De nombreuses endonucléases et des sites de clivage spécifique ont été étudiés chez les eucaryotes [147] et pourraient être transposés à *P. falciparum* car il pourrait posséder l'ensemble de la machinerie permettant de moduler la stabilité des ARN messagers.

3.4.5.b Au niveau de la traduction

L'initiation de la traduction est une étape limitante de ce processus et donc la cible de la plupart des mécanismes de contrôle. Cette initiation est dépendante en grande partie du recrutement, au niveau de la coiffe des ARNm, du complexe eIF-4F composé d'une protéine de liaison à la coiffe eIF-4E, d'une hélicase à ARN eIF-4A, de son activateur eIF-4B et de la protéine d'architecture eIF-4G [148]. Chez *P. falciparum*, l'ensemble de ces sous unités est présent dans le génome, comme l'ensemble des protéines indispensables à la traduction, à l'exception notable de eIF-4B. La partie 5' non traduite possédant la coiffe ou le site interne d'entrée du ribosome est une cible idéale afin d'empêcher la formation du complexe d'initiation de la traduction. En effet, la liaison d'une protéine dans cette région du transcrit pourrait empêcher l'accès du complexe ternaire d'initiation [148]. De nombreux mécanismes ont été aussi décrits chez les eucaryotes permettant l'inhibition de la traduction d'un transcrit après liaison par une protéine de l'extrémité 3' non traduite [149]. Comme dans le génome de

P. falciparum des protéines possédant des domaines de liaison à l'ARN, comme les domaines RRM, Puf, Sm et IRP (iron regulatory protein) [150], ont été identifiées ce parasite semble être en possession de la machinerie complète intervenant dans différentes étapes de la régulation de la traduction des transcrits. Il faut ajouter à cela un homologue de la kinase HRI [151], dont la fonction semble être, chez les eucaryotes, d'inhiber la traduction par la phosphorylation du facteur eIF2 alpha [152].

3.4.5.c Au niveau post-traductionnel

Il existe plusieurs types de régulation au niveau post traductionnel : d'une part, la régulation de la présence d'une protéine et d'autre part la modification ou l'ajout de molécules permettant le contrôle de son activité.

- La modulation de la présence et de la quantité d'une protéine à un moment donné est fonction de l'équilibre entre sa synthèse et sa dégradation. Plus particulièrement, la régulation du niveau d'expression des FT doit permettre une réponse rapide aux changements d'environnement. La dégradation ciblée des protéines peut être accomplie par le système du protéasome. Chez les eucaryotes, la polyubiquitination des protéines les cible vers la protéolyse dépendante de l'ATP dans le complexe protéasome [153]. Ce mécanisme de régulation de l'abondance des protéines permet aussi un contrôle indirect et rétroactif de la transcription en régulant la présence des FT dans la cellule. En conséquence, nombre de FT chez les eucaryotes sont la cible d'une dégradation rapide par le protéasome, permettant ainsi de réguler l'expression de ces TF et de leurs cibles [154]. L'ensemble des sous unités composant le protéasome, en plus de la sous unité bêta du protéasome 20S déjà connue [155], semble présent. De plus, un grand nombre de protéines pouvant catalyser la conjugaison de l'ubiquitine à une protéine est aussi retrouvé dans le génome de *P. falciparum* et a été déjà caractérisé [156]. Cette donnée n'est pas étonnante au vu de l'effet délétère des inhibiteurs du protéasome sur la croissance de *P. falciparum* [157], soulignant l'importance possible de ce mécanisme de régulation de la dégradation protéique lors du cycle érythrocytaire. Dans cette voie, l'ajout de plusieurs ubiquitines sur une protéine est l'évènement initial clef [158].

- Cependant, l'association d'une unique ubiquitine est un mécanisme permettant le contrôle de l'activité des protéines et donc a des effets sur la régulation de l'expression des gènes, lorsqu'elle touche des FT. Cette transformation post-traductionnelle semble être responsable de modifications de l'état chromatinien lorsqu'elle est effectuée sur les histones cœurs mais aussi de variations de l'expression des gènes lorsqu'elle est effectuée sur certains FT [159]. Il existe des voies apparentées à l'ubiquitination, ajoutant des molécules proches de l'ubiquitine sur les protéines. C'est le cas notamment de la protéine SUMO. La conjugaison de SUMO

aux protéines cibles est dépendante de la présence sur celles-ci d'un motif particulier. La SUMOylation est effectuée par le biais d'un mécanisme proche de l'ubiquitination, mais utilisant une machinerie enzymatique spécifique [160]. L'ensemble de ces activités semble être présentes dans le génome de *P. falciparum*, indiquant que cette voie pourrait être utilisée pour la régulation de l'expression des gènes. En effet, la SUMOylation a été décrite comme responsable de l'activation ou de la répression de l'activité de nombreux FT, parmi lesquels les FT de la famille des Myb, chez les eucaryotes [160]. Cette voie serait d'ailleurs antagoniste de l'action de la mono ubiquitination des FT (Figure 13). Le mécanisme d'action de l'ubiquitination et de la SUMOylation reste encore flou mais pourrait impliquer la capacité de ces deux modifications d'influer sur les interactions protéiques [160].

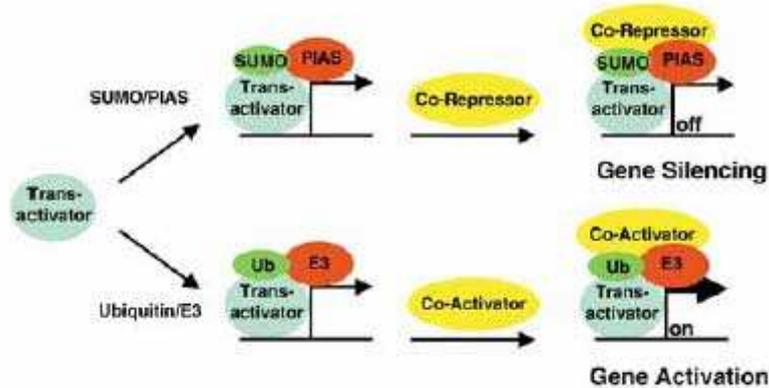


Figure 13. Modèle décrivant l'action des systèmes de l'ubiquitine et de SUMO dans la régulation post-traductionnelle de l'initiation de la transcription d'après [160].

- A l'instar de ces modifications, l'activité des protéines responsables de la régulation de la transcription a été décrite comme sensible, chez les eucaryotes, à d'autres changements post-traductionnels au cours de leur maturation ou en réponse à un stimulus. De ces modifications, la phosphorylation des FT a été la plus étudiée. Le génome de *P. falciparum* code pour environ 60 protéines kinases, un chiffre relativement faible en comparaison à *S. cerevisiae* dont le génome (nombre identique de gènes comparé à *P. falciparum*) présente environ le double de kinases [161]. La phosphorylation et la déphosphorylation des FT auraient au moins 5 effets possibles: contrôler le trafic des FT entre le cytoplasme et le noyau, les cibler en vue d'une dégradation, moduler les interactions protéine-protéine, modifier leur capacité de liaison à l'ADN et modifier la structure de la chromatine [162]. De plus, l'ajout d'un groupement phosphate sur la queue C-terminale de l'ARN polymérase II serait précisément

ordonné pour permettre le recrutement des facteurs généraux de la transcription, l'élongation et la maturation des transcrits [163].

Comme la phosphorylation, l'acétylation est une modification dynamique qui pourrait être impliquée dans la régulation de l'activité des FT. L'ajout et la suppression d'un groupement acétyle sur les FT sont assurés par les protéines déjà décrites dans la section 3.4.2 a. Outre l'activité sur les protéines histone, ces enzymes sont capables de réguler la liaison à l'ADN des FT et notamment de ceux de la famille des HMG box et d'influer sur les interactions protéine-protéine [164].

Cette analyse approfondie du génome de *P. falciparum* et des protéines potentiellement impliquées dans la régulation génétique, dans les limites que nous avons citées qui imposent de valider les mécanismes moléculaires, révèle une grande diversité dans les mécanismes disponibles afin de réguler l'expression de ses gènes. Une grande partie des voies de régulation entre protéines et ADN, connues à ce jour, possède des représentants dans le génome de *P. falciparum*. Cependant, le rôle et l'activité de ces différentes voies lors du cycle de *P. falciparum* sont encore très peu connus.

4. La régulation de l'expression des gènes chez *P. falciparum*

Si l'étude du génome semble indiquer la présence des acteurs de toutes les étapes permettant la régulation de l'expression des gènes, jusqu'ici peu de travaux ont été dédiés à la compréhension des rôles de ceux-ci dans les mécanismes de contrôle qui sous-tendent les variations phénotypiques de *P. falciparum*. Cependant, les études menées jusqu'alors ont valeur d'exemple et attestent de la fonctionnalité de certaines voies de régulation.

La complétion du génome ouvre aussi le champ des investigations possibles et a permis notamment la publication d'études portant sur l'ensemble du génome. Ces dernières donnent la première idée globale de l'expression des gènes au cours du développement de *P. falciparum*.

4.1 Etude à grande échelle de l'expression des gènes

4.1.1 Les études pré-génome

De nombreuses études avaient auparavant souligné le fait que l'expression de certains transcrits de *P. falciparum* semblait spécifique d'un stade du parasite [165]. Cependant l'avènement de nouvelles techniques et notamment des puces à ADN, a permis d'explorer le transcriptome de ce parasite à une autre échelle.

A mesure que la séquence du génome devenait disponible, des expériences tentant de découvrir des différences d'abondance des transcrits entre différents stades du parasite ont été menées. La première étude à grande échelle a utilisé une banque génomique construite grâce à la nucléase du haricot et comprenant 3700 séquences aléatoires [166]. Cette première puce à ADN a servi à la comparaison de l'abondance relative des transcrits entre les stades trophozoïtes et les gamétocytes, réalisant la première description à grande échelle de transcrits spécifiques des stades sexués et asexués. La technique d'analyse quantitative et qualitative en série de l'abondance des transcrits (SAGE) a permis ensuite l'identification de l'expression de 4900 gènes différents durant le cycle érythrocytaire, démontrant qu'une grande partie du génome était exprimée à cette phase du cycle [167]. Puis, une puce à ADN comprenant 944 ADNc a été utilisée afin de mesurer l'abondance des transcrits à 5 temps du cycle érythrocytaire asexué. Cette dernière étude a permis d'acquérir la première image de l'expression temporelle d'une partie des gènes de *P. falciparum* au cours du cycle érythrocytaire. Elle a notamment démontré que certains gènes avaient une expression

inchangée au long des 5 temps utilisés et que 60 % des gènes étudiés semblaient être significativement régulés au cours de ce cycle. De plus, un regroupement des gènes en fonction de leur profil d'expression a permis d'identifier des co-régulations de gènes concourants aux mêmes voies métaboliques [168]. L'ensemble de ces études a démontré la faisabilité d'une approche globale de l'expression des gènes chez *P. falciparum*. Cependant, le faible nombre de gènes ou de stades étudiés, a limité les conclusions qui pouvaient en être tirées.

Après la publication de la séquence de 2 chromosomes de *P. falciparum* l'analyse des gènes du chromosome 2 a été réalisée avec une puce à ADN, utilisant la technologie Affymetrix et couvrant le chromosome 2. L'analyse du transcriptome de 3 points d'une culture synchronisée du cycle érythrocytaire (anneaux, trophozoïtes et schizontes) a ainsi montré que plus d'un tiers des gènes présents sur la puce semblaient significativement régulés. Par ailleurs, cette étude révéla le premier regroupement physique dans le génome de gènes co-exprimés, ces derniers appartenant à la même famille de gènes (SERA) laissant supposer une duplication génique responsable de cette co-régulation [169].

4.1.2 Les études post-génome

La première puce à ADN couvrant la quasi-totalité du génome a été produite en utilisant les données de séquence disponibles afin d'en prédire la structure génique en utilisant un outil de prédiction bioinformatique. Pour chaque gène identifié, un ou plusieurs oligonucléotides de 70 mers ont été choisis en fonction de leur capacité à hybrider une séquence unique dans le génome. Cette puce à ADN a servi à la première étude à échelle génomique comparant deux stades synchronisés du cycle érythrocytaire: les trophozoïtes et les schizontes. Cette étude a démontré l'efficacité des oligonucléotides lors d'une étude globale de l'expression des gènes, et a souligné le nombre important de gènes dont l'expression apparaît strictement contrôlée à un stade donné [170].

Deux études du transcriptome de l'ensemble des gènes de *P. falciparum* ont été menées avec 2 puces à ADN différentes. La première utilisa la puce à ADN où les 5400 gènes du génome sont représentés par environ 7000 oligonucléotides de 70 mers et où chaque heure des 48 h du cycle de *P. falciparum* est analysée [171]. La seconde utilisa la technologie Affymetrix avec 22000 cibles couvrant l'ensemble du génome et en analysant, après amplification des ARN totaux, 9 points du cycle de *P. falciparum*: le stade sporozoïte contenu dans les glandes salivaires, sept points du cycle asexué érythrocytaire des anneaux au mérozoïtes avec deux types de synchronisation différentes et un point du cycle sexué [138]. Ces deux études ont

révéle que plus de 80 % des gènes du parasite sont exprimés au cours du cycle érythrocytaire, laissant peu de gènes dont l'expression est strictement exo-érythrocytaire et soulignant la redondance dans l'utilisation des voies métaboliques tout au long de la vie de *P. falciparum*. Elles ont aussi démontré qu'il existe un programme de co-expression des gènes pour certaines fonctions spécialisées. L'invasion en est un exemple, tout comme les enzymes mitochondriaux codés dans le génome nucléaire démontrant l'importance de la régulation de l'expression dans l'efficacité de la localisation cellulaire [172]. Ces deux publications ont aussi noté qu'il n'existe pas d'organisation chromosomique de gènes co-exprimés, à l'exception de 14 groupes de plus de trois gènes [171]. Dans 8 des 14 groupes, la co-régulation des gènes pourrait s'expliquer par la disposition des gènes, tête-bêche et partageant le même promoteur [171]. Cette donnée implique qu'il n'existe pas chez *P. falciparum* de régulation transcriptionnelle au niveau de larges régions chromosomiques, même si la structure de la chromatine peut être régulée localement. Ceci est par ailleurs confirmé par les données d'un article décrivant l'analyse du contenu protéique des stades du cycle érythrocytaire [173]. Par ailleurs, la comparaison de cette dernière expérience avec les données du transcriptome ont montré une grande similitude entre le stade d'expression du transcrit et celui de la protéine correspondante, ce qui semblerait indiquer une faible participation de la régulation post transcriptionnelle dans la régulation de la présence des protéines lors du cycle érythrocytaire [174]. Cependant, un délai entre l'apparition de l'ARNm et de la protéine a été identifié, notamment pour des gènes impliqués dans la différenciation sexuée [174].

Par ailleurs, l'influence du type de synchronisation effectué ne semble pas être très importante au vu de la corrélation des valeurs d'expression pour chaque point du cycle.

Globalement, les deux séries de données provenant des deux expériences de transcriptome sont semblables même si des différences dans l'évaluation du profil d'expression de certains gènes sont apparues.

Les seules différences entre les deux publications résident dans l'interprétation des résultats. Tout d'abord, alors que Le Roch et collaborateurs divisent leurs données en quinze groupes fonctionnels, Bozdech et collaborateurs ont représenté une partie (la moitié des transcrits) de leurs profils d'expression de manière à ce qu'ils forment une vague continue lors des 48 heures du cycle érythrocytaire (Figure 14). Cette dernière observation implique que les différents stades jusqu'alors décrits par leur morphologie, ne sont pas marqués par des pics d'expression distincts. Cette cascade transcriptionnelle reflète plutôt un développement continu dont l'expression transcriptionnelle n'a pas de lien nets avec les stades décrits

Introduction

morphologiquement. Pour finir, l'évaluation du nombre de gènes variant tout au long du cycle érythrocytaire diffère entre les deux publications, Bozdech et collaborateurs dénombrent plus de 80 % des gènes et Le Roch et coll. n'en identifiant que 45 %. Cette différence est sans doute à attribuer à la variation dans la classification et le regroupement des données.

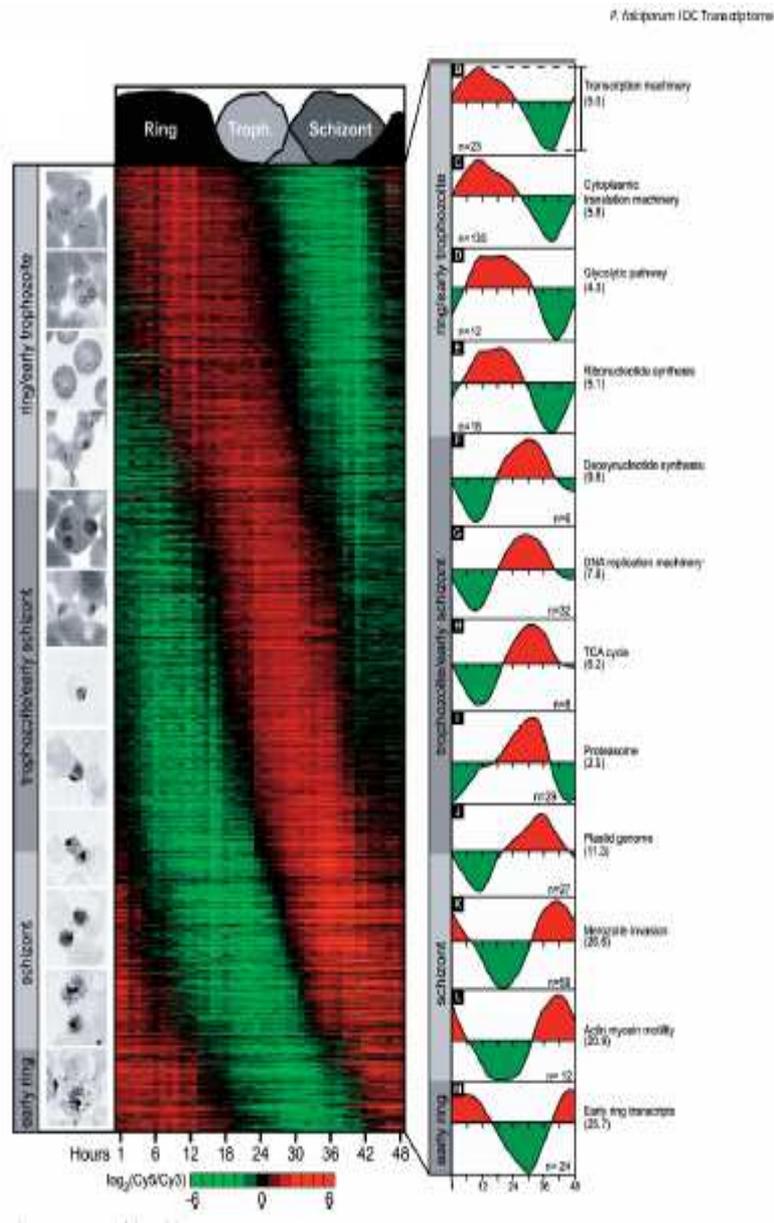


Figure 14. Vue d'ensemble du transcriptome d'après [171].

D'autres expériences ont évalué la réponse des parasites aux variations de l'environnement, au niveau de l'expression relative des transcrits. Ces expériences ont montré que *P. falciparum* est capable d'une réponse limitée face à un stress causé par l'ajout d'une drogue dans le milieu [175]. De même, le changement de la concentration en glucose dans le milieu,

entraîne des variations de l'expression des gènes en réponse à ce stress [176] et notamment une expression différentielle des loci codant pour les ARNr [177]. Ces données révèlent la capacité d'adaptation et de réponse au niveau transcriptionnel de *P. falciparum*.

Conjointement, ces expériences ont donné une vue d'ensemble de l'expression de la totalité des gènes de *P. falciparum*. Elles ont montré que l'expression des gènes chez *P. falciparum* peut être régulée en partie au niveau de l'initiation de la transcription. L'extraordinaire précision de l'orchestration de la régulation transcriptionnelle au cours du cycle érythrocytaire implique l'existence de facteurs dédiés à ce contrôle. Si nous avons vu précédemment que le génome recelait un grand nombre d'acteurs potentiels de cette régulation, les séquences régulatrices contenues dans les promoteurs sont co-responsables de ces modulations de l'expression.

4.2 La régulation transcriptionnelle

La régulation transcriptionnelle nécessite la participation de deux acteurs : les séquences ADN régulatrice agissant en *cis* (séquences promotrices, enhancers...) et les facteurs protéiques agissant en *trans* (facteurs généraux de la transcription, FT...). La régulation fine des interactions entre ces deux acteurs permet le contrôle transcriptionnel. La compréhension du contrôle de ces interactions est un élément clef afin d'appréhender la régulation au niveau transcriptionnel.

4.2.1 La structure des promoteurs

La première description d'un promoteur chez *P. falciparum* a été réalisée en 1992, lors de l'étude d'une partie du chromosome 10. Cette étude a permis de démontrer que la transcription de deux gènes, dont un code pour la protéine GBP130, était monocistronique et unidirectionnelle. Elle a aussi déterminé que le lieu de l'initiation de la transcription du gène *gbp130* et le lieu de terminaison du gène précédant se trouvaient dans la région de 2 kb séparant les deux régions codantes. Ces expériences montrent pour la première fois chez *P. falciparum* que les régions intergéniques recèlent les éléments probablement impliqués dans l'initiation et la terminaison de la transcription [178]. S'en suivent des études des unités de transcription, et notamment la caractérisation du site d'initiation de la transcription, permettant ainsi d'identifier les régions transcrites et non traduites, et les régions dans lesquelles devraient se trouver les éléments permettant l'initiation et l'activation de la

transcription. C'est notamment le cas dans une étude du gène *hsp86*, dans laquelle les auteurs ont identifié dans le promoteur des séquences de régulation qui pourraient être à l'origine de l'initiation ou du maintien de la transcription: un module CCAAT, une séquence riche en GC et enfin des séquences TATA [179]. De plus, une séquence du promoteur du gène *gbp130*, ayant une homologie avec une région « enhancer » de SV40, a été identifiée et possède la capacité de lier spécifiquement des protéines nucléaires [178]. De même, une séquence différente, située en aval du site d'initiation de la transcription du promoteur du gène *kahrp*, semble posséder la même capacité [180]. Ces descriptions de séquences, permettant la liaison spécifique de protéines nucléaires, pourraient impliquer la participation de motifs de régulation et de protéines régulatrices dans le contrôle de la transcription chez *P. falciparum*.

La réelle avancée dans la compréhension des séquences d'ADN responsables de la régulation de la transcription chez *P. falciparum* est l'avènement de la technique de transfection. Cette dernière a permis des études plus complètes des promoteurs et plus particulièrement de séquences indispensables à la régulation transcriptionnelle de certains gènes. Dans la publication référence de cette méthode, les auteurs étudient deux promoteurs afin de valider la technique. Ils observent alors que les régions 5' et 3' flanquantes de ces deux gènes ont une importance capitale sur l'expression du gène rapporteur. L'effet des délétions des parties 5' et 3' flanquantes peut engendrer une instabilité du transcrit et ne permet pas d'identifier des régions régulatrices de l'initiation ou du maintien de la transcription [181]. Une seconde étude permettra ensuite d'identifier une région responsable de l'initiation et une région promotrice en amont du site d'initiation de la transcription, responsable d'une partie de l'expression du rapporteur. Ces expériences constituent la première description du caractère bipartite d'un promoteur de *P. falciparum* [178]. Pour finir, une étude plus complète de trois promoteurs (contrôlant l'expression des gènes de la *Calmoduline* et de la *DHFR-TS* de *P. falciparum* et du gène de la *DHFR-TS* de *P. chabaudi*), grâce à l'utilisation de délétions, a permis de confirmer le caractère bipartite des promoteurs de *P. falciparum*, mais a aussi démontré la présence de séquences possédant la capacité de promouvoir l'expression du gène rapporteur. Ces régions retiennent cette qualité quelque soit leur orientation, ce qui les définit comme les régions «enhancers» qu'on retrouve chez les eucaryotes [182]. Ces dernières expériences ont donné une vision plus fine des promoteurs de *P. falciparum*, les intégrant dans la conception bipartite canonique des promoteurs eucaryotes. Elles ont donné une idée générale de la complexité des séquences permettant la régulation des gènes chez *P. falciparum* (Figure 15).

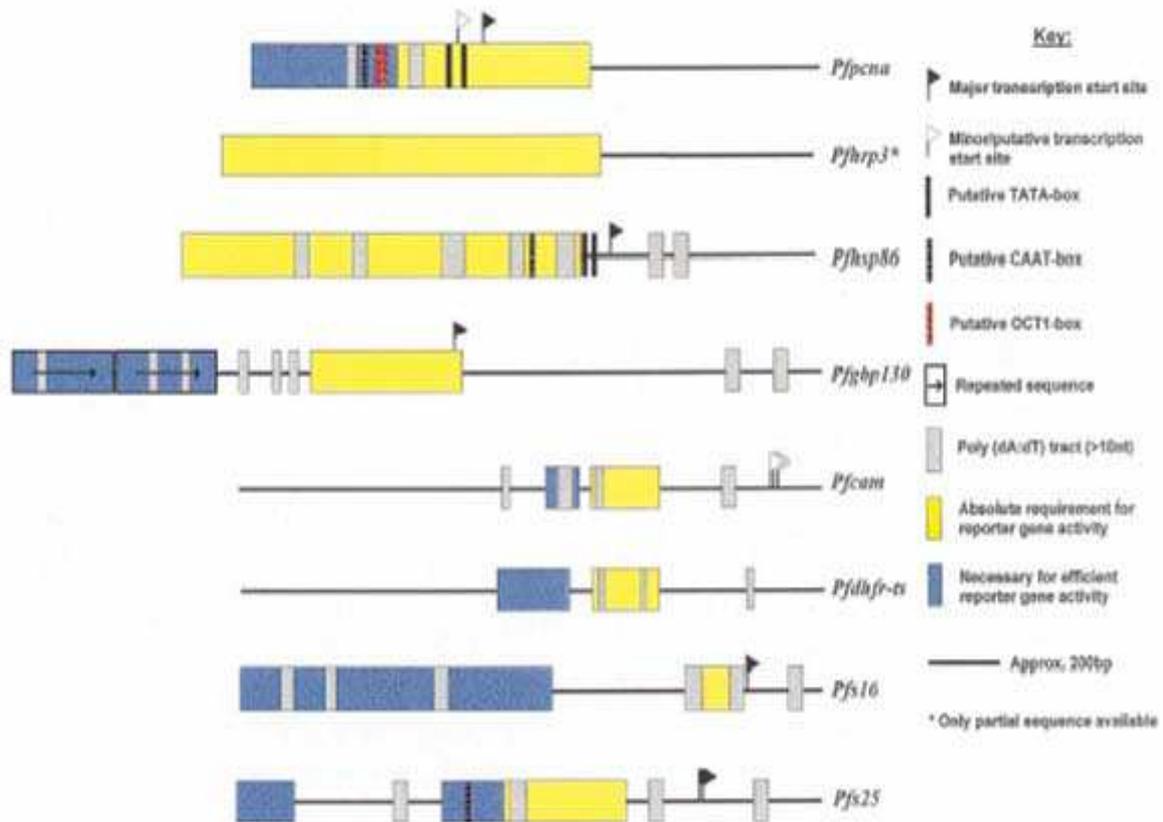


Figure 15. Représentation schématique de promoteurs de *P. falciparum* d'après [165]. Les séquences importantes pour l'expression du gène rapporteur sont indiquées dans la légende.

Elles ont par ailleurs permis de prouver que des éléments de régulation provenant d'une autre espèce de *Plasmodium* (en l'occurrence *P. chabaudi*), sont capables aussi d'activer l'expression d'un gène rapporteur chez *P. falciparum*, permettant d'imaginer qu'il existe une évolution commune des modes de régulation.

4.2.2 Les éléments de régulation

Une analyse utilisant des délétions plus fines des promoteurs a permis la découverte et la caractérisation d'éléments de régulation proprement dits. L'étude du promoteur du gène *PfCDS*, rapporte l'isolement d'une séquence responsable d'une partie de l'activation transcriptionnelle. Cette séquence semble avoir les caractéristiques d'une région « enhancer » car elle semble capable d'activer l'expression du gène rapporteur indépendamment de l'orientation dans laquelle elle est située. De cette région, les auteurs ont identifiés une séquence de 24 pb qui semble responsable de cette activité et qui possède la capacité de lier spécifiquement des facteurs protéiques provenant du noyau [183]. Cette étude met en

évidence un schéma classique de régulation transcriptionnelle où une séquence courte d'ADN est responsable d'une partie de l'activation. Cependant, les auteurs ont mis en évidence la complexité de l'interaction entre les protéines de l'extrait nucléaire et cette séquence, soulignant la possibilité d'un complexe protéique interagissant avec le motif de régulation.

L'étude du promoteur du gène *gbp130* a révélé qu'une régulation plus complexe est possible chez *P. falciparum*. En effet, les auteurs montrent dans ce travail qu'il existe des séquences d'ADN responsables de l'activation et de la répression de l'expression du gène rapporteur. Ces séquences sont d'ailleurs contenues à l'intérieur d'une répétition en tandem qui diffère par quelques nucléotides. La proximité de ces deux répétitions a permis l'identification d'une séquence de quatre nucléotides capable d'engendrer l'activation de l'expression du gène rapporteur. Cette séquence est liée de manière spécifique par des facteurs protéiques contenus dans les extraits nucléaires [184]. Le fait que l'orientation et la position de cette séquence soient déterminantes pour son activité souligne l'importance des facteurs qui pourraient potentiellement agir en coopération et qui auraient un site de liaison sur ce même promoteur. Des facteurs se liant spécifiquement à d'autres séquences du promoteur ont d'ailleurs été identifiés [178, 184], et pourraient être co-responsables de l'activation de la transcription du gène *gbp130*. L'étude de ce promoteur souligne le fait que la régulation de la transcription chez *P. falciparum* est un phénomène coopératif nécessitant la présence de plusieurs acteurs comme décrit chez les eucaryotes.

L'étude du promoteur du gène *hsp86* est la première ayant mêlé analyse bioinformatique et analyse biologique par transfection. Cette combinaison a permis de déterminer que deux séquences en amont du site d'initiation de la transcription sont responsables de l'activation transcriptionnelle du gène rapporteur. L'analyse bioinformatique préalable des promoteurs de ces gènes sur l'ensemble des espèces de *Plasmodium* dont le génome a été séquencé, démontre la conservation d'une séquence palindromique riche en guanine et appelée boîte-G. Lors de l'analyse biologique par transfection, une autre séquence en amont du site d'initiation de la transcription est identifiée comme responsable d'une partie de l'expression du rapporteur, à l'instar de la boîte-G identifiée par bioinformatique. Si ce motif de régulation boîte-G semble conservé et actif, il ne semble pas capable de lier de manière spécifique des facteurs protéiques contenus dans l'extrait nucléaire dans les conditions utilisées [185]. Il semble donc que ce module est responsable de la liaison faible voire non spécifique de facteur, favorisant peut être une conformation de l'ADN permettant la liaison de protéines. Ces dernières pourraient participer à la liaison séquence spécifique d'autres facteurs et ainsi concourir à l'activation de l'expression de ce gène. Ces expériences complètent la panoplie de

dispositifs impliqués dans la régulation des gènes de *Plasmodium* et montrent l'importance possible de facteurs régulant localement la structure de la chromatine.

Une autre étude a porté sur l'analyse de deux promoteurs responsables de l'expression de deux gènes impliqués dans la phase sexuée du parasite. Le premier *pfs16* est exprimé dans les jeunes anneaux qui vont devenir gamétocytes [186] et le second, *pfs25*, est exprimé dans les gamètes et lors du passage dans le moustique [187]. Les expériences présentées dans cet article démontrent que ces deux promoteurs sont actifs uniquement au moment de l'observation de la présence de la protéine (Pfs16 ou Pfs25), aux stades précoces de la gamétocytogénèse pour Pfs16 et chez le moustique pour Pfs25. Cette observation implique une régulation transcriptionnelle de la présence de ces deux protéines. De plus, des éléments de régulation semblent être à l'origine de l'activation de l'expression du gène rapporteur. Dans le promoteur du gène *pfs25*, un élément de régulation semble posséder la capacité de lier des facteurs spécifiquement exprimés dans les gamètes et pas dans les formes asexuées [188]. Cette dernière observation tend à démontrer qu'une partie de la régulation transcriptionnelle de ce gène est dépendante de la disponibilité des facteurs en question.

En complément, il semble que certains promoteurs aient la capacité d'être bidirectionnels. En effet, chez *P. falciparum*, les gènes codant pour les histones cœur sont organisés dans le génome de telle manière que chaque couple d'histones partage le même promoteur et leurs transcrits synthétisés dans un sens opposé. Cette organisation suggère que ce promoteur est bidirectionnel, permettant ainsi la co-expression de ces deux gènes. Cette particularité a été démontrée pour le promoteur des gènes *ef-1 alpha* et *beta* chez *P. berghei* [189], même si cette structure ne semble pas conservée pour ces deux gènes chez *P. falciparum*.

Pour finir, un élément participant au contrôle de l'expression mais situé en dehors d'un promoteur a été identifié dans une étude de la régulation de l'expression des gènes *var* [190]. Cette région, localisée dans l'intron du gène *var7b*, possède la capacité d'induire la répression de l'expression du gène rapporteur, quelle que soit son orientation, lorsqu'elle est associée au promoteur de ce même gène. Cette aptitude est dépendante de cette association car cette région n'est plus capable de réprimer l'expression du gène rapporteur contrôlée par un promoteur différent [191]. La présence des éléments de régulation présents dans le promoteur et dans l'intron serait donc nécessaire à la régulation transcriptionnelle de l'expression ce gène.

L'étude fine de cette poignée de promoteurs met en évidence les nombreuses voies utilisées par *P. falciparum* pour assurer la régulation transcriptionnelle. Néanmoins, ces quelques données soulignent la complexité des phénomènes contrôlant la régulation du niveau de

transcription pour chaque gène. Elles indiquent aussi la nécessité de l'intervention coopérative de plusieurs facteurs afin de permettre le contrôle transcriptionnel de l'expression des gènes. Enfin, elles permettent de faire entrer *P. falciparum* dans le schéma classique du promoteur bipartite, des éléments de régulation positifs ou négatifs, de la coopération et la diversité des facteurs impliqués dans ces régulations. Elles tendent à montrer qu'au delà de ce cadre commun, *P. falciparum* a évolué vers une différenciation des éléments de régulation. Cette remarque est tempérée par le faible nombre d'élément de régulation connu jusqu'alors. Par ailleurs, les contingences techniques dues à la structure des épisomes et à leur maintien impliquent que d'autres phénomènes pourraient être impliqués dans ces régulations [192, 193].

4.2.3 Les particularités

Enfin, si l'ensemble des promoteurs testés chez *Plasmodium* semble être capable de diriger l'expression du gène rapporteur dans n'importe quelle espèce de *Plasmodium*, la fonctionnalité de ces promoteurs dans des cellules eucaryotes n'a pu être démontrée. De même des éléments de régulation eucaryotes n'ont pas prouvé leur efficacité dans la régulation de l'expression des gènes chez *P. falciparum* [165]. Même si nous avons vu que *Plasmodium* pourrait avoir retenu l'essentiel de la structure canonique eucaryote de régulation transcriptionnelle, ces données impliquent une diversification chez *P. falciparum* des éléments de régulation, des facteurs se liant à ces éléments ou bien du mode d'initiation de la transcription.

4.2.3.a site d'initiation de la transcription

L'extrême richesse en base A et T des promoteurs (>90 %) pose la question de l'initiation de la transcription autour de la séquence consensus canonique eucaryote TATA, même si d'autres sites d'initiation existent chez les eucaryotes. En effet, chez *P. falciparum* et dans la totalité des promoteurs étudiés, le site d'initiation de la transcription ne correspondait pas avec la présence d'un consensus de boîte TATA. Cette absence, même relative car provenant d'un petit groupe de promoteurs, est corroborée par l'étude de plusieurs promoteurs de l'apicomplexe *Toxoplasma gondii* [194-196]. Cependant, chez la drosophile, près de 60 % des gènes n'ont pas de boîte TATA [197]. L'initiation de la transcription est alors réalisée au niveau d'une séquence consensus appelée Initiator (Inr) ou grâce à un élément proche mais situé en aval du site d'initiation (downstream promoter element: DPE) [197]. De plus, un élément différent de ces 2 derniers et responsable de l'initiation de la transcription du gène *SAG1* chez *T. gondii* a été identifié [195]. La fonctionnalité de l'ensemble de ces éléments n'a

pu être prouvée pour les promoteurs de *P. falciparum* étudiés, sauf dans le cas de l'Inr dont une version dégénérée, identique du consensus de l'Inr de *Trichomonas vaginalis*, a montré son importance dans l'activation de l'expression, mais pas explicitement dans l'initiation de la transcription, conduite par le promoteur du gène *var7b* chez *P. falciparum* [191]. Cependant, une séquence de régulation proche mais en aval du site d'initiation de la transcription du gène *pgs28* chez *P. gallinaceum* a été caractérisée comme responsable d'une partie de l'activité transcriptionnelle du promoteur [198]. Ce motif de régulation pourrait être apparenté fonctionnellement au DPE, cependant les expériences menées ne peuvent exclure la participation de ce motif dans la régulation de la stabilité ou de la traduction de l'ARNm. La caractérisation des éléments permettant l'initiation de la transcription chez *Plasmodium* bute pour l'instant sur le faible nombre de promoteurs étudiés. Il est d'ailleurs possible que *Plasmodium* ait évolué ses propres séquences de régulation, ou les modalités de la liaison de la TBP, rendant l'analyse du site d'initiation de la transcription par les standards eucaryotes inutile. Cependant, il semble que le choix d'un site initiation ne soit pas un phénomène aléatoire. En effet, le gène *pfb7* présente deux tailles de transcrit différentes dans les formes sexuées ou asexuées de *P. falciparum* et *P. berghei*. Cette variation dans la longueur de la partie 5' non traduite s'explique par un site d'initiation de la transcription alternatif suivant le stade de développement du parasite [199]. L'utilisation de deux sites d'initiation de la transcription différents lors du cycle de *P. falciparum* souligne que ce phénomène est régulé de manière stricte ce qui implique la présence de signaux spécifiques dont la découverte se fait attendre.

4.2.3.b Les régions homopolymériques

La composition en base inhabituelle du génome de *P. falciparum*, pourrait impliquer une éventuelle fonction dans la régulation transcriptionnelle. L'élément déterminant permettant cette interrogation réside dans la découverte d'un biais de composition dans les séquences intergéniques de *P. falciparum*. En effet, la comparaison de l'enchaînement des bases dans le génome de *P. falciparum* par rapport à une séquence aléatoire de même composition, montre qu'il existe une forte surreprésentation des régions homopolymériques. Celle-ci s'explique d'ailleurs par une plus grande occurrence des régions homopolymériques en A et/ou T, dont la taille excède 10 pb, dans les introns et séquences intergéniques, au contraire des séquences codantes [200]. Ce phénomène se retrouve chez certains eucaryotes, notamment chez l'homme et *S. cerevisiae* [200]. Les régions homopolymériques (dA:dT) sont connues pour faciliter la courbure de l'ADN [201]. Cette courbure peut alors être impliquée dans la régulation de l'activité transcriptionnelle en activant ou en inhibant les interactions ADN-protéines, en

favorisant les interactions protéines-protéines ou en influençant la position des nucléosomes [201]. L'influence de ces régions sur l'activité des promoteurs n'a pas jusqu'alors fait l'objet d'études précises. Cependant, il a été montré que ces régions pouvaient agir sur la courbure du promoteur d'une DNA polymérase chez *P. falciparum* [202]. De plus, la délétion de parties du promoteur possédant des régions homopolymériques semble avoir un effet positif sur l'expression du gène rapporteur [203]. L'ensemble de ces données pourrait indiquer que ces régions homopolymériques auraient une influence sur l'activité transcriptionnelle chez *P. falciparum* comme décrit chez *S. cerevisiae* [204].

4.2.4 Les facteurs de transcription

Comme nous l'avons vu au cours de l'exploration du génome, *P. falciparum* semble posséder un grand nombre d'acteurs de la régulation de l'initiation de la transcription. Cependant, même si une faible proportion des facteurs généraux de la transcription a été identifiée et leur cinétique d'expression déterminée, aucune des études menées n'a abouti à une caractérisation de la fonction de ces facteurs. Très récemment, une publication a utilisé un système d'identification de facteurs de transcription avec des extraits nucléaires de *P. falciparum* [205]. Ces expériences qui permettent de mettre en évidence des protéines capables de se lier à des motifs de régulation sont sensés identifier les classes de FT présents dans un extrait protéique. Cependant, la spécificité de ces expériences, prévues à l'origine pour être utilisées chez les mammifères, peut être largement mise en doute. Les auteurs montrent d'ailleurs que la spécificité de seuls 2 des 15 signaux positifs a pu être vérifiée par des expériences de retard sur gel [205]. Il est tout même important de noter que ces deux signaux pourraient révéler la présence dans les extraits nucléaires de FT de la famille des Myb et de la famille des boîtes MADS. Ces 2 familles de protéines ont été retrouvées lors de l'analyse du génome de *P. falciparum*. Il est donc très important d'étudier la fonction moléculaire de ces facteurs afin de pouvoir mieux comprendre la régulation de l'expression des gènes chez *P. falciparum*.

4.3 Régulation post-transcriptionnelle

4.3.1 Régulation de l'épissage

Le contrôle de l'épissage et l'épissage alternatif ont une fonction très importante chez les eucaryotes, et notamment chez les métazoaires où ils permettent de créer une diversité de transcrits à partir d'un même locus génétique, selon l'état de développement de la cellule

[206]. Ce mécanisme permet de créer de la diversité à partir d'un petit nombre de locus, ce qui paraît être le cas pour *P. falciparum* au regard de la complexité de son cycle et du nombre relativement faible de ses gènes. Si les protéines permettant l'épissage n'ont pas été caractérisées chez *Plasmodium*, les signaux indiquant les bornes des introns semblent être les mêmes que chez les eucaryotes [207]. Jusqu'alors le phénomène d'épissage alternatif n'a été identifié que dans un nombre limité de gènes chez *Plasmodium* et la complexité de celui-ci semble limitée [199, 208-210]. De ces phénomènes d'épissage alternatif, celui intervenant dans la création de variants de l'aldénylate cyclase est particulièrement intéressant. En effet, les différentes isoformes produites par ce mécanisme semblent avoir des fonctions différentes en fonction du nombre de domaines transmembranaires qu'ils possèdent [209].

4.3.2 Régulation de la stabilité des ARNm et de la traduction

La longueur exceptionnelle des parties transcrites et non traduites des messagers de *P. falciparum* suppose que leur conservation au cours de l'évolution doit avoir une cause biologique. Par exemple, la taille des parties non traduites du transcrite, produit du gène *gbp130*, est une des plus grandes connue à ce jour [178]. Ces parties de l'ARNm sont connues pour être responsables de la stabilité des ARN messagers et participer à l'efficacité de leur traduction. Lors des expériences de transfection, il a été montré que ces régions non traduites interviennent dans l'expression du gène rapporteur mais l'analyse de leurs rôles respectifs dans la stabilité ou la traduction n'a pas encore été réalisée [182-185, 203]. Dans ces expériences, aucune séquence particulière n'a été identifiée comme impliquée dans l'un ou l'autre de ces mécanismes.

La régulation de la traduction n'a pas pour l'instant fait l'objet de nombreuses études. La première constatation d'une potentielle régulation de la traduction a été faite suite à la découverte de l'expression du transcrite du gène *pbs21* dans les gamétocytes et de l'expression de la protéine uniquement au stade ookinète [211]. Chez les eucaryotes, ce mode de régulation temporelle de la traduction d'un transcrite est d'ailleurs largement conservé et en particulier dans le développement embryonnaire [212]. Ce gène et ses homologues chez d'autres espèces de *Plasmodium* ont fait l'objet de nombreuses études cherchant à comprendre le rôle de sa région 3' UTR dans la régulation de sa traduction. Il a notamment été montré que la délétion d'une partie des régions 5' et 3' flanquantes du gène *pbs21* permettait son expression lors du cycle érythrocytaire asexué, soulignant le rôle potentiel de ces régions dans la régulation de l'expression temporelle du transcrite [213]. Deux régions nécessaires à l'expression d'un gène rapporteur ont été alors identifiées dans la partie 3' UTR

du transcrit du gène *pgs28* [214]. Ces deux régions, une séquence riche en uridine et un site de polyadénylation, ont été étudiées afin d'identifier leurs actions respectives dans l'expression du gène *pgs28*. Ces études ont montré leur importance dans la polyadénylation du transcrit *pgs28*, mettant en exergue la ressemblance de ces régions avec les signaux de polyadénylation des levures ou des plantes [215]. Cependant, ces mêmes sites ne semblent pas être fonctionnels lors de l'expression d'un gène rapporteur dans des cellules humaines [216]. L'ensemble de ces expériences démontre qu'il existe bien une régulation de la traduction chez *P. falciparum*, mais les acteurs et les mécanismes restent largement encore inconnus à ce jour. Enfin, quelques gènes transcrits mais non traduits sont aussi observés chez *Plasmodium*, mais pourraient dériver d'une duplication génétique car ils contiennent un décalage du cadre de lecture [217, 218]. Ces études démontrent la fonctionnalité des codons stop chez *Plasmodium* et dans la régulation de la traduction même si *P. falciparum* semble être capable dans certains cas de passer outre cette information [219].

La régulation de l'expression du gène *dhfr-ts* de *P. falciparum* semble être dépendante de la liaison spécifique de cette enzyme sur son ARNm bloquant ainsi sa traduction [220]. Cette rétro-inhibition de la traduction peut être levée par des inhibiteurs spécifiques de cette enzyme, permettant ainsi l'augmentation de la quantité de protéine [221].

Ces expériences montrent, une fois de plus, que *Plasmodium* a certainement retenu bon nombre des modes de régulation de l'expression des gènes similaires aux eucaryotes, mais a adapté ces mécanismes dans une telle mesure qu'ils requièrent une machinerie spécifique des espèces *Plasmodium*.

4.3.3 ARN anti-sens

Chez les eucaryotes, la présence d'ARN anti-sens naturels a été décrite à maintes reprises. Un grand nombre de ces transcrits sont par exemple exprimés chez l'homme, avec plus de 8 % du génome produisant un ARN anti-sens [222]. La présence d'ARN anti-sens naturels a été aussi démontrée chez *T. brucei* [223], *G. lamblia* [224] et *D. discoideum* [225]. L'implication de ces ARN anti-sens endogènes dans la régulation de l'expression des gènes est maintenant établi chez bon nombre d'eucaryotes [226]. Cependant, ces messagers peuvent parfois être traduits en protéines, comme pour la protéine ASP du virus VIH-1 [227]. De manière générale, la présence d'ARN anti-sens naturels concomitante à celle du messenger sens conduit à la dégradation de ce dernier, mais peut conduire dans certains cas à sa stabilisation [226]. La formation de duplexes entre ARN sens et anti-sens pourrait causer la dégradation du complexe, l'inhibition de sa maturation et de son transport, ou l'inhibition de la traduction du

messenger sens. De plus, la présence de molécules d'ARN anti-sens en excès peut entraîner la séquestration des protéines responsables de la stabilisation du messenger sens et ainsi conduire à sa dégradation ou l'inverse si ces protéines sont responsables de sa déstabilisation. La transcription à partir du brin anti-sens pourrait aussi empêcher l'initiation ou l'élongation de la transcription du brin sens. Chez l'homme, la dégradation des doubles brins d'ARN peut être causée soit par un mécanisme non spécifique [228] ou spécifique tel que l'interférence par ARN [229].

La régulation de l'expression de certains gènes par le biais de messagers anti-sens est devenue apparente lors de la découverte d'une relation inverse entre les niveaux de messagers sens et anti-sens par exemple pour le gène *eb4-psv* de *D. discoideum* [225]. Ce même type d'accumulation inverse de transcrit sens et anti-sens a été décrite pour le gène *Pfmsp2* de *P. falciparum* [230]. De plus, ce ratio d'expression inverse entre les messagers sens et les anti-sens, est en coordination avec le profil d'expression du transcrit et de la protéine *Pfmsp2*. Dans ce cas précis, même s'il est tentant d'attribuer un rôle de régulateur à l'ARN anti-sens, le fait que celui-ci soit produit à partir du site d'initiation de la transcription d'un gène adjacent et orienté à l'opposé relativise cette possibilité. En effet, la cinétique d'expression du gène adjacent est exactement inverse à celle du gène *mfp2*. De plus, une expérience de transfection plaçant un gène *mfp2* (différent du gène endogène) dans un autre contexte chromosomique, mais sous le contrôle de son promoteur, n'altère aucunement son profil d'expression, suggérant que les régions 5' flanquantes sont suffisantes à la bonne régulation de son expression [231]. Cependant, des expériences de SAGE sur le cycle érythrocytaire de *P. falciparum* ont montré la présence d'une grande quantité d'ARN anti-sens naturels [167, 232]. Ces expériences ont montré que près de 12 % des loci étudiés produisaient des tags anti-sens et 5 % des tags dans les deux sens [232]. Dans ce dernier cas, les auteurs ont constaté que la relation entre la quantité de tags sens et anti-sens était inverse. De plus, près de 70 % des loci ayant une transcription du côté anti-sens ne génèrent pas de transcrit sens [232]. Ces observations tendent à démontrer qu'il existe une régulation de l'expression transcriptionnelle de ces gènes par la présence de l'ARN anti-sens, d'autant plus que 90 % de ces loci n'ont pas de gènes adjacents d'où pourraient être initié la transcription [232]. La machinerie qui serait impliquée dans la régulation de l'expression des gènes et donc capable de prendre en charge ces anti-sens et les double brins résultant de la présence concomitante des messagers sens reste pour l'instant inconnue. Cependant, la capacité d'oligonucléotides anti-sens [233, 234] ainsi que de double brin d'ARN [235, 236] à induire une baisse du transcrit voire de la

protéine correspondante semble être des preuves raisonnables de la présence d'une telle machinerie.

L'ensemble de ces données reflète l'état des connaissances sur la régulation de l'expression des gènes chez *P. falciparum*: de nombreuses évidences d'une proximité avec les autres eucaryotes accompagnées d'une grande méconnaissance des processus fins et des acteurs responsables de ceux-ci. Il apparaît aussi que *P. falciparum* ait évolué vers une adaptation de ses acteurs de la régulation de la transcription, rendant ainsi leur analyse moins aisée.

5. Présentation des travaux de thèse

Lors du cycle complexe décrit par *Plasmodium falciparum*, chez le moustique et chez l'homme, la phase érythrocytaire est cruciale car responsable de la pathologie et de la transmission du parasite *via* le moustique. La régulation des événements de prolifération et de différenciation implique un contrôle fin des activités transcriptionnelles de chaque gène impliqué dans ces processus même si la régulation post-transcriptionnelle pourrait jouer un rôle non négligeable.

Les travaux que j'ai réalisés durant ma thèse sont le fruit du constat contradictoire opposant l'évidence du rôle crucial de la régulation de l'initiation de la transcription dans le cycle de vie de *P. falciparum*, à la faible quantité d'études sur les acteurs de ce contrôle. Dans ce cadre, nous avons tenté de comprendre le rôle des protagonistes de la régulation de l'expression des gènes pendant les différents stades du cycle érythrocytaire de *P. falciparum*. Afin de mieux appréhender les mécanismes impliqués dans le contrôle transcriptionnel de l'expression des gènes au cours du cycle érythrocytaire, les acteurs de régulation qui pourraient avoir un rôle dans les événements-clés du cycle érythrocytaire ont été analysés. D'une part, la transition de la prolifération asexuée à la différenciation sexuée nous a particulièrement intéressée. Puis, nous avons tenté de découvrir le rôle d'un facteur de transcription spécifique de séquence dans la régulation transcriptionnelle.

Dans la première partie de ma thèse, nous avons cherché à comprendre le rôle, dans la transition de la prolifération asexuée à la différenciation sexuée, d'un certain nombre de protéines dont la fonction prédite pouvait avoir un rôle dans la transduction du signal, le cycle cellulaire et la régulation transcriptionnelle. Pour cela, l'analyse du transcriptome lors du développement érythrocytaire de deux clones de *P. falciparum* a été initiée. Le clone 3D7 qui produit des gamétocytes et le clone F12 dérivant d'une culture continue de 3D7 mais qui a perdu la capacité de produire des gamétocytes, ont été utilisés lors de cette étude. Les données du transcriptome ont été acquises grâce à l'utilisation d'une puce à ADN construite et produite au sein de notre laboratoire, représentant cent cinquante gènes de *P. falciparum*.

Dans la seconde partie du travail, nous avons décidé d'étudier un facteur de transcription spécifique de séquence, PfMyb1 (famille des protéines à domaines tryptophane). L'analyse bioinformatique de la séquence de ce facteur de transcription a permis de révéler qu'il possédait toutes les caractéristiques d'un facteur de transcription de la famille des Myb. Il a

été ensuite montré que la protéine PfMyb1, présente dans les extraits nucléaires, était capable de lier spécifiquement des séquences d'ADN issues des promoteurs de deux gènes de *P. falciparum* et contenant des éléments de régulation identifiés *in silico* comme cibles des protéines Myb. Nous nous sommes alors intéressés aux conséquences de l'extinction partielle de l'expression du transcrit et de la protéine *pfmyb1* à son pic d'expression lors du stade érythrocytaire. Utilisant la puce à ADN ciblée, nous avons étudié les variations de la représentativité des transcrits, nous permettant ainsi d'identifier une partie des gènes cibles sous le contrôle de PfMyb1.

MATERIEL & METHODES

1. Matériels

1.1 Clones de *P. falciparum*

Les deux clones de *P. falciparum* utilisés, au cours de ce travail, ont été gracieusement fournis par le Dr. Walliker (clone 3D7) et le Dr. Alano (clone F12). Ces clones ont été cultivés et synchronisés comme décrit dans les articles ci-joints (page 71).

1.2 Souches bactériennes et plasmides

Les bactéries *Escherichia coli* TOP' 10 (Invitrogen) et SURE (Stratagene), transformées avec le plasmide TOPO II (Invitrogen), nous ont permis de réaliser le clonage des produits de PCR. La souche d'*E. coli* SG13001 (Qiagen), transformée avec le plasmide pQe30 (Qiagen), a été utilisée pour la production de protéine recombinante PfMyb1.

Les cellules transformées avec le vecteur de clonage sont cultivées en milieu LB (Luria-Bertani) en utilisant une sélection antibiotique avec de l'ampicilline (100 µg/ml) et/ou de la kanamycine (50 µg/ml).

1.3 Anticorps

- Anti-PfMyb1 : Le sérum anti-peptide a été préparé par Eurogentec. Deux lapins ont été immunisés avec deux peptides (RKKYEYKKNSWTKK et KRRYLRLISATNNMEQ), correspondants aux aa 103-117 (en amont du domaine R1) et 400-414 (en toute fin du domaine R3) des domaines tryptophane de la protéine PfMyb1. Un sérum capable de reconnaître la protéine recombinante PfMyb1 a été obtenu après trois inoculations (jour 100) et a donc été sélectionné pour être immuno-purifié avec un des deux peptides. Toutes les études ont été réalisées avec le sérum anti-PfMyb1 purifié contre le peptide 400-414 (0.46 mg/mL). Le sérum préimmun, immun et immuno-purifié ont été utilisés pour les analyses en Western blot et pour les retards sur gels.

- L'anticorps anti-p50, un cadeau du Pr. A. Israël, est dirigé contre la protéine p50 du facteur de transcription NF-κB.

- L'anticorps anti-PfPk5 est un don du Dr. C. Doerig.

2. Méthodes

2.1 Puce à ADN ciblée

2.1.1 Les gènes sélectionnés

Lorsque le projet de création d'une puce à ADN ciblée a été décidé, le génome de *P. falciparum* n'était que partiellement accessible dans les banques de données. Les phases ouvertes de lecture, sélectionnées pour être immobilisées sur la puce, sont donc essentiellement le fruit de recherches par similarité de séquence avec des protéines eucaryotes dont la fonction était déjà connue. La sélection s'est alors opérée en ciblant ces phases ouvertes de lecture dont le rôle potentiel est compris dans les domaines suivants : régulation de l'expression des gènes, du cycle cellulaire et de la transduction du signal. La plupart de ces gènes codent donc pour des kinases, et des protéines intervenant dans la réplication (ADN polymérases, ADN topo isomérases, hélicases, protéines nucléaires, etc.), les mécanismes généraux de la transcription (TBP, ARN polymérases), la régulation de l'initiation de la transcription (facteurs de transcription annotés dans laboratoire : PfMyb1, HMG, ...), la stabilité et structure des ARN messagers (ARNr méthylases, hélicases, ...), et la traduction (facteurs d'initiation et d'élongation). Ce travail a été essentiellement accompli par Philippe Refour sous la direction de Catherine Vaquero.

Les séquences ont été amplifiées à partir de l'ADN génomique du clone 3D7 (Eurogentec), vérifiées et référencées avec la nomenclature adoptée lors de la publication du génome complet de *P. falciparum* en octobre 2002 [53]. Il est intéressant de noter que le nombre des kinases putatives prédites à cette époque (environ 60), est très proche du nombre total de kinases prédit à la fin du séquençage de *P. falciparum* [161, 237].

Un certain nombre de témoins nous ont servi au cours des études lors du développement érythrocytaire du parasite. Des gènes dont l'expression avait été décrite dans la littérature comme étant exclusivement exo-érythrocytaire ont été utilisés comme témoins négatifs. Par exemple le gène *LSA-1*, exprimé préférentiellement au stade hépatocytaire [238], et le gène codant l'*actine 2*, exprimé préférentiellement au stade gamétocyte [239]. De la même façon, les gènes *IFN γ* (humain), *gp120* (HIV-1), le vecteur de clonage vide et du tampon de dépôt seul ont aussi été déposés afin de servir de témoins négatifs.

Des gènes dont le profil d'expression tout au long du cycle érythrocytaire avait été décrit dans la littérature, ont été utilisés comme témoins de la qualité de la synchronisation de nos

cultures afin de valider la démarche expérimentale. Par exemple, le gène de l'*histone h2b* [240], *gfp130* [178] ou le gène *kahrp* [180], ont rempli cet office.

La normalisation des signaux après hybridation a été réalisée grâce à la présence sur la puce à ADN de références internes comme des produits de PCR sans similarités de séquence avec le génome de *P. falciparum* (Alien, Stratagene) ainsi que l'ADN génomique de *P. falciparum* à plusieurs concentrations.

2.1.2 Production des ADNc marqués et hybridation

L'ensemble de ces étapes est détaillé dans l'article 1. Ces protocoles sont essentiellement adaptés de la méthode développée par Salin et collaborateurs [241]. Nous avons choisi d'utiliser la radioactivité, *a contrario* des marqueurs fluorescents généralement utilisés, car cette technique semblait particulièrement pertinente dans le contexte de l'acquisition de signaux faibles générés dans le cadre d'une quantité de matériel initial limitée ou dans le cas d'une faible représentation des ARNm dans la population générale. En effet, cette technique nous permet d'acquérir des signaux de manière linéaire pendant des temps longs, et nous autorise ainsi à identifier des signaux de faible intensité (voir l'article 4 en Annexes p.172).

2.1.3 Acquisition et analyse des données

La technologie utilisant la radioactivité a été largement améliorée grâce à l'arrivée d'outils capables de comptabiliser les désintégrations (provenant de chacun des deux isotopes utilisés lors de la co-hybridation des lames) en temps réel et d'attribuer ces désintégrations à chaque isotope. Le Microimager (Biospace mesures) permet donc l'acquisition simultanée de l'information donnée par chaque radio-isotope et la séparation de ces signaux grâce au logiciel Dual Labelling (Biospace mesures).

Les données ainsi produites peuvent ensuite être analysés avec le logiciel β -vision (Biospace mesures) ou le logiciel Arrayvision (Imaging Research).

2.2 RT-PCR quantitative en temps réel

La RT-PCR quantitative en temps réel repose sur l'amplification d'ADNc, issus d'une transcription reverse à partir d'ARN. Théoriquement, il existe une relation directe entre cette amplification et la quantité initiale d'ADNc, donc avec le nombre de copies d'ARNm pour chaque transcrite. Cependant, les variations inhérentes à la réaction de PCR sont connues pour

limiter la capacité à reproduire avec précision cette relation directe et donc à réduire la qualité et la reproductibilité des résultats. La variabilité est généralement due à l'effet des réactifs en concentration limitante sur la qualité de l'amplification. La PCR est caractérisée par une augmentation exponentielle de l'amplification de la cible. Lorsque les réactifs deviennent limitants, le niveau d'amplification de la cible diminue jusqu'à atteindre un plateau où plus aucune amplification n'est possible. L'avantage des techniques de PCR quantitatives en temps réel, réside dans leur capacité à mesurer la quantité de copies de la cible au début de la phase exponentielle, là où l'effet limitant de la concentration des réactifs est moindre. En effet, la quantité de produits de la PCR à la fin de la réaction est dépendante d'une petite variation de la concentration initiale des réactifs [242].

Dans la technique que nous avons employé pour les réactions de PCR quantitatives en temps réel, nous avons utilisé la détection de marqueurs fluorescents capables de se lier à l'ADN. Le marqueur non lié ne peut émettre qu'une faible fluorescence en solution. Au contraire, lorsque le marqueur se lie à l'ADN, au cours de l'élongation du produit de PCR, il est capable de fluorescer. C'est l'augmentation du signal de fluorescence qui est mesurée en temps réel et qui rend compte de l'augmentation du nombre de copies du produit de PCR [242]. La quantité de fluorescence étant mesurée après chaque cycle d'amplification, l'expérimentateur peut suivre cycle après cycle la réaction (Figure 16).

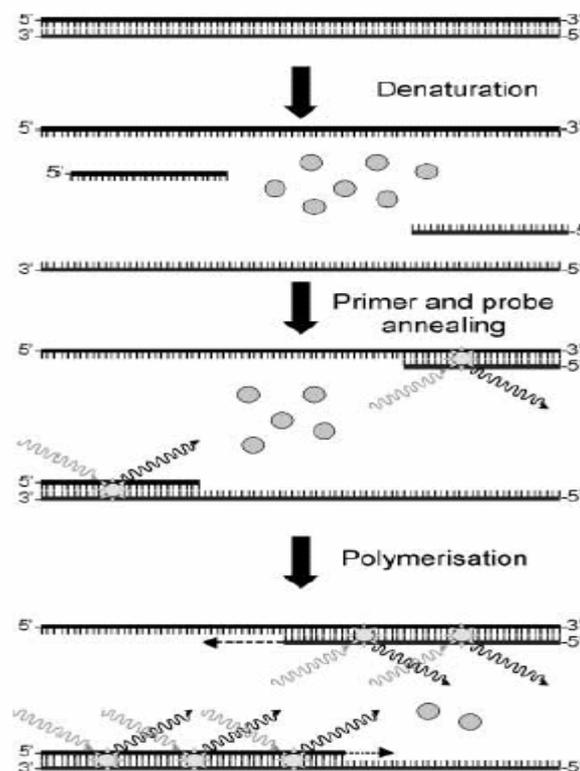


Figure 16. Incorporation du marqueur fluorescent durant l'amplification d'après [242].

L'inconvénient essentiel de l'utilisation de ces marqueurs fluorescents est le manque de spécificité de cette incorporation. En effet, si les amorces utilisées sont capables d'amplifier une autre séquence d'ADN dans le génome, alors la quantité de fluorescence totale sera prise en compte et pourra fausser la mesure du nombre initial de transcrite présent dans le pool d'ARN. La spécificité de la réaction repose donc essentiellement sur les amorces. Afin de vérifier si l'amplification n'a pas donné lieu à la production de plusieurs amplicons, les réactions de PCR sont déposées sur un gel d'acrylamide coloré par le BET.

La reproductibilité et la précision de la mesure est basée sur la détermination du cycle seuil (C_t) à partir duquel l'amplification du produit de PCR est détectable. Plus il existe de copies d'un ADNc initialement dans la réaction, plus le C_t est petit car l'amplification est plus rapidement détectée. Le seuil est déterminé à partir d'une amplification sans ADNc et permet la mesure du C_t (Figure 17).

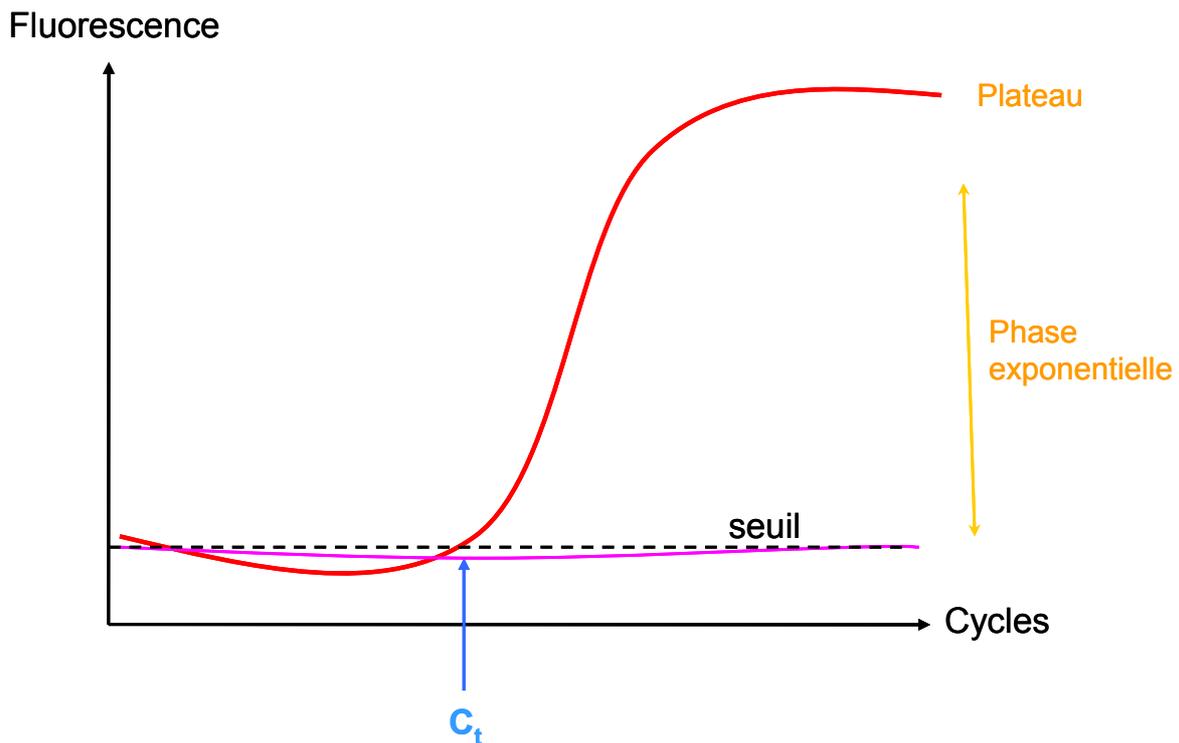


Figure 17. Représentation graphique de l'intensité de fluorescence en fonction du nombre de cycles pour deux amplifications à partir d'ADNc (en rouge) ou sans ADNc (en fuchsia).

L'amplification d'une référence interne pour chaque condition testée, permet la normalisation des réactions. Nous avons choisi d'amplifier le gène *18s* (chr5.rRNA-1-18s-A) de *P. falciparum* afin de pouvoir réaliser le contrôle interne de la quantité initiale d'ARN total pour

chaque échantillon. Le C_t obtenu pour chaque gène dans chaque condition sera normalisé par rapport au C_t de ce gène contrôle. Afin de pouvoir exprimer la variation de la quantité d'ADN par rapport à une autre condition, nous avons utilisé l'équation $2^{-\Delta\Delta C_T}$ [243]. Celle-ci permet de comparer des quantités d'ADNc normalisées et de tracer ainsi une cinétique d'expression d'un transcrit.

2.3 Immuno-précipitation de la chromatine

L'immuno-précipitation de la chromatine est une méthode permettant d'identifier les endroits dans le génome où une protéine interagit avec l'ADN (Figure 18). L'étape initiale est la création dans les cellules vivantes d'une liaison réversible entre ADN et protéines. La formaldéhyde est l'agent permettant cette réaction. L'avantage principal de ce réactif est qu'il est soluble dans l'eau, actif dans une gamme large de température et de pH, et surtout qu'il possède la capacité de traverser les membranes biologiques. Cette liaison chimique peut donc être réalisée sur des cellules vivantes, limitant ainsi les phénomènes de redistribution et de réassociation des protéines sur la chromatine durant la préparation des échantillons. Les cibles chimiques de la formaldéhyde semblent être les groupes amines primaires comme ceux présents dans les lysines [244]. Cette liaison chimique est réversible par une incubation des échantillons à 65 °C. Après liaison des protéines avec l'ADN, les extraits nucléaires sont préparés et ensuite soniqués afin d'obtenir des fragments d'ADN de taille uniforme.

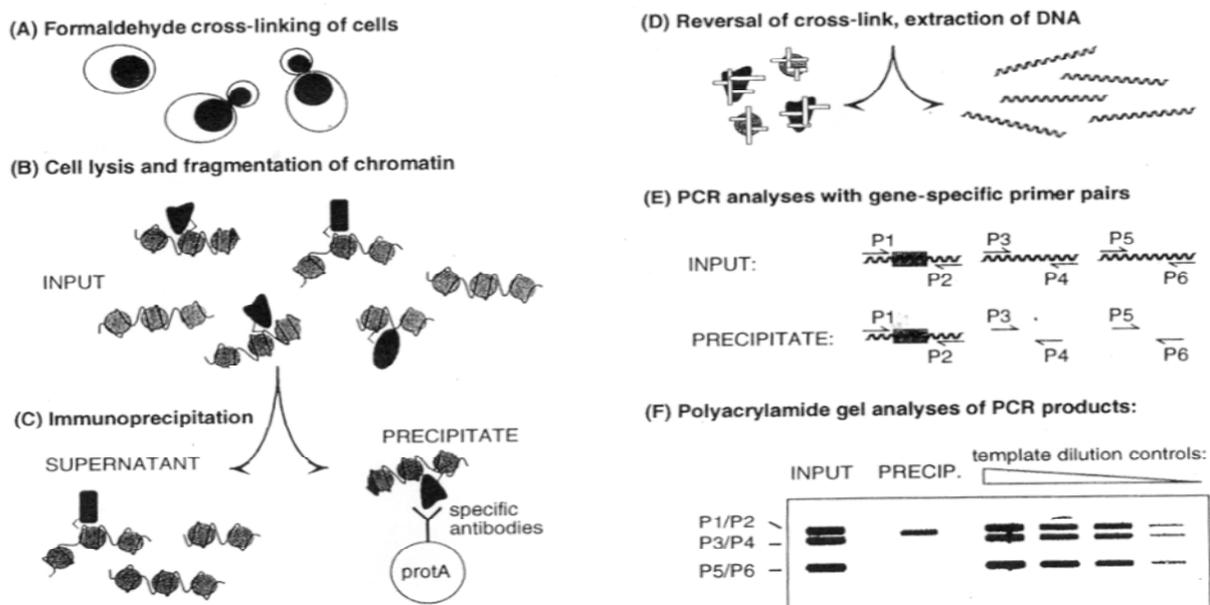


Figure 18. Principe de l'immuno-précipitation de la chromatine d'après [244]

Les extraits contenant les fragments d'ADN liés aux protéines sont alors immuno-précipités à l'aide d'un anticorps spécifique de la protéine d'intérêt. La liaison chimique entre ADN et protéines est alors réversée, et les fragments d'ADN purifiés. Si la protéine d'intérêt est associée de manière spécifique avec une région du génome *in vivo*, les fragments d'ADN provenant de cette dernière seront alors enrichis dans l'extrait immuno-précipité par rapport à la population totale. La présence ou l'absence d'une région d'ADN dans l'extrait immuno-précipité peut être déterminée grâce à une réaction de PCR utilisant des amorces spécifiques de cette région.

Cette technique a été utilisée afin de déterminer si la protéine PfMyb1 pouvait interagir avec un certain nombre de promoteurs. L'anticorps anti-PfMyb1 nous a servi pour les immuno-précipitations. Les détails de cette approche sont décrits dans l'article 3 (p.98).

RESULTATS

1. Recherche de facteurs impliqués dans la transition de la prolifération asexuée à la différenciation sexuée.

1.1 Résumé

Nous avons vu dans l'introduction que l'arrêt de prolifération asexuée et l'entrée dans la différenciation sexuée étaient influencés par de nombreux facteurs environnementaux. Le déclenchement du cycle sexué, impliqué dans l'échappement du parasite hors de l'hôte vertébré, reste pour l'instant largement inconnu. Différents travaux ont permis de caractériser des clones de *P. falciparum* incapables de produire des gamétocytes [36]. L'absence de production de gamétocytes et de transcrits connus pour être les premiers marqueurs de la différenciation sexuée dans le clone F12, nous a conduit à étudier l'expression de ces transcrits tout au long du cycle érythrocytaire afin de pouvoir découvrir des gènes dont l'expression pourrait être impliquée dans la gamétocytogénèse.

L'implication, dans l'initiation de la gamétocytogénèse, des voies de transduction du signal et des facteurs de transcription ayant été suggérée, nous avons utilisé une puce à ADN, construite et produite au sein de notre laboratoire. Celle-ci contient environ 200 fragments de PCR dont 150 sont représentatifs de gènes de *P. falciparum*, potentiellement impliqués dans la régulation transcriptionnelle, traductionnelle et dans la transduction du signal et le cycle cellulaire.

La comparaison des profils d'expression obtenus avec le clone 3D7 avec ceux décrits dans la littérature pour les clones 3D7 [138] et HB3 [171], nous a permis de confirmer la solidité des résultats obtenus avec le clone F12. Nous avons donc pu identifier plusieurs gènes dont les profils d'expression sont différents entre les deux souches. Trois gènes dont la fonction probable pourrait les assigner à la régulation de la transcription, parmi lesquels un facteur de transcription de la famille des protéines Constans, ont particulièrement retenu notre attention. En effet, ces gènes pourraient être impliqués dans la régulation transcriptionnelle de la gamétocytogénèse.

Article 1 :

Transcriptome of 3D7 and its gametocyte-less derivative F12 *Plasmodium falciparum* clones during erythrocytic development using a gene-specific microarray assigned to gene regulation, cell cycle and transcription factors.

Mathieu Gissot, Philippe Refour, Sylvie Briquet, Charlotte Boschet, Stéphane Coupe, Dominique Mazier et Catherine Vaquero.

Accepté à la publication dans *Gene*.

Transcriptome of 3D7 and its gametocyte-less derivative F12 *Plasmodium falciparum* clones during erythrocytic development using a gene-specific microarray assigned to gene regulation, cell cycle and transcription factors

Mathieu Gissot, Philippe Refour, Sylvie Briquet, Charlotte Boschet, Stéphane Coupé¹,
Dominique Mazier, Catherine Vaquero*

INSERM U511, CHU Pitié-Salpêtrière, Université Paris 6, 91 boulevard de l'Hôpital, 75013 Paris, France

Received 28 March 2004; received in revised form 21 June 2004; accepted 5 July 2004

Available online 13 September 2004

Abstract

During the complex life cycle of *Plasmodium falciparum*, through mosquito and human, the erythrocytic cycle is responsible for malarial disease and transmission. The regulation of events that occur during parasite development, such as proliferation and differentiation, implies a fine control of transcriptional activities that in turn governs the expression profiles of sets of genes. Pathways that underline gametocyte commitment are yet poorly understood even though kinases and transcription factors have been assumed to play a crucial role in this event. In order to understand the molecular mechanisms controlling the variation of gene expression profiles that might participate in early gametocytogenesis, the transcriptome of two clones, 3D7 and its gametocyte-less derivative F12, was compared at five time points of the erythrocytic asexual development. We have used a thematic DNA microarray containing 150 PCR fragments, representative of *P. falciparum* genes involved in signal transduction, cell cycle and transcriptional regulation. We identified several genes eliciting different expression profiles among which some implicated in gene regulation or encoding putative transcription factors. The differential expression of transcription factor and kinase transcripts observed in the two clones may enlighten genes that might have a role in impairment of the early gametocytogenesis of the F12 clone.

© 2004 Elsevier B.V. All rights reserved.

Keywords: *Plasmodium*; Transcription factor; Microarray; Erythrocytic development; Gametocytogenesis

1. Introduction

Plasmodium is responsible for 1.5–2.7 million deaths annually, mostly children and pregnant women. Among the

four species of parasites infecting humans, *Plasmodium falciparum* causes the highest morbidity and mortality. Global efforts to eradicate malaria have failed and no effective vaccine is available. Therefore, insight into the parasite biology and mechanisms of gene and cell cycle regulation during the development of *Plasmodium* is urgently required to provide new targets for the control of malaria. The cell cycle of *Plasmodium* between vertebrate host and mosquito passes by a succession of different stages characterized by an active cellular division (as the exoerythrocytic and erythrocytic schizogony) and an arrest of proliferation associated to differentiation (as for gametocytes and sporozoites). Actually, two major events of the *P. falciparum* life cycle occur in the erythrocyte, an asexual multiplication responsible for the clinical manifestations and the initiation of sexual differentiation responsible for dissemination of the disease via the mosquito. Those two

Abbreviations: aa, amino acid; CO, Constans; ES, early schizont; ET, early trophozoite; LS, late schizont; LT, late trophozoite; NAP, nucleosome assembly protein; NLS, nuclear localization site; ORF, open reading frame; PCR, polymerase chain reaction; qPCR, real time quantitative PCR; R, ring; RT, reverse transcription; TF, transcription factor; dNTP, deoxyribonucleoside triphosphate; DNase, deoxyribonuclease; RNase, ribonuclease; UTR, untranslated region(s); cDNA, DNA complementary to RNA; IFN, interferon; SDS, sodium dodecyl sulfate; SSC, 0.15 M NaCl/0.015 M Na₃ citrate pH 7.6.

* Corresponding author. Tel.: +33 1 40778111; fax: +33 1 45838858.

E-mail address: vaquero@ext.jussieu.fr (C. Vaquero).

¹ Present address: Laboratoire de Parasitologie-Mycologie, Centre Hospitalier Universitaire Lariboisière-Saint-Louis, 75010 Paris, France.

events are under the control of coordinated expression of distinct sets of genes governed by transcriptional regulation even though post-transcriptional events could also be implicated.

Pathways that control asexual cycle regulation and switch to sexual commitment are yet poorly understood although a number of genes described to command cell cycle in other eukaryotes, e.g. kinases and transcription factors, have been predicted in our laboratory or by the sequencing project of *P. falciparum* (Gardner et al., 2002). Two publications have described the whole genome examination of the transcriptome of HB3 (Bozdech et al., 2003) and 3D7 (Le Roch et al., 2003) clones of *P. falciparum* giving the first global idea of the expression profiles of almost all genes during erythrocytic asexual cycle and at gametocyte stage, therefore opening new perspectives for the discovery of gene implication in those pathways. Various genes were reported to be associated with induction of gametocytogenesis, as *pfg27* and *pfs16* (Dyer and Day, 2000b). Moreover, implication of cyclic AMP-dependent pathway (Inselburg, 1983) and trimeric G proteins (Dyer and Day, 2000a) suggests participation of kinases and transcription factors in this phenomenon. Sexual commitment have been reported to be sensitive to many environmental conditions including host immunity, anti-malarial drugs, host hormones, erythrocyte intracellular environment and autocrine growth control (reviewed in Dyer and Day, 2000b). The parasite response to environment that influences the developmental decision, point out the key role of signaling pathways comprising cell cycle regulators, kinases and transcription factors leading to transcriptional control of sexual commitment. Clones of *P. falciparum* lacking or with a reduced ability to produce gametocytes have been reported. Although those alterations in sexual development have been considered to be irreversible, drug-driven induction of gametocytogenesis was observed in clones thought to be impaired in sexual development (Ono and Nakabayashi, 1990) suggesting a reduced sensitivity to environmental changes. Subtelomeric deletion of chromosome 9 also induces almost total loss of gametocyte production (Day et al., 1993). However, total or partial failure of the ability to produce gametocytes has been shown to occur in clones, derived from 3D7, carrying an intact chromosome 9 and no visible changes in karyotype (Alano et al., 1995). F12 clone has been reported to be one of the gametocyte-less 3D7 derivative. In addition, loss and decrease of mRNA expression of the two earliest landmarks of gametocyte differentiation, the cytoplasmic protein *Pfg27* and the membrane protein *Pfs16*, respectively, have been stated (Alano et al., 1995). However, the role of *Pfg27* in sexual differentiation remains unclear. Its expression have been described to be developmentally regulated (Alano et al., 1996), required for early stage of gametocytogenesis and reported to play a role in a part of *Pfs16* expression induction (Lobo et al., 1994, 1999). Moreover, crystal structure of *Pfg27* raised the hypothesis of potential binding

to other proteins such as kinases through SH3 domains and to RNA (Sharma et al., 2003). Nature of the impairment of *pfg27* transcript expression could be due to absence or altered-expression of genes implicated in the control of its expression, since major rearrangement upstream to the *pfg27* coding sequence was not observed in F12 clone (Alano et al., 1996). The absence of expression of *pfg27* transcript and therefore loss of gametocyte production ability remain unclear in F12 clone; whether it has lost the sensitivity to external stimuli or the ability to respond to these stimuli also remains to be elucidated.

In the investigation described here, we constructed a gene-specific microarray encompassing 153 *P. falciparum* PCR products representing genes that might be implicated in regulation of cell cycle and transcription. We hybridized cDNA, obtained from total RNA prepared from five time points of asexual erythrocytic culture of the two clones, 3D7 and F12, and labeled with two distinct radioisotopes. We compared expression profiles of the 153 genes in these two clones grown under normal culture condition. Indeed, we tried to identify genes with differential expression profiles in the F12 clone that does not produce gametocytes and in 3D7 that bear the potentiality to produce gametocytes. This analysis was carried out to underline differentially expressed genes particularly focusing on putative transcription factors. Differential expression of transcription factors and kinases might have a role in the impairment in early stage of gametocytogenesis.

2. Materials and methods

2.1. *P. falciparum* clones and cultures

P. falciparum 3D7 and F12 (Alano et al., 1995) were provided by Dr. D. Walliker and Dr. P. Alano, respectively, and were cultured as described in Trager and Jensen (1976) with slight modifications. To obtain synchronized erythrocytic stages, ring cultures were enriched by three successive treatments of 5% sorbitol at 44-h intervals. Five stages of development, rings (10–14 h), early (20–22 h) and late (26–28 h) trophozoites, and early (32–34 h) and late (40–44 h post-invasion) schizonts, were determined according to time and morphological analysis by GIEMSA staining. At each indicated stage, parasite culture at 8% parasitemia was collected by centrifugation and used for preparation of transcripts for further analyses.

2.2. Amplification of PCR products and microarray construction

Microarrays were prepared by printing 170 PCR products 300–1500 bp long produced from sequences cloned in the Topo 2.1 vector (Invitrogen), representing mostly genes implicated in signal transduction and cell cycle of *P. falciparum*. PCR fragments from all clones (Eurogentec)

were purified by isopropanol precipitation and resuspended in spotting buffer (Tris 10 mM pH 7.4, EDTA 1 mM, DMSO 50%) adjusted to an average concentration of 200 ng/ μ l and produced with specific pairs of primers found in the vector; T7 (5' TAATACGACTCACTATAGGG 3') and M13R (5' GGATAACAATTTACACAGG 3').

Seventeen spots, representing positive and negative controls such as sequences unrelated to *P. falciparum*, were printed. Alien PCR products from Stratagene, with no apparent cross-similarity with *P. falciparum* genomic sequence, were used for normalization of the reverse transcription. 3D7 genomic DNA was also printed in order to normalize the total amount of each cDNA present during the hybridization. All targets were spotted in duplicate with a Qarray arrayer (Genetix) onto poly-L-lysine-coated slides prepared as described previously (Schena et al., 1995) with a spot spacing of 400 μ m, center to center onto a 1 cm² area. DNA was cross-linked by UV irradiation at 3 kJ with Stratalinker model 1800 UV Illuminator (Stratagene). Neutralization of free poly-L-lysine was performed by chemical treatment with succinic anhydride (Sigma). Prior to hybridization, the slides were incubated for 1 min at 95 °C to denature the spotted DNA and dehydrated with a 96% ethanol immersion for 1 min and dried by centrifugation.

2.3. Probe labeling, hybridization and data analysis

Parasitized red blood cells were rapidly lysed in TRIZOL (Invitrogen) and total RNA was extracted at different times during erythrocytic development from 3D7 and F12 *P. falciparum* cultures. Concentration was determined by spectrophotometer and integrity of the RNA preparations was verified by ethidium bromide staining on agarose gel and by Agilent 2100 bioanalyser.

Ten micrograms of total RNA extracted at different times of erythrocytic development were labeled during reverse transcription reaction in a sample mixture containing 2 μ g of 15-mer poly-T primer (Invitrogen), 5 mM of MgCl₂, 10 mM of DTT, 80 U of RNAout (Invitrogen), 0.2 mM of each dCTP, dGTP and dTTP, 50 μ Ci/ μ l of ³⁵S-dATP (Amersham, catalog number SJ1134) or 40 μ Ci/ μ l of ³H dATP (Amersham, catalog number TRK633). Reaction was performed at 42 °C for 2 h with 250 U of Superscript II reverse transcriptase and subjected to ribonuclease (RNase)-H treatment with 2 U of enzyme for 20 min at 37 °C. Labeled cDNA were purified on QIAquick nucleotide removal columns as indicated by the manufacturer (Qiagen) and eluted in 40 μ l. Amount of radioactivity contained in 1 μ l of purified eluate was evaluated using a Beckman Coulter LS 6500 apparatus.

For hybridization, the labeled probe mixture corresponding to 1.6×10^6 disintegrations per minute for each cDNA was used per slide. Each labeled cDNA was added to the hybridization buffer (3.5 \times SSC, 0.3% SDS, 0.5 μ g/ μ l DNA Salmon Sperm and 0.5 μ g/ μ l yeast tRNA), heated to 95 °C for 2 min, cooled down at room temperature and added on

the microarray slide under a cover slip (Easy seal, Hybaid). Hybridization was performed in a cassette chamber (Tele-Chem) submerged in a water bath at 62 °C for 14–16 h. Arrays were washed at room temperature in a 2 \times SSC, 0.1% SDS solution prior to washing in a 2 \times SSC and in a 0.2 \times SSC solution successively, each step for 2 min, and finally dried by centrifugation.

Acquisition of arrays was performed as described previously on a Microimager (Biospace Measure) (Salin et al., 2002), and stopped after 24 h acquisition. Data were collected as an image file and the ARRAYVISION software (Imaging Research) was used for quantification of the hybridization intensities and for normalization. The local background was subtracted from each spot intensity. The intensity of each spot was normalized according to the median value of intensities of spots corresponding to Alien cDNA PCR products on each array in order to assess normalization of the reverse transcription. Furthermore, genomic DNA was used to normalize the amount of cDNA present in the hybridization. For all genes, expression profile was expressed as the log₂ ratio of labeling intensity of experiment over reference. Two independent experiments and one technical replicate were performed.

2.4. Genomic DNA labeling

Genomic DNA (30 ng) isolated from F12 and 3D7 clones were ³⁵S-labeled using a random-primed DNA labeling kit (Amersham). Labeled genomic DNAs were purified using QIAquick columns (Qiagen) and hybridized as described above.

2.5. Real time quantitative RT-PCR (qPCR)

Primers were designed using PrimerExpress software (Applied Biosystem). Total RNA (10 μ g) was treated with 10 U of RNase-free deoxyribonuclease (DNase) I (Qiagen), and cDNA synthesis was performed for 50 min at 42 °C using 1 μ g of RNA, 50 U of MMLV reverse transcriptase (Superscript II, Invitrogen), 0.5 ng of random hexamers, 5 mM of MgCl₂, 20 U of RNaseout (Invitrogen), 10 mM of DTT and 0.25 mM of each dNTP. Finally, reaction was subjected to RNase-H treatment with 2 U of enzyme for 20 min at 37 °C. Real time quantitative PCR reaction was performed in a 20 μ l reaction volume containing 5 μ l of a dilution of cDNA preparation, 10 μ l of SYBR green PCR master mix (Sigma) and 0.375 μ M of gene-specific primers (Invitrogen). Amplification and detection of specific products were performed with the MX4000 light cycler (Stratagene) with the following cycle profile: one cycle at 95 °C for 2 min, 40 cycles with 30 s denaturation at 95 °C and 1 min annealing-elongation at 60 °C. Size of each PCR product and number of bands was verified on a 10% acrylamide gel. Two additional reactions were performed, either without reverse transcriptase or without the RNA sample, to verify the absence of DNA contamination. The

quantity of cDNA for each experimental gene was normalized to the 18S concentration (chr5.rRNA-1–18s-A) in each sample. For each time point of each gene, experiments were performed twice and in triplicates. Relative gene expression was expressed via the \log_2 of the ratio of the expression time point over ring stage using the $2_T^{-\Delta\Delta C}$ method (Livak and Schmittgen, 2001).

3. Results

3.1. *P. falciparum* gene-specific microarray, experimental design and validation of the results

We constructed a microarray encompassing 170 PCR products among which 153 amplified from the *P. falciparum* genomic sequences. This thematic microarray could evolve by adding new genes of interest when needed. All sequences to be amplified were localized within the open reading frame (ORF) close to the 3' end and were sequenced prior to be arrayed on glass slides.

The *Plasmodium* genes were selected for their homology to known proteins previously described in eukaryotes to be involved in genetic regulation from DNA to proteins. Indeed, glass slides were printed with PCR fragments coming from genes encoding putative kinases controlling signal transduction and cell cycle, proteins involved in replication (DNA polymerases, DNA topoisomerases and helicases), proteins involved in general transcription machinery (TATA binding protein, small nuclear ribonucleoprotein, RNA polymerases), transcription factors (TFs) (such as three members of the Myb family, two members of the HMG family and several proteins encompassing zinc finger domains), as well as proteins involved in the stability and the structure of mRNA (ribosomal RNA methylases, helicases), in translation (initiation and elongation factors), and in chromatin remodeling (histone deacetylase, nucleosome assembly protein). Several of these proteins have already been described in the literature, others annotated in databases such as PlasmoDB (www.plasmodb.org) or annotated by our group among which putative transcription factors and kinases. For DNA printing, several negative controls were used among which *pflsa1* only expressed during exo-erythrocytic cycle, unrelated genes as human gamma IFN and HIV-1 gp120 envelope protein, empty plasmid vector used to clone each PCR product and mouse or human genomic DNA. Several *Plasmodium* genes reported to be expressed during asexual erythrocytic cycle, among which *hsp70* (PF08_0054), *H2A* (MAL6P1.249) and *H3* (MAL6P1.248), *kahrp* (PFB0100c), *pcna1* (PF13_0328) and *gbp130* (PF10_0159), were spotted as positive controls at different positions on the microarray. Finally, PCR products corresponding to the Alien (Stratagene) synthetic RNAs, added into total RNA samples to normalize the effectiveness of the cDNA probe production, as well as two different *Plasmodium* genomic DNA

concentrations, used to normalize the total amount of each cDNA present in the hybridization, were also printed.

We performed radioactive labeling of the cDNA reverse transcribed from all the total RNA samples, in order to minimize background signal, enhance signal detection of low abundant mRNA species and avoid signal saturation. This experimental approach was also further adapted to very low concentrations of biological material (Refour et al., submitted for publication). A CCD camera determined, in a real time manner, the signals obtained after hybridization of the microarray with the two differentially ^{35}S or ^3H labeled cDNA. For each *P. falciparum* clone, two independent technical replicates as well as two different biological RNA preparations from independent culture were used for cDNA synthesis. A representative picture of co-hybridization data obtained with the cDNA labeled with the two radioisotopes is presented in Fig. 1. Negative controls showed no signal as for *pflsa1* (PF10_0356) (a), human genomic DNA (b), human gamma IFN (c) and HIV-1 gp120 protein (d), in contrast to positive controls, genes expressed during asexual erythrocytic cycle (1–3). In addition, three PCR products along the sequence of PF14_0316 ORF were spotted to control the PCR product sequence effect on the expression profile (4–6). Moreover, the same PCR product of PF11_0098 ORF (7–8) was spotted twice to control the localization effect and all gave similar signals. Duplicate spots appeared to give similar signal levels. In addition, self-hybridization between ring stage cDNA labeled with ^{35}S and ring stage cDNA labeled with ^3H gave ratio very close to one (data not shown).

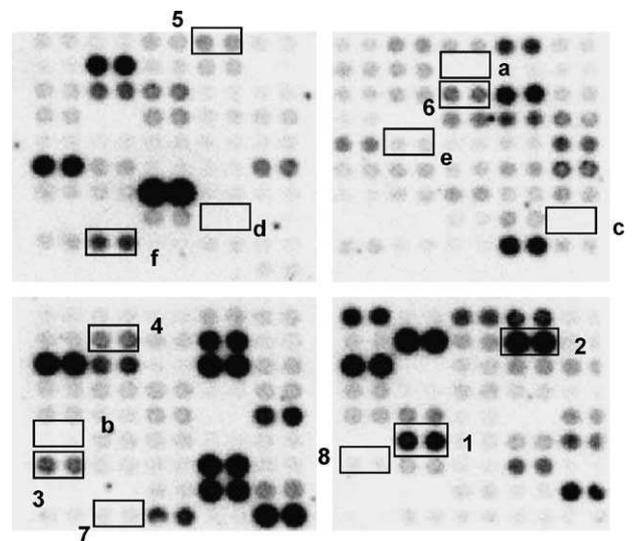


Fig. 1. A representative picture of a co-hybridization of two cDNA labeled in the presence of ^{35}S dATP and ^3H dATP. Framed duplicated spots represent (a) *pflsa1* (PF10_0356), (b) human genomic DNA (20 ng/ μl), (c) human gamma IFN, (d) HIV gp120 protein, (e) Alien 1 from Stratagene, (f) Alien 2 from Stratagene; (1) *pflhsp70* (PF08_0054), (2) *histone H2A* (MAL6P1.249), (3) *gbp130* (PF10_0159); (4, 5, 6) three different PCR products for PF11_0098; (7, 8) the same PCR fragment for PF14_0316 spotted twice.

When performing the first time course experiment for the 3D7 clone, we selected five genes as positive controls for which expression has been previously described. We compared throughout development the peaks of expression obtained either by our microarray or qPCR to published data (Table 1). The different expression profiles were equivalent and in good agreement with the published data, therefore validating the quality of our experimental approach.

Finally, prior to the first microarray differential experiments, genomic DNA from F12 clone were labeled and hybridized to the microarray in order to verify that the level of homology between F12 and 3D7 sequences was sufficient to determine and compare the abundance of all messengers of each clone. Actually, all immobilized PCR products amplified from 3D7 clones hybridized to labeled genomic DNA extracted from F12 (data not shown). In addition, we verified by real time quantitative RT-PCR that total RNA prepared from the F12 different stages was devoid of *pfg27* transcript (data not shown).

3.2. Comparison between 3D7 and F12 expression profiles

The expression profiles of 3D7 clone and its gametocyte-less derivative F12 clone were monitored during the erythrocytic cycle. Highly synchronized cultures were harvested at five time points during the asexual cycle. Five stages of development, rings (10–14 h), early (20–22 h) and late (26–28 h) trophozoites, and early (32–34 h) and late (40–44 h post-invasion) schizonts, were assessed according to time and morphological analysis. The cDNA obtained from ring total RNA of each *Plasmodium* clone was selected as the reference RNA, labeled with ³⁵S dATP and co-hybridized with the cDNA labeled with ³H dATP

prepared from total RNA extracted at each indicated time of development. The five time points are rings (R), early (ET) and late trophozoites (LT), and early (ES) and late schizonts (LS). For all genes, the steady state RNA profile was expressed as the log₂ ratio of labeling intensity of experiment (R, ET, LT, ES, or LS) over reference (R). In addition, signal intensity was normalized twice, first with the Alien cDNA to normalize the RT efficacy and second with the genomic *Plasmodium* DNA in order to normalize the amount of labeled cDNA loaded during the hybridization process.

Among the 153 *Plasmodium* genes spotted on the array, 23 genes including the seven negative controls, reported to be expressed only during exo-erythrocytic cycle, gave no detectable signal. Among the remaining 130 genes, 16 exhibited too low expression intensity and therefore no reproducible results for the three replicates assayed. When the 114 expression profiles obtained for 3D7 clone were compared to the two recently published transcriptome data for HB3 (Bozdech et al., 2003) and 3D7 clones (Le Roch et al., 2003), more than 80% (82,5%) of our expression results (see Supplementary data 1) were similar to at least one of these published data.

We compared the expression profiles of the 114 genes expressed during the erythrocytic cycle of the two clones and found that 106 genes shared similar profiles, consistent with the fact that F12 clone derived from 3D7 (Alano et al., 1995) (see Supplementary data 2). However, eight genes displayed differential expression and those results were confirmed by real time quantitative PCR experiments. These genes were clustered in three groups highlighting the differences between the two clones (Fig. 2). First, two kinases genes (*MAL6P1.146* and *PF13_0258*) with nearly no modification in expression throughout 3D7 cycle in contrast to what was observed during F12 development with a maximal expression at LT for *MAL6P1.146* and ES for *PF13_0258* (Fig. 2A). Second, two genes (*PFC0485w* and *PFB0865w*) presented opposite profiles. The *PFC0485w* gene, a putative Ser/Thr kinase, was found to have a maximal expression during ring stage and a minimal expression during schizont stage in 3D7 clone in contrast to the expression profiles of F12 clone. The expression of *PFB0865w*, a putative small nuclear ribonucleoprotein, peaked in 3D7 clone in LT and was always higher than the level observed in rings, whereas mRNA relative amount decreased from ring to schizont in F12 (Fig. 2B).

Third, the maximal peak of expression of genes represented in Fig. 2C was reached for 3D7 at different stages of development forwarded or delayed when compared to the F12 clone. *PF10930c*, a nucleosome assembly protein and *PFC0805w*, a putative DNA-directed RNA polymerase II largest subunit, showed a maximal peak at ET stage in 3D7 clone whereas the maximal peak was observed later in F12 cycle. On the contrary, in the case of *PF07_0073*, a seryl-tRNA synthetase and *PF14_0383*, a Constans-like putative transcription factor, the maximal

Table 1
Comparison of the stage where maximal expression was observed for five control genes analyzed in 3D7 clone by microarray experiments, real time quantitative PCR or published data

Gene	Technical procedure		
	Microarray ^a	qPCR ^b	References
Histone H2A MAL6P1.249	S ^c	S	S (Lobo and Kumar, 1999)
KAHRP PFB0100c	R	R	R (Lanzer et al., 1992a)
GBP130 PF10_0159	LT	LT	LT (Lanzer et al., 1992b)
Pfs 40 ^d PF11_0098	S	S	S (La Greca et al., 1997)
PCNA1 PF13_0328	ES	ES	ES (Horrocks et al., 1996)

^a The stage of maximal expression was determined from the microarray data for the 3D7 clone.

^b By real time quantitative RT-PCR experiments.

^c S stands for schizonts, R for rings, LT for late trophozoite and ES for early schizonts.

^d Results are given for the three different PCR products as stated in Materials and methods.

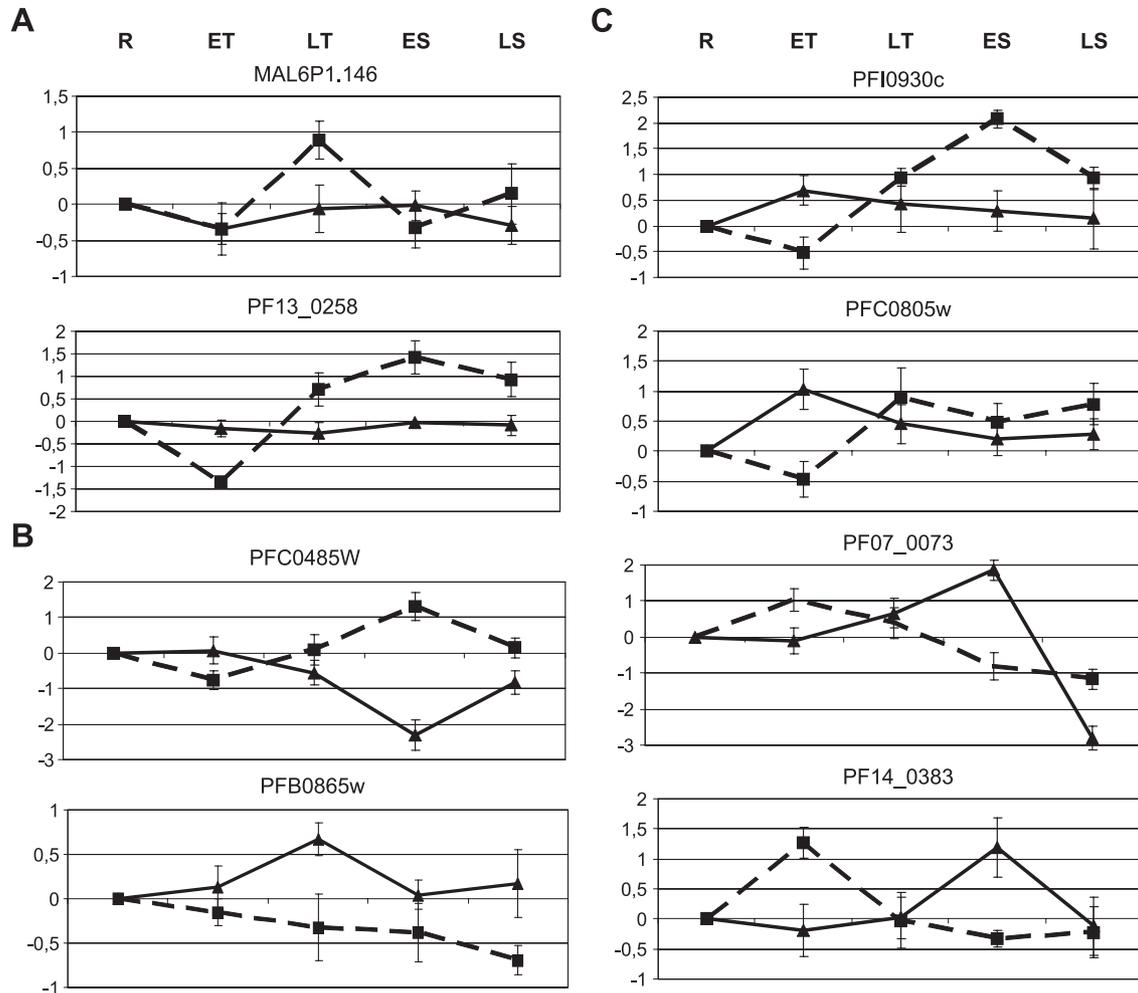


Fig. 2. Graphic representation of the expression profile of eight genes showing differential expression between the two clones 3D7 and F12. For each gene, the steady state RNA profile was expressed as \log_2 ratio of labeling intensity of experiment (any time point) over reference (ring). 3D7 expression profile is represented by a continuous line with black triangles and F12 expression profile by a dash line with black squares. R stands for rings, ET for early trophozoites, LT for late trophozoites ES for early schizonts and LS for late schizonts, stated at the top of the figure. For each clone, the average of three experiments are shown as well as standard deviation. The genes have been clustered in three groups according to their expression profiles: (A) two kinases MAL6P1.146 and PF13_0258, (B) PFB0865w, a putative small nuclear ribonucleoprotein, and PFC0485w, a putative serine/threonine kinase, and, (C) PFC0805w, a putative DNA-directed RNA polymerase II largest subunit, PFI0930c, a nucleosome assembly protein, PF14_0383, a Contrans-like putative transcription factor, and PF07_0073, a seryl-tRNA synthetase.

peak was reached in the ES in 3D7 and in ET for the F12 clone (Fig. 2C).

3.3. Comparison of the transcriptional machinery proteins in 3D7 and F12 clones

We particularly focused our attention on the putative transcription factors and proteins of the transcriptional machinery, some previously annotated by our group, among which several are under molecular and functional investigation.

All 15 transcription factors were developmentally regulated during erythrocytic asexual cycle of the two clones with essentially similar patterns of expression. In Fig. 3, we selected five putative TFs presenting different expression patterns throughout erythrocytic development however similar in 3D7 and F12 clones by microarray experiment

(Fig. 3, panel A) and confirmed the data by real time quantitative PCR experiment (Fig. 3, panel B). *Pfhmg2* (PFL0290w) was expressed at a similar level throughout the asexual cycle with slight variations when compared to ring stage. In contrast, the relative amount of *pfphD2* declined strongly more than four times from R to ES re-increasing thereafter. For the last three transcripts, maximal expression was observed in ET for *pfmyb3* (PF10_0143), in ES for *pfphDB* (PFL1905w) and LS for *pfkrox* (MAL13P1.76). All the messengers coding for these TF were developmentally regulated peaking at different times during the erythrocytic asexual cycle.

However, as already stated in Fig. 2, several messengers encoding various proteins of the transcription machinery exhibited differential expression profiles when comparing the 3D7 to F12 clone. Those transcripts encode a putative nuclear ribonucleoprotein (PFB0865w), a DNA-directed

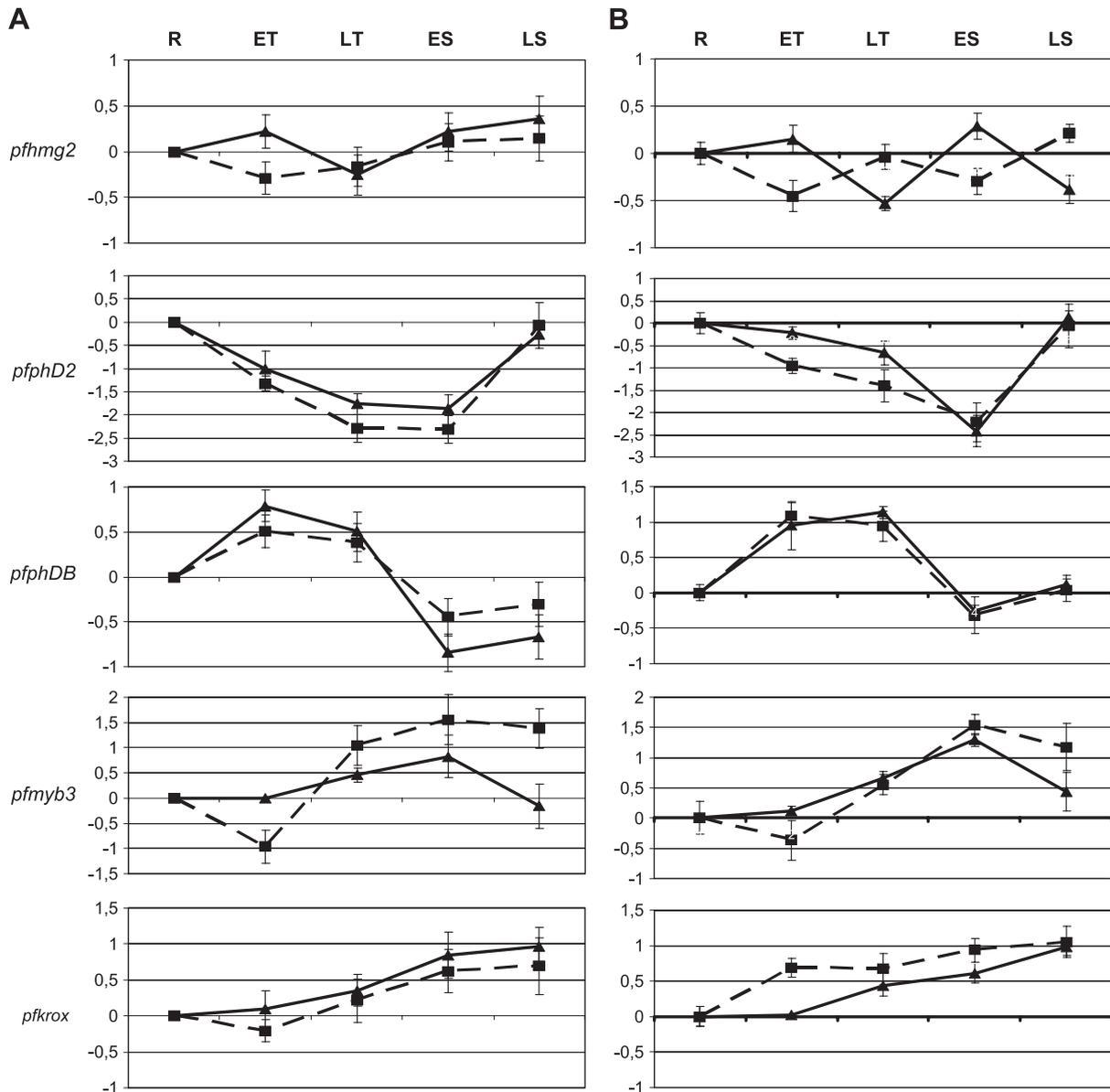


Fig. 3. Graphic representation of the expression profile of five putative transcription factors in the two clones 3D7 and F12 monitored via microarray and qPCR. For each gene, expression profile was expressed and represented as mentioned in Fig 2. (A) Expression profile obtained from microarray experiments. (B) Expression profile derived from real time quantitative PCR experiments.

RNA polymerase II (PFC0805w), a nucleosome assembly protein (NAP) (PFI0930c) and a Constans-like transcription factor (PF14_0383). These messengers and proteins should be analysed to determine if their differential expression could participate in the transcriptional regulation of varied sets of genes during the asexual parasite cycle.

4. Discussion

In order to determine differential transcript expression during erythrocytic development of *P. falciparum* clones 3D7 and its gametocyte-less derivative F12, we have used a gene-specific microarray printed with 153 PCR products

representing genes potentially implicated in cell cycle and transcriptional regulation. F12 clone has been isolated as a gametocyte non-producer from a long-term culture of the 3D7 clone. In contrast to 3D7 clone, F12 is impaired for early stages of gametocytogenesis without apparent difference in chromosomes number or size (Alano et al., 1995). Due to the genomic proximity of these two clones, one might expect that differential gene expression during the erythrocytic asexual cycle would participate in the early stage of gametocytogenesis including sensitivity to environment stimuli known to induce differentiation into gametocyte. Therefore, we designed a gene-specific microarray to study and compare temporal expression of roughly 150 messengers, during erythrocytic asexual development, of

highly synchronized cultures to identify some genes that might play a role in the parasite differentiation.

Before performing the time course transcriptome experiments we verified the clone properties, the quality of synchronization procedure and the feasibility of using F12 derived cDNA with a microarray encompassing 3D7 derived PCR product. First, we monitored expression of *pfgr27* transcript by real time quantitative RT-PCR and verified that no expression of the transcript could be detected in F12 in contrast to 3D7. Second, we carefully synchronized the parasite cultures, by three successive treatments with sorbitol at 44-h intervals. Subsequently, the five stages of erythrocytic asexual development were harvested with a minimum of 90% enrichment at each stage. Finally, we verified that all the PCR fragments obtained after amplification of the 3D7 genomic sequence and printed on the array could cross-hybridize with the genetic material prepared from F12 clone.

The radioactive labeled cDNAs were prepared from total RNA extracted at the indicated times of asexual life cycle. Formerly, the quality of the RNAs was controlled by Agilent analyzer and by fractionation on agarose gel. We performed radioactive-labeling of the cDNA probes, in order to minimize the background signal, avoid signal saturation and enhance signal detection especially for transcription factor mRNA that are believed to have a low abundance level. At the level of the microarray experimental protocol, we controlled the novel radioactive labeling and hybridization procedures carried out with the two probes (Salin et al., 2002; Refour et al., submitted for publication). Indeed, this was necessary to certify the reliability and reproducibility of this novel experimental approach.

Fig. 1 presents a representative picture corresponding to a co-hybridization with two cDNA labeled with ^{35}S dATP and ^3H dATP. Printing and hybridization were verified by the (i) appropriate responses given by the positive and negative controls, (ii) quality of the duplicates, and (iii) similarity of the signals given by the same PCR product (PF11_0098) spotted twice at different places in the array as well as that given by three different PCR products along the PF14_0316 ORF. In addition, each raw signal was normalized twice, first with the Alien cDNAs and second with the genomic *Plasmodium* DNA. Finally, some of the microarray data were verified using qPCR and compared with the results reported in the literature (Table 1).

The comparison of the steady state profiles of mRNA expression between the two clones, 3D7 and F12, led us to list 106 genes with essentially similar profiles and only eight genes with differential expression (Fig. 2). Additionally, we focused our attention on all the transcripts probably implicated in the transcriptional machinery and in particular those coding for transcription factors. Indeed, among the transcripts of the 15 putative transcription factors, most of them gave similar profiles throughout erythrocytic development of 3D7 and F12 clones, as shown for five of them verified by quantitative real time PCR (Fig. 3A and B). This

is not surprising since F12 was derived from 3D7. In addition, the conserved transcriptional profiles of several of these putative transcription regulators during erythrocytic cycle in 3D7, F12 and HB3 clones (Bozdech et al., 2003), favour their significant role for asexual growth and not for the impairment of sexual differentiation observed in the F12 clone.

All genes and in particular those encoding kinases and proteins of the transcriptional machinery displaying different expression profiles, were compared to results reported in the Scripps/GNF Malaria Array data (Le Roch et al., 2003) and the De Risi *P. falciparum* HB3 time course microarray data (Bozdech et al., 2003). Most of our data obtained with the 3D7 clone were in accordance with these two published experiments. However, some differences were observed that might be due to the clones features or/and to the cDNA preparation and labeling protocol (fluorescence and use of RNA amplification).

Even though most of the transcript profiles were similar between 3D7 and F12 clones, differential profiles were observed for genes potentially coding for kinases. First, two kinases, *MAL6P1.146* (haem-regulated inhibitory (HRI) kinase) and *PF13_0258* (a putative serine/threonine protein kinase) were clustered since expression was quite similar throughout 3D7 cycle in contrast to a maximal peak of expression occurring in the F12 clone, at LT for *MAL6P1.146* and ES for *PF13_0258* (Fig. 2A). Since *PF13_025* expression peaked in the schizont stage for 3D7 (Scripps/GNF Malaria Array data) and HB3 clones as well as for F12 clone, this gene does not appear to be essential for either asexual or sexual growth.

MAL6P1.146, also known as PfPK4, has been annotated as a protozoan eIF-2 α -related protein kinase due to its homology to HRI kinase, and reported to participate to the inhibition of protein synthesis, via the phosphorylation of the small subunit of the eIF-2 initiation factor (Clemens, 2001). Indeed, activity of the recombinant PfPK4 has been described to be inhibited by haemin (Mohrle et al., 1997). In erythrocytes infected by the K1 isolate, the level of the *MAL6P1.146* protein appeared quite similar during the parasite cycle even though its cellular distribution varied (Mohrle et al., 1997). Since PfPK4 accumulates in rhoptries, it probably plays a role in invasion and early development of the ring stage. In addition, erythrocytes contain small concentrations of free haemin and this level is increased in aged erythrocytes and pathological cells of sickle cell anaemia (Shaklai et al., 1985). The inhibition of PfPK4 activity by haemin might play a role in sensing internal erythrocyte environment. Since differential expression takes place between 3D7 and F12 clones and since enhanced gametocyte production occurs in young erythrocytes and in blood from patients with sickle cell anaemia (Trager and Gill, 1992), PfPK4 might be implicated in commitment of the young ring to gametocytogenesis by sensing internal erythrocyte environment. Over-expression of this gene in trophozoite might be at least in part

responsible for a decreased sensitivity to erythrocyte internal environment leading to its inability to respond to the commitment stimulus.

Second, *PFC0485w* encoding a putative Ser/Thr kinase, displayed opposite expression profiles (Fig. 2B) with a maximal expression for F12 in early schizonts, and a minimal for expression 3D7 clone. Interestingly, expression has also been shown to be increased in 3D7 gametocytes (Le Roch et al., 2003), and therefore, might be involved in gametocytogenesis providing an explanation for the gametocyte-less phenotype of F12 clone.

Finally, *PF07_0073*, a seryl-tRNA ligase, expression has been shown to be maximal in the early schizonts in 3D7 as also observed in HB3 in contrast to F12 where expression was maximal at early trophozoite stage (Fig. 2C) as seen for 3D7 (Scripps/GNF Malaria Array data, sorbitol synchronization). The expression profile of this gene may not be crucial for sexual development since it is different in the gametocyte producing clones.

In addition to kinases, several transcripts coding for proteins involved in transcriptional machinery showed differential profiles between the two clones even though most of the TF transcripts investigated presented similar expression during erythrocytic development.

The relative level of *PFB0865w* transcript, encoding a putative small nuclear ribonucleoprotein probably involved in RNA splicing, was higher than ring in 3D7 clone, in good agreement with the Scripps/GNF Malaria Array data, and continuously lower than ring in F12, as stated for the HB3 clone. Since HB3 clone is a gametocyte producer, this gene may not play a crucial role in controlling gametocyte production.

The maximal peak of expression of *PFC0805w* (a putative DNA-directed RNA polymerase II largest subunit) was reached at different stages of development, delayed in F12 when compared to 3D7. The delay observed in F12 clone may have very limited effects considering the crucial role of RNA polymerase II in transcription.

Finally, two putative transcriptional regulators *PFI0930c* and *PF14_0383*, showed distinct expression profiles in the two clones. Expression of *PFI0930c* transcript encoding a protein also known as PfB7, bearing a NAP motif, is not greatly altered throughout 3D7 parasite while it is markedly enhanced in F12 early schizonts. Proteins bearing a NAP motif have been shown to act as histone chaperones, shuttling both core and linker histones from their site of synthesis in the cytoplasm to the nucleus (Akey and Luger, 2003). These proteins may be involved in regulating gene expression by interaction either with chromatin remodelling factors therefore promoting transcriptional silencing (Singer et al., 1998) or with architectural factors such as HMGB proteins interacting with nucleosomes and promoting their sliding (Agregi and Bianchi, 2003). Interestingly, as we have annotated various HMGB proteins and since they are expressed during erythrocytic cycle (to be published), PfB7 might be

responsible for the modulation of the functional activity of these proteins. We have identified orthologs protein of *PFI0930c* in the four *Plasmodium* species accessible in PlasmoDB that showed a very high level of similarity (about 95% of the amino acid content), as stated before (Birago et al., 1996), enlightening the crucial role of this protein (data not shown). Moreover, evidence of two differently-sized mRNAs, encoding this protein, have been obtained in *P. falciparum* (Pace et al., 1998). In *P. berghei*, two different lengths of the 5' untranslated region (UTR) have been proposed to explain these two forms of mRNAs, one accumulating during gametocyte differentiation and the other specific for asexual growth. Since the two transcripts share an identical coding sequence, the PCR product used in this study detects either asexual or sexual specific transcript. Different species of *PfB7* mRNA have been hypothesized to be the results of two different promoters specific for asexual or sexual growth. Therefore, *PFI0930c* regulation has been demonstrated to be linked to gametocyte differentiation in *P. berghei* (Pace et al., 1998) and one can assume that it is also the case in *P. falciparum*. Differential expression in 3D7 and F12 clone may therefore be a consequence of differential expression of transcription factors and kinases implicated in the gametocyte differentiation. Indeed, chromatin remodelling may play a role in the impairment of F12 to produce gametocytes.

In the case of *PF14_0383* transcript coding for a Constans-like transcription factor, the maximal peak was obtained in the early schizont stage in 3D7 when it was observed during early trophozoite stage in F12 clone (Fig. 2C). The *PF14_0383* protein comprises two B-box motifs as seen in Constans (CO) proteins (Robson et al., 2001) and two nuclear localization signals (NLS) in tandem. Indeed, the B-box zinc finger motifs are present in a large number of proteins involved in transcription regulation, usually associated with RING finger and coil-coil motifs to form a tripartite motif, and if mutated, associated with human disease and cancer (Torok and Etkin, 2001). In plants, CO belongs to a family of transcription factors characterized by two zinc-finger B-boxes in tandem in their N-termini, believed to mediate protein-protein interaction (Robson et al., 2001; Samach et al., 2000). CO is involved in transcriptional regulation of flowering in *Arabidopsis thaliana* and promotes the expression of specific sets of genes (Samach et al., 2000). Again, we identified protein orthologs in the four *Plasmodium* species accessible in PlasmoDB (data not shown) showing conserved region located in the two B-boxes and in two other locations of the coding region including the two putative NLS. The first B-box matches the exact consensus (C x2 C x8 C x7 C x2 C x4 H x8 H) derived from CO B-box (Robson et al., 2001) where *x* represents any amino acid (aa). In addition, two other amino acids conserved in CO motif were shown, if mutated, to result in a flowering phenotype mutant of *A. thaliana* (Koornneef et al., 1991), and were conserved in the *PF14_0383* first B-box motif. The second B-box appeared

quite similar to the canonical B-box consensus [C x2 H x(4–9) C x2 C x4 C x2 H (C)] (Torok and Etkin, 2001). Bioinformatics analyses of the two B-boxes appeared to reveal a genuine CO-like transcription factors. Although its function as a transcription factor, has not yet been demonstrated, it possesses highly conserved motifs that have been involved in transcription regulation. Differential expression of this gene between the two clones may promote differential expression of sets of genes implicated in gametocytogenesis.

This study is the first attempt to analyse, by using a gene-specific microarray, the expression profile of messengers encoding proteins implicated in the cell cycle and transcription machinery, among which potential transcription factors not involved in the basal transcription but in the modulation of transcription level. Our results led us to identify, during the erythrocytic development of the 3D7 and F12 clones, transcription factors and kinases differentially expressed. They might in turn lead to differential expression of different sets of genes directly or indirectly implicated in the gametocyte differentiation. Nonetheless, true functional involvement in gametocytogenesis remains to be demonstrated. More work is needed and already engaged in order to decipher the potential involvement of each differentially expressed gene in F12 clone gametocyte-less phenotype.

Acknowledgements

We are indebted to Dr. Walliker and Dr. Alano for the kind gift of the *P. falciparum* clones. We are grateful for the microarray facilities provided by L. Galio and A. Doukani at P3S (Pitié-Salpêtrière). M.G., P.R. and C.B. were financially supported by the Ministère de l'Éducation Nationale, de la Recherche et de la Technologie and the PAL+ program. This work was supported by INSERM, France, to C.V. and D.M.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.gene.2004.07.004](https://doi.org/10.1016/j.gene.2004.07.004).

References

- Agresti, A., Bianchi, M.E., 2003. HMGB proteins and gene expression. *Curr. Opin. Genet. Dev.* 13, 170–178.
- Akey, C.W., Luger, K., 2003. Histone chaperones and nucleosome assembly. *Curr. Opin. Struck. Biol.* 13, 6–14.
- Alano, P., Roca, L., Smith, D., Read, D., Carter, R., Day, K., 1995. *Plasmodium falciparum*: parasites defective in early stages of gametocytogenesis. *Exp. Parasitol.* 81, 227–235.
- Alano, P., Silvestrini, F., Roca, L., 1996. Structure and polymorphism of the upstream region of the pfg27/25 gene, transcriptionally regulated in gametocytogenesis of *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* 79, 207–217.
- Birago, C., Pace, T., Barca, S., Picci, L., Ponzi, M., 1996. A chromatin-associated protein is encoded in a genomic region highly conserved in the *Plasmodium* genus. *Mol. Biochem. Parasitol.* 80, 193–202.
- Bozdech, Z., Llinas, M., Pulliam, B.L., Wong, E.D., Zhu, J., De Risi, J.L., 2003. The transcriptome of the intraerythrocytic developmental cycle of *Plasmodium falciparum*. *PLoS Biol.* 1, 5.
- Clemens, M.J., 2001. Initiation factor eIF2 alpha phosphorylation in stress responses and apoptosis. *Prog. Mol. Subcell. Biol.* 27, 57–89.
- Day, K.P., Karamalis, F., Thompson, J., Barnes, D.A., Peterson, C., Brown, H., Brown, G.V., Kemp, D.J., 1993. Genes necessary for expression of a virulence determinant and for transmission of *Plasmodium falciparum* are located on a 0.3-megabase region of chromosome 9. *Proc. Natl. Acad. Sci. U. S. A.* 90, 8292–8296.
- Dyer, M., Day, K., 2000a. Expression of *Plasmodium falciparum* trimeric G proteins and their involvement in switching to sexual development. *Mol. Biochem. Parasitol.* 110, 437–448.
- Dyer, M., Day, K.P., 2000b. Commitment to gametocytogenesis in *Plasmodium falciparum*. *Parasitol. Today* 16, 102–107.
- Gardner, M.J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R.W., Carlton, J.M., Pain, A., Nelson, K.E., Bowman, S., Paulsen, I.T., James, K., Eisen, J.A., Rutherford, K., Salzberg, S.L., Craig, A., Kyes, S., Chan, M.S., Nene, V., Shallom, S.J., Suh, B., Peterson, J., Angiuoli, S., Pertea, M., Allen, J., Selengut, J., Haft, D., Mather, M.W., Vaidya, A.B., Martin, D.M., Fairlamb, A.H., Fraunholz, M.J., Roos, D.S., Ralph, S.A., McFadden, G.I., Cummings, L.M., Subramanian, G.M., Mungall, C., Venter, J.C., Carucci, D.J., Hoffman, S.L., Newbold, C., Davis, R.W., Fraser, C.M., Barrell, B., 2002. Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* 419, 498–511.
- Horrocks, P., Jackson, M., Cheesman, S., White, J.H., Kilbey, B.J., 1996. Stage specific expression of proliferating cell nuclear antigen and DNA polymerase delta from *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* 79, 177–182.
- Inselburg, J., 1983. Stage-specific inhibitory effect of cyclic AMP on asexual maturation and gametocyte formation of *Plasmodium falciparum*. *J. Parasitol.* 69, 592–597.
- Koornneef, M., Hanhart, C.J., van der Veen, J.H., 1991. A genetic and physiological analysis of late flowering mutants in *Arabidopsis thaliana*. *Mol. Gen. Genet.* 229, 57–66.
- La Greca, N., Hibbs, A.R., Riffkin, C., Foley, M., Tilley, L., 1997. Identification of an endoplasmic reticulum-resident calcium-binding protein with multiple EF-hand motifs in asexual stages of *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* 89, 283–293.
- Lanzer, M., de Bruin, D., Ravetch, J.V., 1992a. A sequence element associated with the *Plasmodium falciparum* KAHRP gene is the site of developmentally regulated protein–DNA interactions. *Nucleic Acids Res.* 20, 3051–3056.
- Lanzer, M., de Bruin, D., Ravetch, J.V., 1992b. Transcription mapping of a 100 kb locus of *Plasmodium falciparum* identifies an intergenic region in which transcription terminates and reinitiates. *EMBO J.* 11, 1949–1955.
- Le Roch, K.G., Zhou, Y., Blair, P.L., Grainger, M., Moch, J.K., Haynes, J.D., De La Vega, P., Holder, A.A., Batalov, S., Carucci, D.J., Winzeler, E.A., 2003. Discovery of gene function by expression profiling of the malaria parasite life cycle. *Science*.
- Livak, K.J., Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2-(Delta Delta C(T)) Method. *Methods* 25, 402–408.
- Lobo, C.A., Kumar, N., 1999. Differential transcription of histone genes in asexual and sexual stages of *Plasmodium falciparum*. *Int. J. Parasitol.* 29, 1447–1449.
- Lobo, C.A., Konings, R.N., Kumar, N., 1994. Expression of early gametocyte-stage antigens Pfg27 and Pfs16 in synchronized game-

- toocytes and non-gametocyte producing clones of *Plasmodium falciparum*. Mol. Biochem. Parasitol. 68, 151–154.
- Lobo, C.A., Fujioka, H., Aikawa, M., Kumar, N., 1999. Disruption of the Pfg27 locus by homologous recombination leads to loss of the sexual phenotype in *P. falciparum*. Mol. Cell. 3, 793–798.
- Mohrle, J.J., Zhao, Y., Wemli, B., Franklin, R.M., Kappes, B., 1997. Molecular cloning, characterization and localization of PfPK4, an eIF-2alpha kinase-related enzyme from the malarial parasite *Plasmodium falciparum*. Biochem. J. 328 (Pt. 2), 677–687.
- Ono, T., Nakabayashi, T., 1990. Gametocytogenesis induction by ammonium compounds in cultured *Plasmodium falciparum*. Int. J. Parasitol. 20, 615–618.
- Pace, T., Birago, C., Janse, C.J., Picci, L., Ponzi, M., 1998. Developmental regulation of a *Plasmodium* gene involves the generation of stage-specific 5' untranslated sequences. Mol. Biochem. Parasitol. 97, 45–53.
- Refour, P., Siau, A., Gissot, M., Briquet, S., Boschet, C., Mazier, D., Vaquero, C., 2004. High-accuracy of nanogram total RNA amounts for gene profiling without amplification (submitted for publication).
- Robson, F., Costa, M.M., Hepworth, S.R., Vizir, I., Pineiro, M., Reeves, P.H., Putterill, J., Coupland, G., 2001. Functional importance of conserved domains in the flowering-time gene CONSTANS demonstrated by analysis of mutant alleles and transgenic plants. Plant J. 28, 619–631.
- Salin, H., Vujasinovic, T., Mazurie, A., Maitrejean, S., Menini, C., Mallet, J., Dumas, S., 2002. A novel sensitive microarray approach for differential screening using probes labelled with two different radioelements. Nucleic Acids Res. 30, e17.
- Samach, A., Onouchi, H., Gold, S.E., Ditta, G.S., Schwarz-Sommer, Z., Yanofsky, M.F., Coupland, G., 2000. Distinct roles of CONSTANS target genes in reproductive development of *Arabidopsis*. Science 288, 1613–1616.
- Schena, M., Shalon, D., Davis, R.W., Brown, P.O., 1995. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science 270, 467–470.
- Shaklai, N., Shviro, Y., Rabizadeh, E., Kirschner-Zilber, I., 1985. Accumulation and drainage of hemin in the red cell membrane. Biochim. Biophys. Acta 821, 355–366.
- Sharma, A., Sharma, I., Kogkasuriyachai, D., Kumar, N., 2003. Structure of a gametocyte protein essential for sexual development in *Plasmodium falciparum*. Nat. Struct. Biol. 10, 197–203.
- Singer, M.S., Kahana, A., Wolf, A.J., Meisinger, L.L., Peterson, S.E., Goggin, C., Mahowald, M., Gottschling, D.E., 1998. Identification of high-copy disruptors of telomeric silencing in *Saccharomyces cerevisiae*. Genetics 150, 613–632.
- Torok, M., Etkin, L.D., 2001. Two B or not two B? Overview of the rapidly expanding B-box family of proteins. Differentiation 67, 63–71.
- Trager, W., Gill, G.S., 1992. Enhanced gametocyte formation in young erythrocytes by *Plasmodium falciparum* in vitro. J. Protozool. 39, 429–432.
- Trager, W., Jensen, J.B., 1976. Human malaria parasites in continuous culture. Science 193, 673–675.

2. Caractérisation fonctionnelle et rôle de PfMyb1, un facteur de transcription spécifique de séquence, dans la régulation des gènes au cours du cycle érythrocytaire

2.1 Caractérisation fonctionnelle de PfMyb1

2.1.1 Résumé

Etant donné le faible nombre d'acteurs de la régulation de la transcription décrits chez *P. falciparum*, l'annotation et la caractérisation de facteurs de transcription spécifiques de séquence semblaient être une étape indispensable dans l'étude de la régulation de l'expression des gènes chez *P. falciparum*. Les analyses du génome effectuées au sein du groupe, avant qu'il ne soit complètement séquencé, ont montré la présence de FT. Ces protéines, identifiées par des techniques bioinformatiques, ont cependant besoin de la validation de leur fonction biologique avant leur annotation définitive. Les FT sont définis et classés selon leur domaine de liaison à l'ADN. Après l'identification dans le génome de *P. falciparum* d'un gène codant pour la protéine PfMyb1 par homologie de séquence avec des FT Myb eucaryotes, une analyse bioinformatique plus fine a permis de le classer dans la classe des FT à domaines tryptophane de la famille des protéines Myb (Figure 1, article 2, p.88). Une étude de sa séquence second-tertiaire, a confirmé que cette protéine possédait toutes les caractéristiques d'un FT de la famille des Myb. Enfin, nous avons tenté de réaliser la validation biologique de la fonction de PfMyb1 en tant que FT de la famille des Myb.

La fonction d'un FT spécifique de séquence se définit essentiellement par sa capacité à lier l'ADN de manière spécifique. Sa caractérisation fonctionnelle passe donc par l'analyse de sa liaison à l'ADN. Les facteurs de la famille Myb sont connus pour lier spécifiquement une séquence d'ADN consensus appelée Myb Related Element (MRE) (Figure 1, article 2, p.88).

Lors de cette étude, il a été montré qu'une ou plusieurs protéines contenues dans les extraits nucléaires étaient capables de se lier spécifiquement à plusieurs séquences d'ADN contenant un MRE (Figure 2, article 2, p.89). Ces motifs de régulation provenaient de deux séquences consensus retrouvées dans deux promoteurs de *P. falciparum* et d'une séquence provenant d'un promoteur du gène *mim-1* de poulet connue pour interagir avec des protéines de la famille Myb. Cette interaction, même si elle semble complexe, est dépendante de la présence

de la protéine PfMyb1 puisqu'elle est chassée par des anticorps spécifiques. Cette protéine possède donc tous les attributs d'un FT spécifique de séquence de la famille Myb.

Afin d'évaluer son éventuelle implication dans l'initiation de la gamétocytogénèse, nous avons comparé la cinétique d'interaction de la protéine avec un MRE au cours du cycle érythrocytaire de deux clones de *P. falciparum*, un producteur de gamétocytes (3D7) et l'autre incapable d'en produire (F12). Ces deux cinétiques sont différentes au cours du cycle érythrocytaire. En effet, ces expériences ont révélé une absence de complexes entre le MRE et les extraits nucléaires d'anneaux et de schizontes du clone F12.

Dans des expériences complémentaires, l'expression du transcrit *pfmyb1* et la cinétique d'expression de la protéine PfMyb1 ont été caractérisés dans les deux clones (Figure 2 et 3, résultats complémentaires, p.94 et 96). Ces études ont montré que l'expression du transcrit semble similaire dans les deux clones. Les premiers résultats obtenus indiquent une grande corrélation de la cinétique d'expression de la protéine PfMyb1 pour les deux clones.

Article 2 :

Characterization of PfMyb1 transcription factor during erythrocytic development of 3D7 and F12 *Plasmodium falciparum* clones.

Charlotte Boschet *, Mathieu Gissot*, Sylvie Briquet, Zuhail Hamid, Clotilde Claudel-Renard et Catherine Vaquero.

* : les deux auteurs ont contribué également à ce travail.

Accepté à la publication dans *Molecular and Biochemical Parasitology*.

Short communication

Characterization of PfMyb1 transcription factor during erythrocytic development of 3D7 and F12 *Plasmodium falciparum* clones[☆]

Charlotte Boschet^{a,1}, Mathieu Gissot^{a,1}, Sylvie Briquet^a, Zuhail Hamid^{a,2},
Clotilde Claudel-Renard^b, Catherine Vaquero^{a,*}

^a INSERM U511, CHU Pitié-Salpêtrière, 91 boulevard de l'Hôpital, 75013 Paris, France

^b Department of Biochemistry, University of Pretoria, Pretoria, South Africa

Received 8 May 2004; received in revised form 19 July 2004; accepted 20 July 2004

Keywords: *Plasmodium*; Transcription factor; Expression; Erythrocyte; Development

Two major events of the *Plasmodium falciparum* life cycle occur in the erythrocyte: an asexual multiplication responsible for clinical manifestations and initiation of sexual differentiation responsible for dissemination of the disease via mosquito. These developmental pathways require the coordinated and modulated expression of different sets of genes, involving transcriptional controls, even though post-transcriptional regulation could also be implicated. These gene clusters are therefore likely to share regulatory elements or modules within their promoters [1]. Regulation of transcription is also accomplished by the availability of transcription factors, which is presumably modulated throughout parasite development.

In *Plasmodium* very little is known at the level of *cis*- and *trans*-regulatory elements, and to our knowledge only a few regulatory elements and no clearly defined transcription factors have been reported in the literature. Nevertheless, *Plasmodium* transcriptional machinery appeared to share com-

mon features with that of eukaryotes [2]. Moreover, regulatory sequences within the promoters were reported to contain motifs in common with the binding sites of eukaryotic transcription factors [2–5] and our own computer analyses confirmed this statement, including DNA motifs potentially interacting with Myb factors (Boschet and Vaquero, to be published elsewhere).

Herein, we present the first description of a *Plasmodium* transcription factor: a Myb-related protein of the tryptophan cluster family. The Myb proteins are highly conserved in eukaryotes and generally, their characteristic DNA-binding domain (DBD) contains three tandem repeats (R1, R2 and R3) of approximately 50 residues with 3 regularly spaced (18 or 19 amino acids) tryptophan residues [6,7]. They have been shown to bind DNA in a sequence-specific manner and to participate in the regulation of the expression of genes implicated in growth control and differentiation [7]. To identify Myb proteins in the *P. falciparum* genome, more than 200 non-redundant eukaryotic Myb proteins were aligned with MultAlin and ClustalW [8,9] to generate a consensus sequence corresponding to the characteristic DNA-binding domain. This consensus was used as query for the *Plasmodium* database and allowed the annotation of a 414 amino acid-long ORF: PfMyb1. Since PFSCAN [10] identified only one Myb domain (R2 in Fig. 1A), the complete sequence of PfMyb1 was aligned with the DBD of three proteins, DdMybH, DdMyb2 and DdMyb3 [11–13] of the slime mold *Dictyostelium discoideum*, which has an A + T-rich genome (78% overall) like *Plasmodium* (85%). This resulted in the

Abbreviations: aa, amino-acid; DBD, DNA-binding domain; EMSA, electrophoretic mobility shift assay; MRE, Myb regulatory element; NE, nuclear extract; ORF, open reading frame; PCR, polymerase chain reaction; RT, reverse transcription

[☆] Note: EMBL accession number AJ291747.

* Corresponding author. Tel.: +33 1 40778111; fax: +33 1 45838858.

E-mail address: vaquero@ext.jussieu.fr (C. Vaquero).

¹ The authors wish it to be known that, in their opinion, Charlotte Boschet and Mathieu Gissot should be regarded as joint first authors.

² Permanent address: INMO, University of Gezira, Wad Madani, Box 20, Sudan.

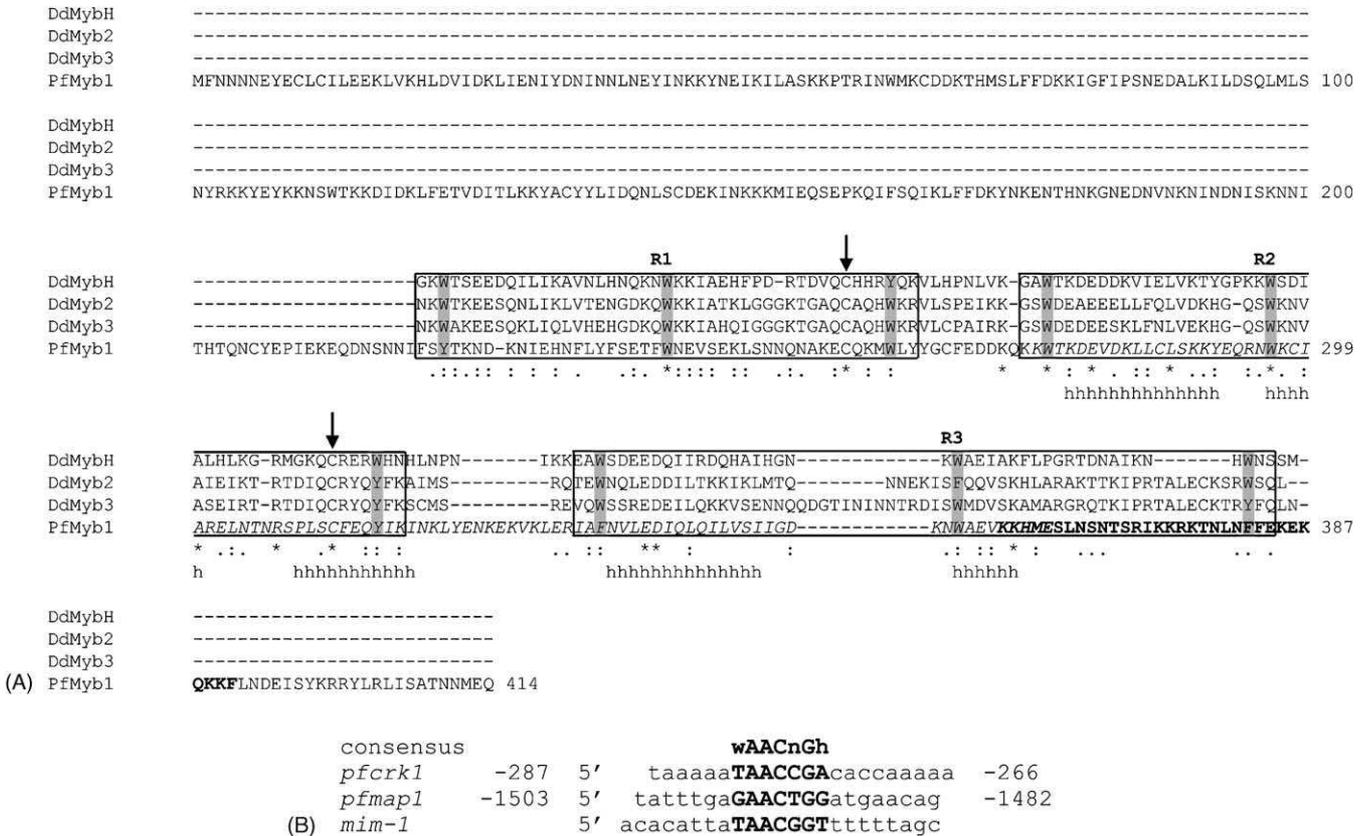


Fig. 1. The DNA-binding domain of the PfMyb1 sequence and three different Myb regulatory elements. (A) The protein consensus of the characteristic DNA-binding domain of the Myb family was obtained with an alignment of eukaryotic Myb selected throughout evolution and allowed the annotation PfMyb1 within the *Plasmodium* chromosome 13. The complete sequence of PfMyb1 is shown as well as an optimal alignment between its putative DNA-binding domain and the three DBD of *Dictyostelium discoideum* (DdMybH, DdMyb2 and DdMyb3; SwissProt accession numbers: P34127, O15816, Q9GUB3, respectively). Identities are indicated by asterisks, strong similarities by two points and weak similarities by one point. The residues W, Y or F that characterize the Myb domains are highlighted in gray and R1, R2 and R3 are boxed. The highly conserved cysteine residues are marked by an arrow. The sequence from Lys 275 to Glu 364 threaded by Meta-Server and TITO [21,35] is indicated in italics and aa predicted to take part in α -helices are marked with an h. A putative nuclear localization sequence, encompassing the unusually long C-terminal part of the third repeat, is shown in bold letters. (B) Putative MRE were localized among others in *Plasmodium pferkl* and *pfmap1* gene promoters using MatInspector [36] according to three matrix families (one from the plant and two from the vertebrate subsections). Core and matrix similarity parameters were left by default. These two potential MRE were aligned with a prototype MRE found in the chicken *mim-1* gene promoter reported to be functional [26] and with the consensus sequence of the Myb binding domain. Letters in upper-case represent the MRE and in lower-case the flanking sequences. The positions are numbered from the first ATG of the open reading frames. W: A or T; H: A, C or T; N: A, C, G or T.

identification of three Myb domains (Fig. 1A), located in the C-terminus of the protein as in DdMyb2 and DdMyb3 of *D. discoideum*, and these two factors were reported to be functional [12,13] whereas in most of the Myb proteins reported so far, the DBD is located in the N-terminus. However, PfMyb1 contains imperfect repeats with a tyrosine or a phenylalanine in place of tryptophan and with an aa insertion in R3 as observed in *D. discoideum* (Fig. 1A) as well as in *Saccharomyces cerevisiae* Bas1 [14,15] and *Kluyveromyces lactis* Reb1 [16]. Moreover, a critical cysteine residue, as highly conserved as the tryptophan residues, [17] was also found in R1 and R2 at position 258 and 312, respectively, and was reported to play a role in redox regulation [18–20]. Furthermore, 3D-structure was determined with Meta-Server [21] and it appeared that the sequence from Lys 275 to Glu 364 (Fig. 1A), encompassing the R2 domain and the first part of

the R3 domain, could be threaded by homology with the human Myb proto-oncogene protein, whose structure has been determined by X-ray diffraction [22]. Indeed, three α -helices were predicted in R2 and only two in R3 in contrast to the R1 repeat. Nevertheless, in Myb proteins, R2 and R3 constitute a minimal DBD sufficient for sequence-specific DNA binding, while R1 does not appear to engage in specific interactions with DNA [23]. In summary, the diverse computational analyses of PfMyb1 provide evidence for a genuine Myb protein, which is strikingly conserved in all *Plasmodium* species listed in PlasmoDB.

After annotation of the putative transcription factor PfMyb1, the characterization of its cognate RNA transcript was carried out by Northern blotting experiment. A messenger of 2.6 kb was evidenced (data not shown) in *P. falciparum* infected erythrocytes. Our aim was to analyze the differences

occurring at the level of transcription factors that might control expression of transcripts and/or proteins involved in sexual differentiation, a process crucial for the dissemination of the disease via the mosquito. Using a semi-quantitative RT-PCR, the temporal expression of *pfmyb1* transcript (data not shown) was analyzed at different times of erythrocytic development of highly synchronized parasites (rings, early and late trophozoites, as well as early and late schizonts). The profile of expression was determined in two clones of *P. falciparum*, 3D7 and the gametocyte-less F12 derived from 3D7. The major difference in the mRNA profiles resides in a lower abundance of the *pfmyb1* transcript in the ring stage of F12 when compared to 3D7, followed by a sharp increase in early F12 trophozoites. In contrast, the mRNA level was

quite similar in both clones in the early and late schizonts after a maximal expression in trophozoites. The 3D7 expression profile is in good agreement with the published transcriptome data of HB3 [24] and 3D7 clones [25].

We then analyzed the presence of the transcription factor in the nuclear extracts of the parasite. The *pfmyb1* ORF has been cloned and the His-tagged recombinant protein was expressed, purified and used as a positive control in Western blots. The PfMyb1 protein was shown to be present in the parasite trophozoite nuclear extracts by a serial immunoprecipitation and Western blotting experiments (Fig. 2A). Briefly, nuclear extracts were immunoprecipitated, using tosylactivated magnetic Dynabeads coated with either anti-PfMyb1 serum (lane 2) or preimmune (lane 3). Immunoprecipitated

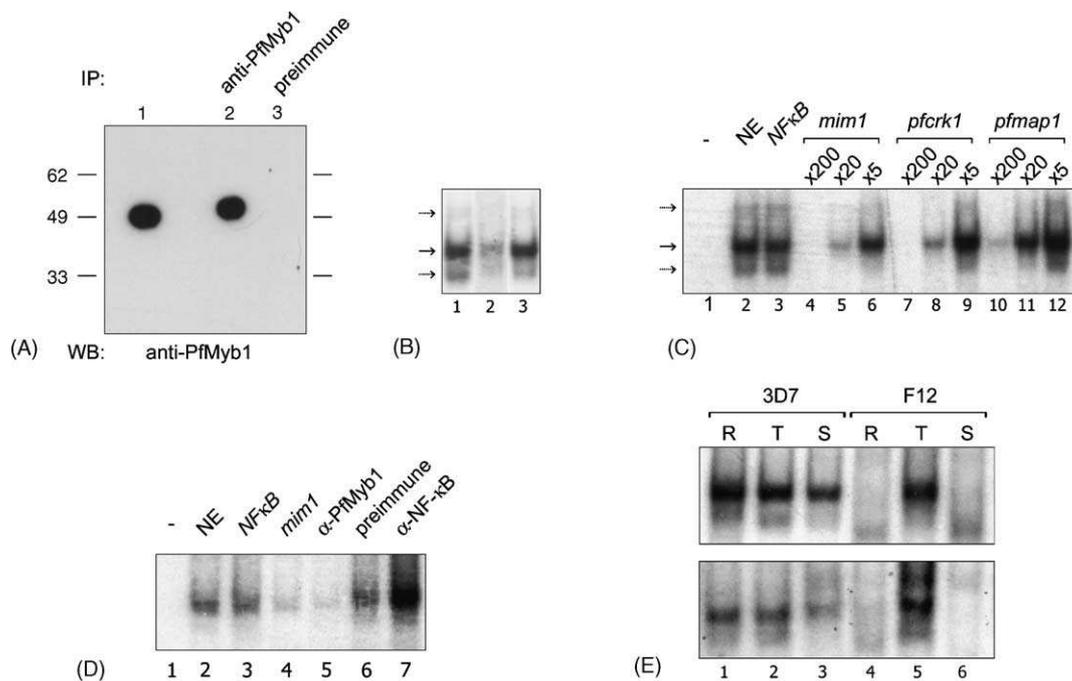


Fig. 2. Functional study of PfMyb1 in *P. falciparum* nuclear extracts. (A) Serial immunoprecipitation and Western blotting experiments. *Plasmodium falciparum* protein extracts were prepared as described in [37]. A polyclonal antipeptide serum was prepared by Eurogentec (Belgium). Two rabbits were immunized with two peptides (RKKYEYKKNNSWTKK and KRRYLRLISATNMEQ), corresponding to amino acids (aa) 103–117 (upstream to the R1 domain) and 400–414 (encompassing the very last C-terminus aa of R3) of the PfMyb1 amino acid sequence, respectively. One hundred micrograms proteins of *P. falciparum* trophozoites were subjected to immunoprecipitation using 2×10^7 magnetic Dynabeads tosylactivated M-280 (DynaL Biotech) coated with 2.5 μ g of anti-PfMyb1 serum (lane 2) or preimmune (lane 3) following manufacturer recommendations. Lane 1 stands for PfMyb1 recombinant protein. Immunoprecipitated proteins were run on a 10% SDS-PAGE and subjected, after transfer onto nitrocellulose membrane, to Western blotting using the anti-PfMyb1 serum and revealed by a peroxidase-conjugated anti-rabbit antibody (Sigma). (B–E) EMSA. Nuclear extracts (NE) from 50 mL of red blood cells at 10%–12% parasitaemia infected with rings (1–18 h), trophozoites (20–28 h) and schizonts (32–42 h post-invasion) were prepared as described by Osta et al. [37]. EMSA experiments were conducted with nuclear extracts of 3D7 and F12 clones for 30 min at 4 °C in binding buffer (25 mM HEPES pH 7.5, 50 mM KCl, 1 mM MgCl₂, 0.5 mM EDTA, 0.2 mM PMSF, 1 mM DTT, 5% glycerol and 2 μ g of poly dI-dC). Addition of 50,000 cpm of labeled probes is followed in a final volume of 10 μ l for an additional 30 min. Samples were loaded on pre-run 6% non-denaturing polyacrylamide gel and electrophoresed at 160 V for 2 h at 4 °C in a low-ionic strength TBE buffer (0.25 \times). (B) Double stranded oligonucleotides (100 ng) of the prototype and putative *Plasmodium* MRE (Fig. 1B) were end-labeled with 40 μ Ci of [γ -³²P] ATP and purified with QIAquick nucleotide removal kit (Qiagen). One micrograms of 3D7 trophozoite nuclear extract was incubated with labeled chicken *mim-1* prototype, *pfmap1* and *pferk1* putative *Plasmodium* MRE (lanes 1–3). A bold arrow shows the major complex and dotted arrows indicate two minor complexes. (C) Competitions were carried out as in (B) only with the labeled chicken *mim-1* MRE prototype in presence of a 200-fold molar excess of unlabeled non-specific *NF- κ B* (lane 3) and a 200-, 20- and 5-fold excess of specific *mim-1* (lanes 4–6), *pferk1* (lanes 7–9) and *pfmap1* (lanes 10–12) oligonucleotides. Controls without and with NE alone are presented in lanes 1 and 2, respectively. (D) Interaction between 3D7 nuclear extract and *mim-1* was assayed in the presence of either 0.5 μ l of anti-PfMyb1 polyclonal antibody (0.6 μ g/ μ l, lane 5), or 0.5 μ l of the corresponding preimmune serum (lane 6), or 0.5 μ l of an irrelevant anti-NF- κ B antibody (lane 7). Lanes 1–4 are as in (C). (E) Labeled *mim-1* oligonucleotide was incubated with 1 μ g of nuclear extracts of 3D7 and F12 rings (R, lanes 1 and 4), trophozoites (T, lanes 2 and 5) and schizonts (S, lanes 3 and 6). Two independent experiments are presented.

proteins were subjected to SDS-PAGE and Western blotting with the anti-PfMyb1 serum. A single band, whose molecular mass (50 kDa) is slightly higher than the recombinant PfMyb1 (lane 1) was observed (lane 2) in the parasite nuclear extracts. In contrast, when using preimmune coated beads (lane 3) no protein was revealed by the anti-PfMyb1 serum.

In order to activate or repress transcription of particular sets of genes, the transcription factors have to bind *cis*-regulatory DNA elements in a sequence-specific manner. A Myb-related protein was evidenced in the 3D7 trophozoite nuclear extracts (NE) since specific interactions were detected with diverse Myb-regulatory elements (Fig. 1B), such as a prototype MRE (*mim-1* gene promoter) [26] and two putative MRE naturally occurring in the promoters of *Plasmodium* *pfmap1* and *pfcrk1* genes. These genes were originally reported to be expressed preferentially during erythrocytic asexual and sexual stages, respectively [27,28]. The EMSA profiles (Fig. 2B) appeared quite similar for the three MRE assayed, all giving rise to one major and two minor complexes, even though the magnitude of interaction was different, with its highest level being observed for *mim-1* and *pfcrk1* MRE, in line with the nucleotide sequence of the elements (Fig. 1B). These retarded bands could result from homo-dimerization of the PfMyb1 protein, as well as from interaction with other proteins, which have been shown to interact with the Myb proteins, as the bZip transcription factor C/EBP β [22] and the poly(ADP-ribose) polymerase (PARP) [29]. The DNA/protein interaction was specific, as shown by competition experiments with unlabeled oligonucleotides corresponding to each MRE (Fig. 2C). The interaction was not competed out by a 200-fold excess of an irrelevant unlabeled NF- κ B element present in the *IL-2R α* promoter [30] (lane 3), in contrast to a 200-fold excess of unlabeled relevant *mim-1*, *pfcrk1* and *pfmap1* (lanes 4, 7 and 10, respectively). In Fig. 2D, the specificity of the interaction was verified by using the anti-PfMyb1 antibody (lane 5), which competed out interaction with the labeled probe as efficiently as the *mim-1* competitor (lane 4), in contrast to the pre-immune serum (lane 6) and an irrelevant anti-NF- κ B antibody [31] (lane 7). Indeed, anti-PfMyb1 antibody, raised against the C-terminus of the DBD, impaired complex formation in place of generating a super-shifted band as already reported for other transcription factors [30,32]. For all these reasons we think that PfMyb1 is a genuine transcription factor. However, it was not possible, under the experimental conditions assayed so far, to demonstrate any interaction between these MRE and the recombinant protein, whether complete or incomplete, encompassing the three Myb domains. This might be due to inappropriate folding, processing, and/or phosphorylation of the recombinant protein, as has already been reported for other recombinant *P. falciparum* proteins expressed in *E. coli* [33].

Finally, we compared the temporal expression of PfMyb1 protein during 3D7 and F12 erythrocytic development with nuclear extracts prepared from the three main forms: rings, trophozoites and schizonts. EMSA experiments using the

three nuclear extracts and labeled *mim-1* oligonucleotide are shown in Fig. 2E and correspond to two representative and independent biological experiments. The major complex was present throughout development of the 3D7 clone, decreasing only slightly up to the schizont stage (lanes 1–3). In contrast, for the F12 clone the major complex, while quite apparent in trophozoites (lane 5), was almost undetectable in rings and schizonts (lanes 4 and 6). A discrepancy between transcript and EMSA profiles was quite apparent, especially in the F12 clone. The messenger was observed in F12 rings, despite absence of Myb/MRE interaction. Moreover, in early and late schizonts of both clones, the *pfmyb1* transcript was present to equal extents, while the DNA/protein interaction was detectable in 3D7 and undetectable in F12 nuclear extracts. The absence of interaction was not due to the degradation or lower concentration of the nuclear extracts used in the EMSA, since the integrity and protein content of all nuclear preparations were verified by SDS-PAGE analysis and Bradford assay (data not shown). Decreased mRNA translation [34] or absence of expression of a co-factor responsible for the formation of EMSA complexes, among other explanations, could account for the absence of retarded band. We are not, however, as yet able to address these questions. Nevertheless, the lack of PfMyb1 or partners in rings and schizonts might participate to the impairment of sexual differentiation towards gametocytes observed in F12 clone.

In summary, by computational annotation we identified, within the genomic DNA of *P. falciparum*, an open reading frame sharing common features with the Myb transcription factors. We assumed that this factor could regulate expression of particular genes since appropriate binding sites were identified within their promoters. In nuclear extracts, a Myb DNA binding activity was observed with a prototype and two putative *Plasmodium* Myb regulatory elements. This interaction was demonstrated to be specific, since it was inhibited by specific competitors and anti-PfMyb1 antibody in bandshift assays. Finally, during erythrocytic development, the bandshift profiles were markedly different in the 3D7 and the gametocyte-less F12 clones, in contrast to the transcript profile. These experiments provide the first evidence of the presence, in *Plasmodium*, of a transcription factor belonging to the highly conserved Myb family that might be implicated in the regulation of gene expression.

Acknowledgements

We thank Pietro Alano for his gift of F12 clone, Leila Gannoun (UMR-CNRS 5539) for the nuclear extract protocol and Alain Israël for the anti-NF- κ B antibody. We are grateful to Jennifer Richardson for critical reading of the manuscript, and to Professors M.M. Magzoub (Minister of Higher Education and Scientific Research, Sudan) and Dominique Mazier, as well as Philippe Refour (INSERM U511), for continuous encouragement and support. C.B. and M.G. were financially supported by the Ministère de l'Éducation Nationale, France,

and Z.H. by the WHO Mediterranean Office and the Ministry of Higher Education and Scientific Research, Sudan. The project was supported by INSERM funds to C.V.

References

- [1] Arnone MI, Davidson EH. The hardwiring of development: organization and function of genomic regulatory systems. *Development* 1997;124:1851–64.
- [2] Horrocks P, Decherig K, Lanzer M. Control of gene expression in *Plasmodium falciparum*. *Mol Biochem Parasitol* 1998;95:171–81.
- [3] Lanzer M, de Bruin D, Ravetch JV. Transcription mapping of a 100 kb locus of *Plasmodium falciparum* identifies an intergenic region in which transcription terminates and reinitiates. *EMBO J* 1992;11:1949–55.
- [4] Lanzer M, Wertheimer SP, de Bruin D, Ravetch JV. *Plasmodium*: control of gene expression in malaria parasites. *Exp Parasitol* 1993;77:121–8.
- [5] Horrocks P, Kilbey BJ. Physical and functional mapping of the transcriptional start sites of *Plasmodium falciparum* proliferating cell nuclear antigen. *Mol Biochem Parasitol* 1996;82:207–15.
- [6] Kanei-Ishii C, Nomura T, Ogata K, et al. Structure and function of the proteins encoded by the *myb* gene family. *Curr Top Microbiol Immunol* 1996;211:89–98.
- [7] Lipsick JS. One billion years of Myb. *Oncogene* 1996;13:223–35.
- [8] Corpet F. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res* 1988;16:10881–90.
- [9] Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 1994;22:4673–80.
- [10] PFSCAN, <http://hits.isb-sib.ch/cgi-bin/PFSCAN>.
- [11] Stober-Grasser U, Brydolf B, Bin X, Grasser F, Firtel RA, Lipsick JS. The Myb DNA-binding domain is highly conserved in *Dictyostelium discoideum*. *Oncogene* 1992;7:589–96.
- [12] Otsuka H, Van Haastert PJ. A novel Myb homolog initiates *Dictyostelium* development by induction of adenylyl cyclase expression. *Genes Dev* 1998;12:1738–48.
- [13] Guo K, Anjard C, Harwood A, Kim HJ, Newell PC, Gross JD. A myb-related protein required for culmination in *Dictyostelium*. *Development* 1999;126:2813–22.
- [14] Tice-Baldwin K, Fink GR, Arndt KT. BAS1 has a Myb motif and activates HIS4 transcription only in combination with BAS2. *Science* 1989;246:931–5.
- [15] Hovring I, Bostad A, Ording E, Myrset AH, Gabrielsen OS. DNA-binding domain and recognition sequence of the yeast BAS1 protein, a divergent member of the Myb family of transcription factors. *J Biol Chem* 1994;269:17663–9.
- [16] Morrow BE, Ju Q, Warner JR. A bipartite DNA-binding domain in yeast Reb1p. *Mol Cell Biol* 1993;13:1173–82.
- [17] Pinson B, Brendeford EM, Gabrielsen OS, Daignan-Fornier B. Highly conserved features of DNA binding between two divergent members of the myb family of transcription factors. *Nucleic Acids Res* 2001;29:527–35.
- [18] Myrset AH, Bostad A, Jamin N, Lirsac PN, Toma F, Gabrielsen OS. DNA and redox state induced conformational changes in the DNA-binding domain of the Myb oncoprotein. *EMBO J* 1993;12:4625–33.
- [19] Guehmann S, Vorbrueggen G, Kalkbrenner F, Moelling K. Reduction of a conserved Cys is essential for Myb DNA-binding. *Nucleic Acids Res* 1992;20:2279–86.
- [20] Grasser FA, LaMontagne K, Whittaker L, Stohr S, Lipsick JS. A highly conserved cysteine in the v-Myb DNA-binding domain is essential for transformation and transcriptional trans-activation. *Oncogene* 1992;7:1005–9.
- [21] Douguet D, Labesse G. Easier threading through web-based comparisons and cross-validations. *Bioinformatics* 2001;17:752–3.
- [22] Tahirov TH, Sato K, Ichikawa-Iwata E, et al. Mechanism of c-Myb-C/EBP beta cooperation from separated sites on a promoter. *Cell* 2002;108:57–70.
- [23] Ogata K, Morikawa S, Nakamura H, et al. Solution structure of a specific DNA complex of the Myb DNA-binding domain with cooperative recognition helices. *Cell* 1994;79:639–48.
- [24] Bozdech Z, Llinas M, Pulliam BL, Wong ED, Zhu J, DeRisi JL. The transcriptome of the intraerythrocytic developmental cycle of *Plasmodium falciparum*. *PLoS Biol* 2003;1:5.
- [25] Le Roch KG, Zhou Y, Blair PL, et al. Discovery of gene function by expression profiling of the malaria parasite life cycle. *Science* 2003;301:1503–8.
- [26] Ness SA, Marknell A, Graf T. The v-myb oncogene product binds to and activates the promyelocyte-specific *mim-1* gene. *Cell* 1989;59:1115–25.
- [27] Doerig C, Horrocks P, Coyle J, et al. Pferk-1, a developmentally regulated cdc2-related protein kinase of *Plasmodium falciparum*. *Mol Biochem Parasitol* 1995;70:167–74.
- [28] Doerig CM, Parzy D, Langsley G, Horrocks P, Carter R, Doerig CD. A MAP kinase homologue from the human malaria parasite, *Plasmodium falciparum*. *Gene* 1996;177:1–6.
- [29] Cervellera MN, Sala A. Poly(ADP-ribose) polymerase is a B-MYB coactivator. *J Biol Chem* 2000;275:10692–6.
- [30] Galio L, Briquet S, Cot S, Guillet JG, Vaquero C. Analysis of interactions between huGATA-3 transcription factor and three GATA regulatory elements of HIV-1 long terminal repeat, by surface plasmon resonance. *Anal Biochem* 1997;253:70–7.
- [31] Blank V, Kourilsky P, Israel A. Cytoplasmic retention, DNA binding and processing of the NF-kappa B p50 precursor are controlled by a small region in its C-terminus. *EMBO J* 1991;10:4159–67.
- [32] Markle D, Das S, Ward SV, Samuel CE. Functional analysis of the KCS-like element of the interferon-inducible RNA-specific adenosine deaminase ADAR1 promoter. *Gene* 2003;304:143–9.
- [33] Pandey KC, Singh S, Pattnaik P, et al. Bacterially expressed and refolded receptor binding domain of *Plasmodium falciparum* EBA-175 elicits invasion inhibitory antibodies. *Mol Biochem Parasitol* 2002;123:23–33.
- [34] Decherig KJ, Thompson J, Dodemont HJ, Eling W, Konings RN. Developmentally regulated expression of pfs16, a marker for sexual differentiation of the human malaria parasite *Plasmodium falciparum*. *Mol Biochem Parasitol* 1997;89:235–44.
- [35] Labesse G, Mornon J. Incremental threading optimization (TITO) to help alignment and modelling of remote homologues. *Bioinformatics* 1998;14:206–11.
- [36] Quandt K, Frech K, Karas H, Wingender E, Werner T. MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. *Nucleic Acids Res* 1995;23:4878–84.
- [37] Osta M, Gannoun-Zaki L, Bonnefoy S, Roy C, Vial HJ. A 24 bp cis-acting element essential for the transcriptional activity of *Plasmodium falciparum* CDP-diacylglycerol synthase gene promoter. *Mol Biochem Parasitol* 2002;121:87–98.

2.1.2 Résultats complémentaires

2.1.2.a Expression de la protéine PfMyb1 recombinante

La découverte d'une phase ouverte de lecture ayant des similarités avec la famille des FT à domaines tryptophane, dans les séquences accumulées peu à peu lors du séquençage du génome de *P. falciparum*, nous a permis de cloner ce gène à partir de l'ADN génomique du clone 3D7 de *P. falciparum*. La séquence complète de la phase ouverte de lecture a été clonée dans le vecteur PGEM-T easy et séquencée. Deux clones indépendants ont donné une séquence identique et celle-ci a été soumise à la banque de données EMBL (décembre 2000 numéro d'accèsion AJ291747).

Lors de la comparaison de la séquence clonée avec celle présente dans la banque PlasmoDB, deux différences dans la région N-terminale ont été observées. Ces variations dans la séquence ont résulté au changement de deux acides aminés lysines, placés en positions 15 et 16, en glutamines. La séquence complète de la phase ouverte de lecture a ensuite été insérée dans le vecteur PQe30 afin de produire la protéine PfMyb1 recombinante ajoutée de six histidines dans la partie N-terminale. Après transformation des bactéries, la protéine PfMyb1 recombinante a été exprimée, purifiée et analysée en utilisant les lysats par électrophorèse SDS-PAGE et Western blot. Même si la protéine semblait largement insoluble, la fraction insoluble a pu être purifiée.

La figure 1 présente une expérience typique. Les lignes 1 et 2 montrent les lysats d'*E. coli* non induits et induits pour la production de la protéine PfMyb1 recombinante, et la ligne 3 les protéines obtenues après élution de la colonne, après une électrophorèse SDS-PAGE et coloration au bleu de Coomassie. Dans la partie droite de la figure 1, la protéine purifiée a été révélée par Western Blot utilisant un anticorps anti-histidine (ligne 4), le sérum anti-PfMyb1 (ligne 6), l'anticorps immuno-purifié anti-PfMyb1 (ligne 7) ou le sérum pré-immun (ligne 5).

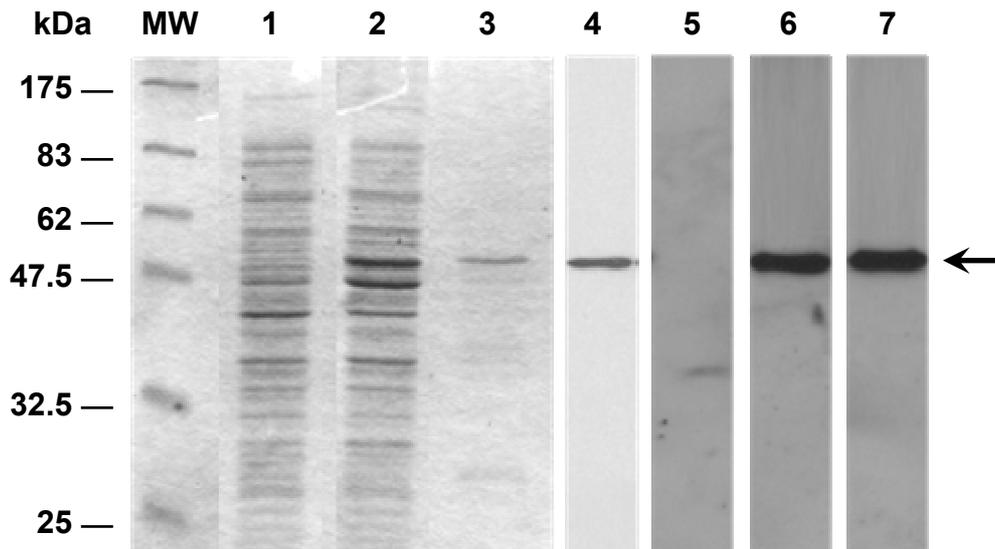


Figure 1. Analyse de la protéine PfMyb1 recombinante. La protéine a été exprimée dans *E. coli* et les lysats non induits et induits ont été fractionnés sur un gel 12 % SDS-PAGE coloré au bleu de Coomassie (lignes 1 et 2). La protéine PfMyb1 recombinante obtenue après élution de la colonne de nickel est présentée en ligne 3. Les lignes 4 à 7 représentent une analyse de la protéine PfMyb1 purifiée en Western blot utilisant un anticorps dirigé contre le tag histidine (ligne 4), un sérum immun anti-PfMyb1 et un anticorps immuno-purifié anti-PfMyb1 (lignes 6 et 7, respectivement), et le sérum pré-immun (ligne 5). Le marqueur de protéines est situé sur la partie gauche et une flèche indique la position de la protéine PfMyb1.

Ces expériences nous ont permis de confirmer la spécificité de l'anticorps qui a été utilisé en figure 2 de l'article 2 (p.89).

2.1.2.b Etude de la cinétique d'expression du transcrit *pfmyb1*

Le transcrit du gène *pfmyb1* a été caractérisé par Northern Blot à partir d'une préparation de d'ARN polyA⁺ isolée d'un mélange d'ARN totaux de différents stades du cycle érythrocytaire du parasite. Seul ce protocole, engageant une grande quantité d'ARN, nous a permis d'identifier un transcrit d'une taille de 2,6 kb, donnant un exemple supplémentaire de parties transcrites mais non traduites exceptionnellement longues (Figure 2A).

La cinétique d'expression temporelle du transcrit a été analysée tout au long du développement érythrocytaire par RT-PCR semi-quantitative (anneaux, jeunes et vieux trophozoïtes, de même que jeunes et vieux schizontes) et comparée entre deux clones de *P. falciparum*: 3D7 et F12. Les stades de développement du parasite ont été collectés, après trois synchronisations successives au sorbitol, en fonction du temps et d'une analyse morphologique de frottis colorés au GIEMSA. L'intégrité des différentes préparations d'ARN

a été vérifiée par coloration au bromure d'éthidium (Figure 2A) et par le bioanalyseur d'Agilent (Figure 2B). La spécificité des produits de la RT-PCR a été confirmée par hybridation avec un oligonucléotide interne à la séquence (données non présentées). Pour chaque expérience, deux RT-PCR contrôles en absence d'ARN et de transcriptase inverse ont été réalisées (données non présentées). De plus, des expériences de RT-PCR réalisées avec des amorces spécifiques du gène MAL13P1.76, situées autour d'un intron, nous ont permis d'exclure une contamination par de l'ADN génomique dans les préparations d'ARN total traitées à la DNase I (données non présentées).

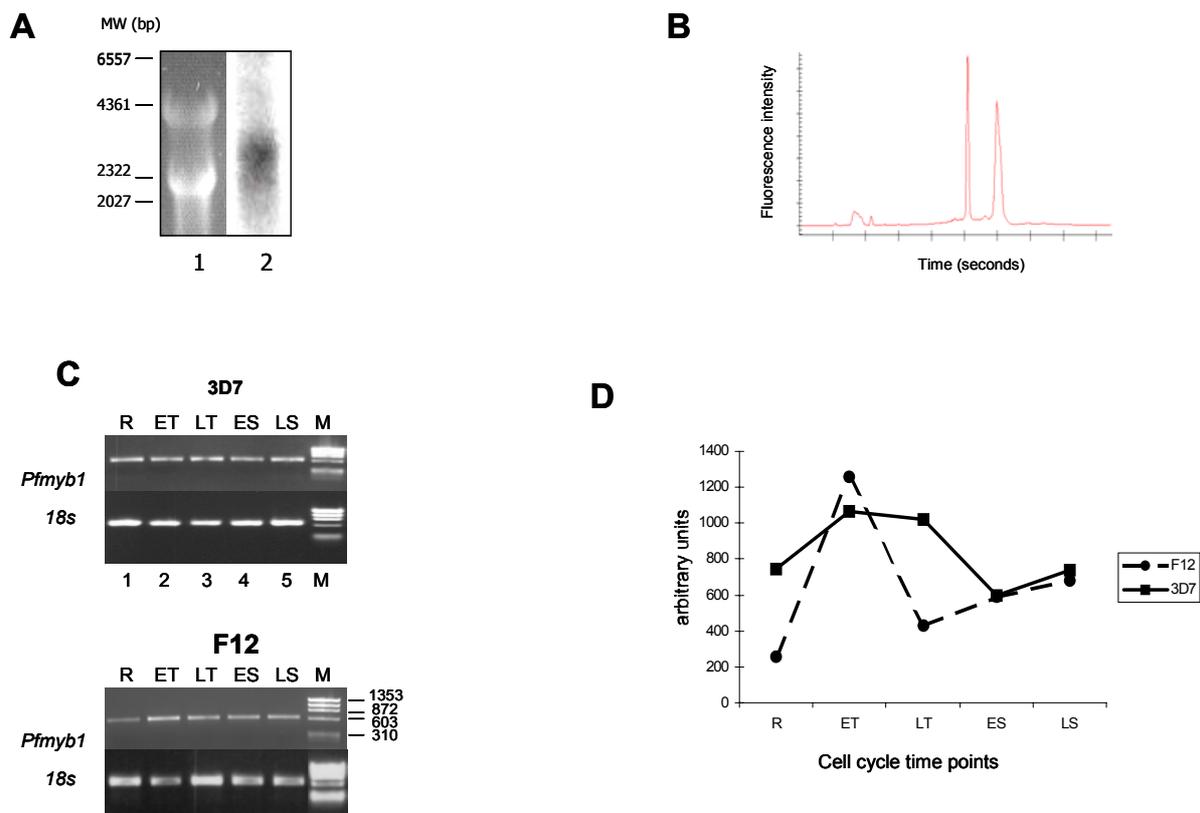


Figure 2. Analyse de l'expression temporelle des ARNm de *pfmyb1*.

L'ARN Poly A⁺ a été purifié à partir d'un mélange de 35 µg d'ARN total issu des différents stades de développement du clone 3D7, dont la coloration au bromure d'éthidium est présentée en ligne 1 (A) et dont le profil Agilent montre un ratio *28S/18S* de 1,86 (B). Le Northern Blot hybridé avec une ribosonde de *pfmyb1* est présenté en ligne 2 (A). Les amplicons de *pfmyb1* et du gène de l'ARN ribosomal *18S* (indiqués sur la gauche) ont été obtenus après extraction d'ARN total à différents temps du cycle érythrocytaire: anneaux (R), jeunes (ET) et vieux trophozoïtes (LT), jeunes (ES) et vieux schizontes (LS). Le marqueur de taille est présenté sur la droite (C). Représentation graphique des intensités des fragments amplifiés après normalisation avec l'ARN ribosomal *18S*. Les figures 2C et 2D sont représentatives d'une expérience sur trois réalisées (D).

Les profils de RT-PCR ont été déterminés après vérification des conditions utilisées, afin que ces expériences soient semi-quantitatives (températures d'hybridation des amorces,

concentration des différents composants et nombre de cycles pour l'amplification des deux séquences *pfmyb1* et *18S*). La cinétique d'expression du transcrit *pfmyb1a* a été suivie en normalisant le fragment de 634 pb provenant de l'ARNm de *pfmyb1* (Figure 3C) par l'intensité des amplicons du gène de l'ARN ribosomal *18S* (Figure 3D). Ces expériences montrent que l'abondance du messenger de *pfmyb1* est régulée au cours du développement érythrocytaire. Le profil d'expression du transcrit *pfmyb1* est similaire entre les clones 3D7 et F12. Cependant, le niveau d'expression au stade anneau est plus haut dans le clone 3D7. Pour les deux clones, le niveau d'ARNm de *pfmyb1* est maximum au stade trophozoïte et diminue de manière équivalente lors du passage au stade schizonte.

2.1.2.c Etude de la cinétique d'expression de la protéine PfMyb1

Le profil d'expression de la protéine PfMyb1 a été déterminé après immuno-précipitation et Western Blot sériés avec le sérum anti-PfMyb1 et à partir d'extraits nucléaires provenant des clones 3D7 et F12 (Figure 3A). Les extraits nucléaires ont été préparés à partir de cultures synchronisées et pour trois points du développement au cours du cycle érythrocytaire : anneaux, trophozoïtes et schizontes. Plusieurs immuno-précipitations, avec le sérum pré-immun ou avec des billes magnétiques seules, suivies d'un Western Blot sérié avec le sérum anti-PfMyb1, n'ont pas donné de signal (données non présentées).

Une quantification de l'intensité donnée par la bande représentant la protéine PfHsp70 nous a permis de nous assurer que, pour chaque point de la cinétique, une quantité équivalente de protéines avait été déposée pour le SDS-PAGE (données non présentées). La cinétique d'expression de la protéine PfMyb1 semble très similaire pour les deux clones 3D7 et F12 (Figure 3B). Le niveau d'expression est relativement plus élevé au stade anneau qu'au stade schizonte. Le pic d'expression de la protéine est observé au stade trophozoïte. Dans cette première expérience, la comparaison entre les cinétiques d'expression du transcrit et de la protéine révèle une certaine proximité entre les deux clones, sauf pour le stade anneau du clone F12 où le niveau du transcrit apparaît inférieur à celui de 3D7.

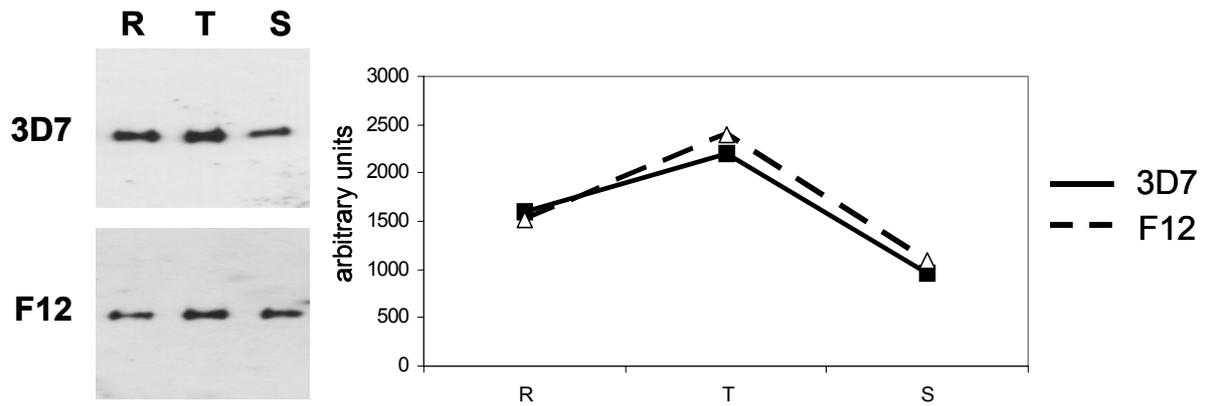


Figure 3. Analyse de l'expression temporelle de la protéine PfMyb1. (A) Immuno-précipitation et Western Blot sériés de la protéine PfMyb1. Les extraits nucléaires de stades anneau (R), trophozoïte (T) et schizonte (S) ont été soumis à une immuno-précipitation utilisant des billes magnétiques recouvertes de sérum anti-PfMyb1, suivie d'un Western Blot avec le sérum anti-PfMyb1. (B) Représentation graphique des intensités des bandes révélées par le Western Blot après normalisation avec l'intensité donnée par les bandes révélées par un Western Blot utilisant un anticorps anti-PfHsp70. Les figures 3A et 3B sont représentatives d'une expérience sur deux réalisées.

Ces résultats sont à mettre en perspective avec ceux obtenus lors des cinétiques d'interaction avec la séquence d'ADN contenant un MRE, par la technique du retard sur gel, et pour les clones 3D7 et F12 (Figure 2 article 2). Tout d'abord, il semble que la cinétique d'expression de la protéine PfMyb1, pour le clone 3D7, soit très proche des résultats obtenus avec la cinétique d'interaction. Cette donnée vient renforcer l'hypothèse que PfMyb1 joue un rôle central dans la détermination de l'interaction avec un MRE pour le clone 3D7. Cependant, nous avons pu remarquer une absence d'interaction pour les extraits nucléaires des stades anneau et schizonte du clone F12. Ces résultats contrastent avec le profil d'expression de la protéine PfMyb1 pour ce clone et suggèrent l'implication d'une autre protéine dans la formation du complexe d'interaction ou une modification différentielle de son activité dans les deux clones par le biais d'une modification post-traductionnelle.

2.2 Rôle de PfMyb1 lors du cycle érythrocytaire et caractérisation de son réseau de régulation.

2.2.1 Résumé.

Après la caractérisation fonctionnelle de la protéine PfMyb1, nous avons acquis assez de données pour inférer sa fonction en tant que FT de la famille des Myb et spécifique de séquence. Le rôle d'un FT est évidemment de réguler l'expression de certains gènes et c'est dans ce cadre que nous voulions induire une baisse de l'expression de PfMyb1 afin d'identifier les gènes cibles de ce FT.

Nous avons utilisé des ARNdb, spécifiques du gène *pfmyb1*, dans le but d'induire une altération de l'expression de PfMyb1 et ainsi tenter d'identifier des gènes cibles de ce FT. Ces expériences pourraient nous permettre de mieux comprendre son rôle et ses modalités d'action pendant le cycle érythrocytaire.

L'ajout d'ARNdb de *pfmyb1* dans le milieu de culture des parasites au stade anneaux semble induire une baisse de croissance par rapport à une culture contrôle analysée 48 h plus tard par trois techniques différentes. Cette baisse de la parasitémie n'est pas observée dans des cultures où l'on a ajouté des ARNdb spécifiques pour un gène du VIH-1.

Lorsqu'on examine l'expression du gène *pfmyb1*, autant au niveau du transcrit qu'au niveau de la protéine, nous observons une baisse significative dans les cultures traitées en comparaison du témoin, au stade trophozoïte, le moment du pic de son expression pendant le cycle érythrocytaire (Figure 1, article 3, p.130).

Les conséquences de l'altération partielle de l'expression du transcrit et de la protéine codant ce FT (Figure 2 et 3, article 3, p.131 et 132), au moment de son expression maximale, ont fait l'objet d'une étude grâce à l'utilisation de la puce à ADN ciblée. Les résultats obtenus ont permis d'identifier huit gènes dont l'expression est altérée dans les cultures traitées par rapport au témoin. Ces modifications de l'expression impliquent que ces gènes soient régulés directement ou indirectement par PfMyb1. Après confirmation de ces altérations, nous avons tenté d'identifier les gènes régulés directement par PfMyb1. Pour cela, l'interaction des promoteurs avec PfMyb1 a été analysée par immuno-précipitation de la chromatine au stade trophozoïte. Ces expériences ont mis en évidence les gènes dont la régulation est directement dépendante à la protéine PfMyb1. Elles ont ensuite souligné la complexité des déterminants de la liaison de PfMyb1 à l'ADN.

Article 3 :

PfMyb1, a *Plasmodium falciparum* transcription factor, is required for intra-erythrocytic growth and controls key genes for cell cycle regulation

Mathieu Gissot, Sylvie Briquet, Philippe Refour, Charlotte Boschet et Catherine Vaquero.

Accepté à la publication dans *Journal of Molecular Biology*.

PfMyb1, a *Plasmodium falciparum* transcription factor, is required for intra-erythrocytic growth and controls key genes for cell cycle regulation.

Mathieu Gissot¹, Sylvie Briquet¹, Philippe Refour^{1,2}, Charlotte Boschet¹ and Catherine Vaquero^{1*}.

¹ INSERM U511, CHU Pitié-Salpêtrière, 91 boulevard de l'Hôpital, 75013 Paris, France.

² Present address: Harvard School of Public Health, Department of Immunology and Infectious Diseases, 665 Huntington Ave, Boston, MA 02115, USA.

Running title: *Plasmodium* transcriptional gene regulation by PfMyb1

***: Corresponding author:**

Catherine Vaquero

Phone: 33 1 40 77 81 11

Fax: 33 1 45 58 88 53

E-mail: vaquero@ext.jussieu.fr

Summary

During the complex life cycle of *Plasmodium falciparum*, divided between mosquito and human hosts, the regulation of morphologic changes implies a fine control of transcriptional regulation. Transcriptional control, however, and in particular its molecular actors, transcription factors and regulatory motifs, are as yet poorly described in *Plasmodium*. In order to decipher the molecular mechanisms implicated in transcriptional regulation, a transcription factor belonging to the tryptophan cluster family was studied. In a previous work, the PfMyb1 protein, contained in nuclear extracts, was shown to have DNA binding activity and to interact specifically with *myb* regulatory elements. We used long *pfmyb1* double stranded RNA (dsRNA) to interfere with the cognate messenger expression. Parasite cultures treated with *pfmyb1* dsRNA exhibited a 40 % growth inhibition when compared with either untreated cultures or cultures treated with unrelated dsRNA, and parasite mortality occurred during trophozoite to schizont transition. In addition, the *pfmyb1* transcript and protein decreased by as much as 80 % in treated trophozoite cultures at the time of their maximum expression. The global effect of this partial loss of transcript and protein was investigated using a thematic DNA microarray encompassing genes involved in signal transduction, cell cycle and transcriptional regulation. SAM software enabled us to identify several genes that were differentially expressed and probably directly or indirectly under the control of PfMyb1. Using chromatin immuno-precipitation, we demonstrated that PfMyb1 binds, within the parasite nuclei, to several promoters and therefore participates directly in the transcriptional regulation of the corresponding genes. This study provides the first evidence of a regulation network involving a *Plasmodium* transcription factor.

Keywords : *Plasmodium*; gene regulation; transcription; transcription factor; PfMyb1; Myb.

Introduction

Plasmodium is responsible for 1.5 to 2.7 millions death annually¹, mostly among children and pregnant women. Of the four species of *Plasmodium* infecting humans, *Plasmodium falciparum* causes the highest morbidity and mortality. As global efforts to eradicate malaria have failed, there is an urgent need to decipher the biology of *Plasmodium* and in particular the mechanisms of gene regulation that govern its developmental cycle, so as to propose novel strategies to fight malaria².

The cell cycle of *Plasmodium* between vertebrate and mosquito hosts implies a high degree of adaptation and strict control of the cellular machinery to correspond precisely to the state of differentiation in the host. These different developmental stages require coordinated modulation of expression of distinct sets of genes, which could be achieved by transcriptional and/or post-transcriptional controls. In *Plasmodium*, as in all eukaryotes, gene expression is governed at the level of transcription by the interaction of elements within promoters acting in *cis* (DNA regulatory elements) and elements acting in *trans* (transcription factors) whose availability is modulated during cellular development. The fine-tuning of transcriptional regulation resides in the interplay between the *cis* and *trans* elements throughout *P. falciparum* development. Little, however, is currently known about these two actors in *Plasmodium*.

The study of *Plasmodium* chromosomes 2 and 3^{3;4} and recently of the complete 3D7 clone sequence⁵, has revealed a genomic organization composed of rather short intergenic sequences, encompassing the basal and distal promoter sequences⁶, between successive open reading frames. The messengers thus far investigated, while few in number, appear to resemble eukaryotic transcripts. They are capped and bear canonic poly-adenylation signals⁷. Moreover, the nucleosome organization⁸, the structure of the promoters and the general transcription factors^{9; 10; 11; 12; 13} are quite similar to those of other eukaryotes. The regulatory sequences within the promoters appear to contain motifs similar to several binding sites encountered in eukaryotes^{6; 14; 15}. Nevertheless, the majority of the regulatory elements reported so far have not been linked to any known transcription factor^{16; 17; 18} or found to bind to unknown nuclear factors¹⁹.

In order to decipher the molecular mechanisms implicated in transcriptional regulation during the erythrocytic development, we have identified, among diverse families of transcription factors, a Myb-related protein belonging to the tryptophan cluster family. In eukaryotes, the Myb proteins are highly conserved throughout evolution and generally contain 3 repeats of approximately 50 residues with 3 regularly spaced tryptophan residues^{20; 21}. Indeed, a large number of Myb-related proteins have been shown to bind DNA in a sequence-specific manner and to participate in the regulation of the expression of genes implicated in growth control and differentiation²¹. PfMyb1 was the first transcription factor not belonging to general transcription factors to be cloned and analyzed in *Plasmodium falciparum*²². This study showed that *pfmyb1* mRNA reached its highest level at the trophozoite stage, at least as regards the 3D7 clone. In *in vitro* experiments, the PfMyb1 protein present in nuclear extracts was able to bind specifically to a prototype *myb*-regulatory element (MRE) from the chicken *mim-1* gene promoter²³ and 2 putative MRE annotated within the *Plasmodium* promoters²², suggesting that PfMyb1 can be considered to be a Myb transcription factor.

We took advantage of long *pfmyb1* double stranded RNA (dsRNA) to interfere with expression of the cognate messenger and to address the role of this transcription factor during the erythrocytic cycle. In particular, we investigated the consequences of decrease in *pfmyb1* transcript and protein on parasite growth and transcriptional regulation of *P. falciparum* genes.

Results

1. Pfmby1 dsRNA alters parasite growth and trophozoite to schizont transition.

We investigated the global effect of double stranded RNA corresponding to part of the *pfmyb1* sequence on *P. falciparum* growth when added to highly synchronized parasite cultures during a 48 hr assay. Parasites were considered to be synchronized after three successive sorbitol treatments leading to an enrichment in rings of at least 95 %. In young ring cultures, *pfmyb1* dsRNA or dsRNA of an unrelated gene (HIV-1 *gp120* AN: **K02013**) was added (25 µg per ml of culture) and parasitaemia was evaluated using three different methods (Fig. 1). As shown in Figure 1A, addition of unrelated dsRNA to culture media (line 2) did not affect asexual growth when compared with control cultures (line 1), i.e. parasites grown in media supplemented with identical volumes of annealing buffer. In contrast, a statistically significant decrease in parasitaemia (approximately 40 %) was observed in cultures treated with *pfmyb1* dsRNA using manual counting (Fig. 1A, line 3) or a ³H-hypoxanthine assay (Fig. 1A, line 4). The result obtained by manual counting (line 3) corresponds to the average of 9 experiments performed in triplicate, while the result obtained by ³H-hypoxanthine assay (line 4) corresponds to two experiments carried out in triplicate. Finally, using a flow-cytometry assay, a similar degree of growth inhibition (40 %) was observed after treatment with *pfmyb1* dsRNA as compared with controls treated with single stranded sense or anti-sense *pfmyb1* RNA (data not shown). Thus three different experimental approaches used for evaluating the inhibition of growth upon addition of dsRNA specific for *pfmyb1* gave essentially the same results (Fig. 1).

We investigated the effect of *pfmyb1* dsRNA on trophozoite to schizont transition (Fig. 1B). During the erythrocytic cycle, smears were performed every 4 hr and the number of parasites in trophozoite (Fig. 1B, lines 1 and 2) and in late schizont stage (Fig. 1B, lines 3 and 4) was evaluated. The parasitaemia observed in the trophozoite cultures was quite similar in the control (Fig. 1B, line 1) and *pfmyb1* dsRNA treated cultures (Fig. 1B, line 2), in contrast to that of late schizonts showing a significant decrease in *pfmyb1* dsRNA treated parasites (Fig. 1B, line 4 versus line 3). Moreover, the

growth inhibition (Fig 1A) and the decrease in schizont number (Fig. 1B) were of an essentially similar magnitude.

2. Decreased expression of *pfmyb1* messenger is observed in *pfmyb1* dsRNA treated culture.

The expression of the *pfmyb1* transcript was analysed at the trophozoite stage, at the peak of expression in the erythrocytic cycle, in cultures treated with *pfmyb1* or *gp120* dsRNA, or left untreated, using a semi-quantitative (Fig. 2A and 2B) and real time quantitative RT-PCR (Fig. 2B). For each set of experiments, PCR control reactions were performed in the absence of RNA or RT (data not presented), as well as on the MAL13P1.76 gene with two primers encompassing an intron (Fig. 2A left panel). These data demonstrated the absence of DNA contamination in the RNA preparations. First, by semi-quantitative RT-PCR (Fig. 2A, right and left panel), we showed that the expression of the *pfmyb1* transcript was decreased in *pfmyb1* dsRNA treated parasites when compared with control parasites, while expression of the MAL13P1.76 transcript was unaltered. After normalization against ribosomal *18s* RNA, a 75 % decrease in the *pfmyb1* transcript was observed in *pfmyb1* dsRNA treated cultures (Fig. 2B, lines 3 and 6) when compared with control cultures (lines 1 and 4) or cultures treated with unrelated dsRNA (lines 2 and 5), using semi-quantitative (lines 1-3) or real time quantitative RT-PCR (RT-qPCR) (lines 4-6). RT-qPCR was performed using a multiplex assay amplifying *pfmyb1* and the *18s* reference in the same experiment. For the experiments presented herein, different couples of primers (see Experimental procedures) were used, all yielding similar results.

3. Decreased expression of PfMyb1 protein is observed in *pfmyb1* dsRNA treated culture.

In order to follow the effect of *pfmyb1* dsRNA treatment on PfMyb1 protein expression, serial immuno-precipitation and Western blotting experiment were performed. Protein extracts from *P. falciparum* trophozoites were immuno-precipitated using magnetic beads coated with anti-PfMyb1 serum. Immuno-precipitated proteins were subjected to SDS PAGE followed by Western blotting with the anti-PfMyb1 serum. No bands were detected after immuno-precipitation of lysates with either

preimmune serum or uncoated beads and Western blotting with the anti-PfMyb1 serum (data not shown). The abundance of the PfMyb1 protein was markedly decreased in *pfmyb1* dsRNA treated parasites (Fig. 3A, lower panel) as confirmed by normalization with PfHsp70 (Fig. 3A, upper panel). Indeed, a reduction of approximately 80 % in the PfMyb1 protein was observed in protein extracts from trophozoite cultures treated with *pfmyb1* dsRNA when compared to untreated cultures (Fig. 3B, lane 2 versus lane 1), while no such reduction was observed in controls treated with unrelated dsRNA (lane 3).

4. Genes are differentially expressed in *pfmyb1* dsRNA treated culture.

We then decided to investigate the effect of the partial loss of the PfMyb1 transcript and protein on the global expression profile, by comparing untreated culture with cultures treated with either *pfmyb1* dsRNA or irrelevant dsRNA at a time when the expression of the *pfmyb1* transcript was highest, i.e. the trophozoite stage.

With the use of a thematic microarray fully described in Gissot *et al.*²⁴ and comprising 180 PCR products printed in duplicate, among which 153 amplified from the *Plasmodium falciparum* genomic sequences, we monitored the steady state of the corresponding transcripts. The genes were selected for their homology to known proteins previously described in eukaryotes to be involved in genetic regulation from DNA to proteins. The microarray quality controls were performed as detailed in Gissot *et al.*²⁴. Briefly, for each growth condition, at least two different total RNA preparations from independent cultures were used for synthesis of cDNA radiolabelled with ³H or ³⁵S radioisotope, and for subsequent hybridization of two sets of arrays. Each raw signal was normalized twice, first with the Alien cDNAs in order to normalize for the efficiency of each reverse transcription and second with 3D7 genomic DNA to normalize for the amount of each cDNA used for the hybridization. Statistically significant differences in gene expression were monitored using the Statistical Analysis for Microarrays (SAM) program²⁵. Among the 153 *P. falciparum* genes printed on the microarray, 139 genes gave signals at least twice as high as background (see supplementary data 1 and 2).

When untreated controls were compared with HIV-1 *gp120* dsRNA treated culture, no differences in the relative expression level of any gene were detected using the SAM program (see supplementary data 2). In contrast, upon comparison of *pfmyb1* dsRNA treated and untreated parasites, 11 genes were identified as statistically different and the 128 remaining genes as similar. Among the 11 genes, one was up-regulated and 10 down-regulated, as shown in the SAM output plot presented in Figure 4.

The differential expression of these 11 genes was verified by real-time quantitative RT-PCR using the *18s* gene as reference. Using two independent biological samples, 3 (numbered 2, 3 and 4 in Figure 4) of the 10 down-regulated genes did not give sufficiently consistent results to be taken into account. The results of the RT-qPCR (closed bars) and microarray (open bars) analyses, for the genes whose expression was modulated by *pfmyb1* dsRNA (7 down and 1 up-regulated gene, also listed in Table I) and diverse additional genes as controls of the experiments, are compared in Figure 5.

Indeed, several genes whose expression profile was similar in control and *pfmyb1* dsRNA treated parasites were selected and their expression compared by DNA microarray or RT-qPCR experimental approaches. PfHSP70 (PF08_0054) and MAL13P1.76 were shown to be equally expressed in treated and untreated cultures when analysed by either experimental approach. This was also the case for PF10_0233, PFD0465w and MAL8P1.154 genes, which showed slight but detectable homology with the *pfmyb1* sequence used to generate *pfmyb1* dsRNA by using the BLASTN tool on the PlasmoDB web site, as well as for *pfmyb2* (PF10_0327) and *pfmyb3* (PF10_0143), belonging to the tryptophan cluster family of transcription factors. Moreover, the RT-qPCR also confirmed the results obtained for the 7 down- and 1 up-regulated gene, attesting to the reliability of the microarray data (Fig. 5).

Finally, the over-expression of the *pfpk5* transcript in *pfmyb1* dsRNA treated samples was also established at the level of protein expression (Fig. 6). Western blotting experiments on whole trophozoite protein extracts showed an increased representation of PfPk5 protein in *pfmyb1* dsRNA treated samples as compared with untreated samples, in good agreement with the transcript data.

5. PfMyb1 is associated with the promoter of genes differentially expressed using ChIP experiments.

In order to determine whether, within the parasite nuclei, the PfMyb1 protein could interact with promoters of genes listed in Table I, and therefore could be involved in the regulation of the targeted genes, chromatin immuno-precipitation (ChIP) experiments were carried out. PfMyb1 protein present in nuclear extracts has been shown to interact with Myb related elements (MRE)²². A computational analysis was performed on the promoters of the genes listed in Table I to identify these MRE, and revealed that among the 8 promoters only that of PGK (PFI1105w) did not contain any putative MRE. The PfMyb1-DNA complexes were cross-linked within the parasite nuclei and subsequently immuno-precipitated by using the anti-PfMyb1 coated beads. No cross-linked PfMyb1-DNA complexes were immuno-precipitated using either beads alone or beads coated with preimmune serum (data not shown). After purification of the DNA fragments, PCR analysis was carried out for the 8 previously selected promoters. In Figure 7, odd numbers correspond to PCR performed on input DNA, and even numbers to immuno-precipitated DNA.

Promoters of 3 genes containing MRE (PF08_0054, PF10_0159, PF11_0096), but whose expression was not altered in parasites cultured in the presence of *pfmyb1* dsRNA in microarray experiments, were amplified in input (Fig. 7 lanes 17, 19 and 21) but not ChIP samples (Fig. 7 lanes 18, 20 and 22), even though expression of their cognate transcript peaked at the trophozoite stage²⁶. In addition, the promoter of the PF07_0073 gene, devoid of MRE, yielded identical results (Fig. 7 lanes 15-16).

In contrast, among the 8 genes whose expression was altered in the presence of *pfmyb1* dsRNA, 6 (MAL13P1_279, PFI1105w, PFL1285c, MAL6P1.248, MAL6P1.249 and PF14_0224) gave rise to amplification products with both anti-PfMyb1 immuno-precipitated (Fig. 7 lanes 2, 4, 10, 12 and 14) and input DNA (Fig. 7 lanes 1, 3, 9, 11 and 13). These 6 genes corresponded to 5 promoters, since the genes encoding histones H3 and H2A (MAL6P1.248 and MAL6P1.249) share a bidirectional promoter. Finally, the promoters of two (PFL1885c and PFL2345c) of these 8 genes gave no amplification products using ChIP (Fig. 7 lanes 6 and 8) in contrast to input samples (Fig. 7 lanes 5 and 7).

Discussion

In light of the crucial role of transcriptional regulation during the complex life cycle of *P. falciparum*, we decided to investigate PfMyb1, a protein belonging to the tryptophan cluster family. Indeed, previous experiments have shown that this protein interacts specifically with several MRE²². These data, together with computational study of the PfMyb1 amino acid sequence²², have led us to consider this protein to be a genuine transcription factor.

In order to investigate the role of PfMyb1 in transcriptional regulation of *P. falciparum*, we decided to interfere, by means of long *pfmyb1* double stranded RNA, with its transcript and therefore its protein expression at the time of its highest expression, *i.e.* during the trophozoite stage^{22; 26; 27}. As shown in Figure 1A, a 40 % growth inhibition was observed when parasite cultures were treated with *pfmyb1* dsRNA as compared with either untreated parasite cultures or cultures treated with unrelated long dsRNA. The observed inhibition thus appeared to depend specifically on *pfmyb1* dsRNA. This observation was made by using three different methods and confirmed by repeated experiments that consistently yielded similar results. The unrelated dsRNA was chosen for its relatively high AT content (60 %) and its length (570 nt), approximating those of the *pfmyb1* dsRNA (70% AT and 634 nt, respectively). Moreover, using the BLASTN tool of PlasmoDB, we did not find any detectable homology between this unrelated sequence and the entire *P. falciparum* genome. We further showed that parasite death occurred during the trophozoite to schizont transition, notably after the *pfmyb1* expression peak (Fig. 1B).

Since addition of *pfmyb1* dsRNA modified parasite growth and development, we investigated the level of the *pfmyb1* transcript at the trophozoite stage. By two different experimental approaches based on RT-PCR (Fig. 2) and performed with different sets of primers, we showed that *pfmyb1* dsRNA treatment diminished the *pfmyb1* transcript in the parasite by 75 %. Moreover, the decrease in the *pfmyb1* transcript was correlated to a decrease in the corresponding protein, as determined by serial immuno-precipitation and Western blotting of the trophozoite protein extracts, using anti-PfMyb1 antibodies²² and after normalization for protein loading. In conclusion, the decreased availability of

pfmyb1 transcript and protein, together with the growth inhibition occurring during the trophozoite to schizont transition, argue in favour of a major function for this transcription factor during the parasite erythrocytic cycle. This assumption is strengthened by the high conservation of *pfmyb1* gene within the diverse *Plasmodium* species listed in PlasmoDB.

In *Plasmodium*, the mechanisms by which gene expression is modified by dsRNA are still not clearly defined. The characterization of such mechanisms, whether involving classical RNA interference as suggested by previous work^{28;29}, or an anti-sense effect as shown by Noonpakdee *et al.*³⁰, is beyond the scope of this study. Herein, we took advantage of long dsRNA to reduce the cognate messenger and encoded protein, so as to study the consequences of this depletion on gene expression. Interestingly, many antisense transcripts have been shown to arise during the erythrocytic cycle^{31; 32} and have been suggested to be potential regulatory elements in gene transcription³³. These studies might imply that the machinery required to process these antisense transcripts and the resulting dsRNA is present in *P. falciparum*.

In order to identify at least some of the transcripts whose expression is governed by the transcription factor, we compared the transcriptome of cultures treated with *pfmyb1* dsRNA and control parasite cultures by means of our thematic DNA microarray. The slides were hybridized with cDNA prepared from total RNA of control and *pfmyb1* treated cultures radiolabelled as previously reported²⁴. Total RNA was extracted at the trophozoite stage prior to parasite death and when maximum expression of the *pfmyb1* transcript was reached (Fig. 1). When unrelated dsRNA treated samples were compared with untreated controls no significant differential expression was identified by SAM (see supplementary data 2). The decrease in *Pfmyb1* transcript and protein mediated by *pfmyb1* dsRNA altered the relative expression of a set of genes in four independent biological experiments normalized and filtered through SAM software (Fig. 4).

In particular, among the 153 *P. falciparum* genes printed on the microarray, the level of expression of 128 transcripts was similar in control and in *pfmyb1* dsRNA treated parasites (see supplementary data 1), suggesting that treatment with *pfmyb1* dsRNA did not markedly alter synchronization and development of the parasites. In this regard, upon examination of GIEMSA stained blood smears no substantial shift was detected between the two growth conditions.

In contrast, 11 genes were identified by SAM as differentially expressed in control and in *pfmyb1* dsRNA treated parasites, of which 1 was up-regulated and 10 down-regulated. All of the microarray data leading to the identification of the genes of interest as well as several control genes were verified by quantitative real time RT-PCR. Three genes identified by SAM were eliminated upon performing independent biological experiments. It is noteworthy that these genes were very close to the limit defined by SAM (2, 3 and 4 of Fig. 4). The remaining 8 genes, listed in Table I, gave similar expression ratios when microarray data were compared with real time RT-PCR data. In addition, the expression ratio of 7 other gene controls (Fig. 5, right part) was verified, confirming the reliability of the microarray experiments. The level of expression of *pfmyb2* (PF10_0327) and *pfmyb3* (PF10_0143), two genes belonging to the tryptophan cluster family of transcription factors, was not modified by *pfmyb1* dsRNA. This was also the case for three genes showing slight but detectable homology to the *Pfmyb1* DNA sequence used to generate dsRNA. Taken together, these data indicate that the activity of *pfmyb1* dsRNA is sequence-dependent.

Among the 8 genes cited as differentially expressed in *pfmyb1* dsRNA treated samples, the up-regulation of PfPk5 (MAL13P1_279) expression observed at the transcript level was also detected at the protein level by Western blotting (Fig. 6). It is worthy of note that PfPk5 has been annotated as a *cdc2* related cyclin-dependent kinase bearing 60 % identity to the human *cdc2* gene³⁴. In all eukaryotes, cyclin-dependent kinases are key regulators of cell cycle control and progression³⁵. PfPk5 has been implicated in regulation of the nuclear division cycle³⁶ and is active during the erythrocytic cycle, and shares common features with the *cdc*-related kinase activation process³⁷. Taken together, these data suggest a crucial role for PfPk5 in regulating cell cycle progression of *P. falciparum*. A clear link between the classical cell cycle (G1, S, G2 and M phases) and the *Plasmodium* erythrocytic cell cycle has not been established³⁸. Nevertheless, DNA synthesis appears to occur during the trophozoite stage and asynchronous nuclear division during the trophozoite to schizont transition³⁹. These two critical phenomena, however, may overlap during trophozoite to schizont transition. Indeed, PfPk5 may play a crucial role during the trophozoite to schizont transition and its over-expression in *pfmyb1* dsRNA treated culture may explain the growth inhibition and parasite death at that particular time in the erythrocytic cycle.

Furthermore, in *pfmyb1* dsRNA treated samples, a decreased expression in the gene PFL1285c, a homologue of proliferating cell nuclear antigen (PCNA), was observed. In eukaryotes, PCNA has been shown to be a key player in DNA replication, DNA repair and cell cycle control⁴⁰. In *Plasmodium*, PCNA1 expression was maximum at the trophozoite stage at the time of DNA replication⁴¹ and its altered expression could therefore have played a role in the death of *pfmyb1* dsRNA treated parasites.

Alteration of the glucose metabolism *via* the partial inhibition of expression of the phosphoglycerate kinase (PGK, PFI1105w) might also have participated in the alteration of the parasite life cycle. PGK has been found to be a key enzyme of the glucose pathway and its activity has been detected during the erythrocytic cycle⁴². Of course all of the observed variations in gene expression might contribute in a cooperative manner to the alteration of parasite development and therefore participate in promoting parasite death. Finally, two histone (H3 and H2A) genes have been shown to be under-expressed in *pfmyb1* dsRNA treated samples, in contrast to the two other histones (H4 and H2B). The two histone (H3 and H2A) genes appeared to be under the control of a bidirectional promoter. This observation provided another argument in favour of an activity of PfMyb1 as a transcription factor.

In order to determine, among the 8 genes identified by microarray, those whose expression might be directly controlled by PfMyb1, we performed a chromatin immuno-precipitation assay with a specific anti-PfMyb1 antibody, followed by a PCR identification of gene promoters. For six genes (MAL13P1_279, PFI1105w, PFL1285c, MAL6P1.248, MAL6P1.249 and PF14_0224), an interaction was observed *ex vivo* between the promoter and the transcription factor PfMyb1, suggesting that they might be directly regulated by PfMyb1. By contrast, two genes (PFL1885c and PFL2345c, see Figure 7, lanes 5-8) were not amplified in ChIP samples. These genes had in fact shown small variation in the level of gene expression in *pfmyb1* dsRNA samples (spots 5 and 6 of Fig.4). The absence of amplification might be attributable to a weak interaction between their potential MRE and PfMyb1, below the threshold of detection of ChIP methodology, or to poor accessibility of the MRE within the chromatin. Alternatively, these genes might be regulated by a transcription factor itself under PfMyb1 control. Similar results were obtained with three control genes (PF08_0054, PF10_0159, PF11_0096) displaying no significant differential expression in microarray experiments and sharing the same peak

of expression during the trophozoite stage. Identical results were obtained as regards the PF07_0073 promoter that lacks MRE (Fig. 7).

Finally, the PFI1105w promoter, despite the apparent absence of MRE, gave rise to a positive signal in CHIP experiments (Fig. 7, lanes 3 and 4). This result has been shown to arise with a Myb protein from *S. cerevisiae* incapable of binding conventional MRE⁴³. *P. falciparum* might have evolved new PfMyb1/DNA binding specificities that might depend on chromatin context and/or availability of the protein partner. Indeed, identification of binding sites for PfMyb1 in the PGK promoter is a new challenge for understanding PfMyb1 dependent transcriptional regulation. Nevertheless, this experiment clearly demonstrated the ability of the PfMyb1 protein to bind to the promoter of a set of genes of critical importance in the *P. falciparum* cell cycle and therefore to regulate their expression in a direct manner. Moreover, these data suggest that PfMyb1 may participate directly in either repression (for expression of PfPk5) or activation (for the 5 other genes) of transcriptional gene expression. This is a common feature for Myb proteins, among which the c-Myb factor reported to mediate either up or down regulation of different genes²⁰. Hence, depending on post-translational modifications of the transcription factors^{44, 45}, the protein partners present on each given promoter and the chromatin context, the activity of transcription factors could be modulated either for activation or repression of transcription.

In conclusion, this paper describes the first attempt to unravel the PfMyb1 regulation network. Although more work is needed to understand the multiple determinants of PfMyb1 mediated regulation, these experiments provide important clues for future analysis of factors determining regulation of gene expression by PfMyb1. The data reported herein suggest that PfMyb1 is an essential transcription factor for the erythrocytic cycle of *P. falciparum* and, moreover, that PfMyb1 directly regulates key genes involved in cell cycle regulation and progression.

Materials and Methods

Plasmodium falciparum culture

The 3D7 clone of *P. falciparum* was provided by Dr D. Walliker and was cultured as described in ⁴⁶ with slight modifications. To obtain synchronized erythrocytic stages, ring cultures were enriched by 3 successive treatments with 5 % sorbitol at 44 hr intervals ⁴⁷. Synchronization of culture was verified by morphological analysis of GIEMSA stained blood smears at different time points.

dsRNA preparation

A polymerase chain reaction (PCR) fragment (634 nt) representing the *pfmyb1* gene (PF13_0088) was amplified from 3D7 genomic DNA using forward (5'GGATGAAATGTGATGATAAAAC) and backward primers (5'TACTTCATCTTTTGTCCATTTT). The PCR product was cloned into TOPO PCR II vector (Invitrogen). A PCR fragment (570 nt) representing the portion of the *env* gene encoding gp120 (GeneBank accession number **K02013**) was amplified from HIV-1 genomic DNA using forward (5'ATGCCCCAGACTGTGAGTTG) and backward primers (5'CTACTGTAATTCAACACAAC) and cloned into pBluescript SK- vector. The clones were sequenced by dideoxy sequencing reactions. Plasmids were used as a template to generate sense RNA (sRNA) and antisense RNA (asRNA) using T7, SP6 and T3 RiboMAX Express Large Scale RNA Production System (Promega). The dsRNA was prepared as follows: equal amounts of sRNA and asRNA were separately denatured at 95 °C for 1 min, mixed and then heated to 95 °C for an additional minute. Annealing was performed by slow cooling over several hours and thereafter dsRNA samples were analyzed on native 1 % agarose gel.

dsRNA assay

For each experiment using dsRNA, 25 µg of dsRNA were added per mL of highly synchronized young ring culture (4 to 6 hr post invasion). Parasitaemia was set as described in “Parasite growth assay section” for manual counting, ³H-hypoxanthine uptake assay and flow cytometry in 24 well

plates and incubated for 48 hr. Twenty mL culture of 5 % parasitaemia of highly synchronized young ring stages was set for RNA and protein extraction. This culture was incubated in 75 cm³ culture flask for 18-20 hours as described in “*P. falciparum* culture” section. Control cultures were carried out as above and an equivalent volume of dsRNA annealing buffer was added to the medium.

Parasite growth assay

³H-hypoxanthine uptake was evaluated as described in ⁴⁸ with slight modifications. In brief, 200 µl of ring stage parasitized erythrocytes (parasitaemia, 0.5%; hematocrit, 1.8 %) were dispensed in 96-well plates preloaded with 25 µg dsRNA per mL in triplicate and with serial dilutions of chloroquine in positive control wells. After 24 hr, ³H-hypoxanthine was added to each well and plates were then incubated for an additional 24 hr. Parasites were harvested, and incorporation of radioactivity was determined by liquid scintillation counting. Experiments were repeated twice. Data are reported as mean and standard deviation (S.D.) of percentage parasite growth inhibition in triplicate experiments relative to untreated controls after correcting for background ³H-hypoxanthine incorporation in uninfected erythrocytes. Alternatively, growth inhibition was determined by microscopic examination of Giemsa-stained thin blood smears. Smears from untreated cultures were always used as controls. Parasitaemia was measured by counting 3000 red blood cells and expressed as percentage of total parasitized erythrocytes. Twenty five µg of dsRNA was added to ring stage culture (1.5 % parasitaemia, 5 % hematocrit) and parasitaemia was evaluated 48 hr later. Experiments were repeated 9 times in triplicate. Finally, flow cytometry was performed on 48 hr treated culture (1.5 % rings, 5 % hematocrit) as described in ⁴⁹. Experiments were repeated 3 times in triplicate. Statistical analysis was performed using a Student t-test.

RT-PCR and RT-qPCR

RT-PCR: total RNA (1 µg) was treated with 2 U of RNase-free DNase I (Qiagen) and purified using RNeasy columns with the Qiagen clean-up procedure to avoid DNA contamination. Synthesis of cDNA was performed for 50 min at 42 °C using 1 µg of RNA (in 40 µL final volume), 50 U of MMLV reverse transcriptase (Superscript II Invitrogen), 0.5 ng of random hexamers, 5 mM of MgCl₂,

20 U of RNaseout (Invitrogen), 10 mM of DTT and 0.25 mM of each dNTP. Samples were then subjected to RNase-H treatment with 2 U of enzyme for 20 min at 37 °C. PCR was carried out with 2 µL of each cDNA sample (or the negative control) in the presence of 1 µM of primers; *18s* forward (5'GACTCAACACGGGGAACTCACTAGTT) and backward primers (5'ACAATTCATCATCTTTCAATCGGTA) or *pfmyb1* forward (5'GGATGAAATGTGATGATAAAAC) and backward primers (5'TACTTCATCTTTTGTCCATTTT or 5'TCTCCTTTTGTAAGATATTCATCATTCA) or MAL13P1.76 forward (5'ATGCAAAATCCTCAAAAT) and backward (5'TTATGTATCGTTAATTAAAC) primers, 200 µM dNTP, and 2 U of Taq Polymerase enzyme. After denaturation at 95 °C for 2 min, 30 cycles with annealing at 55 °C, elongation at 60 °C and denaturation at 95 °C, each step for 1 min, were performed to amplify the *pfmyb1* and MAL13P1.76 fragments. Twenty cycles of the same amplification profile were used to amplify the ribosomal *18s* subunit. We verified that, under these conditions and number of cycles, the magnitude of the signals remained proportional to RNA concentrations, making the RT-PCR a semi-quantitative procedure. Fragments of expected lengths (657 bp, 1204 bp and 634 bp or 1015 bp for *18s*, MAL13P1.76 and *pfmyb1*, respectively) were observed by 1 % agarose gel electrophoresis and relative amounts of mRNA were determined by densitometry (Densylab, Microvision Instruments, Evry, France).

Real-time quantitative PCR reaction was performed as described in ²⁴. Specific primers are listed in supplementary material 1.

Multiplex real-time quantitative PCR was performed with LUX-primers (Invitrogen). Reactions were performed in a 25 µL volume containing 5 µL of 25-fold diluted cDNA preparations, 12.5 µL of Platinum PCR mix (Invitrogen) supplemented with 3 mM of MgCl₂, and 1875 nM of *pfmyb1* primers and 94 nM of *18s* primers. Amplification and detection of specific products were performed with the MX4000 light cycler (Stratagene) with the following cycle profile: one cycle at 50 °C for 2 min, one cycle at 95 °C for 2 min, 40 cycles with 15 sec denaturation at 95 °C and 30 sec annealing-elongation at 60 °C. *18s* primers were CACCATTTTCTTGATTTCTTGGATGG labelled with JOE, and GATCTCGTTCGTTATCGGAAT. *pfmyb1* primers were CACATCTGCAAAGGAATGTCAAAGATG labelled with FAM, and CACGACAAAGGACTTCTATTGGT.

The size of each PCR product and number of bands were verified on a 10 % acrylamide gel. Two additional reactions were performed, either without reverse transcriptase or without the RNA sample, to verify the absence of DNA contamination. The quantity of cDNA for each experimental gene was normalized to the *18s* concentration (chr5.rRNA-1-18s-A) in each sample. For each sample and each gene, experiments were performed twice and in triplicate. Relative gene expression was expressed as the \log_2 of the ratio of expression at each time point over that of the ring stage using the $2^{-\Delta\Delta CT}$ method.

Serial immuno-precipitation and Western blotting

Protein extracts of *P. falciparum* trophozoites were prepared as described in ¹⁶. One hundred μg of protein were subjected to immuno-precipitation using 2×10^7 magnetic tosylactivated Dynabeads (M-280, Dynal Biotech) coated with 2.5 μg of either anti-PfMyb1 or preimmune serum, according to the manufacturer's recommendations. Immuno-precipitated proteins were resolved by 10 % SDS-PAGE and subjected, after transfer onto nitrocellulose membrane, to Western blotting using the anti-PfMyb1. The membrane was first incubated overnight at 4 °C with a 1/1000 dilution of preimmune or anti-PfMyb1 serum and then incubated with a 1/10,000 dilution of peroxidase-conjugated anti-rabbit antibody (Biolabs). Concomitantly, 5 μg of the supernatant proteins were resolved by 10 % SDS-PAGE and blotted onto nitrocellulose membrane. The membrane was first incubated overnight at 4 °C with a 1/10,000 dilution of anti-PfHsp70 serum and then with a 1/10,000 dilution of peroxidase-conjugated anti-mouse antibody (Biolabs).

Western blotting experiments on PfPfk5 protein were performed on 60 μg of protein extracts prepared as described in ⁵⁰. Proteins were resolved by 12 % SDS-PAGE and subjected, after transfer onto nitrocellulose membrane, to Western blotting using the anti-PfPfk5 serum, a kind gift of C. Doerig. The membrane was first incubated overnight at 4 °C with a 1/1000 dilution anti-PfPfk5 serum and then with a 1/10,000 dilution of peroxidase-conjugated anti-rabbit antibody (Biolabs). PfHSP70 protein was detected as described previously.

All membranes were developed using a chemiluminescent substrate, according to the manufacturer's instructions (ECL, Amersham-Pharmacia Biotech). Relative amounts of proteins were determined by densitometry (Densylab, Microvision Instruments, Evry, France).

Microarray construction, probe labeling, hybridization and data analysis

Microarrays were prepared as described in Gissot *et al.* ²⁴. Briefly, seventeen spots representing positive and negative controls such as sequences unrelated to *P. falciparum* were printed. Alien PCR products from Stratagene, with no apparent similarity to the *P. falciparum* genomic sequence, were used for normalization of the reverse transcription. 3D7 genomic DNA was also printed in order to normalize the total amount of each cDNA present during the hybridization. All targets were spotted in duplicate with a Qarray arrayer (Genetix) onto poly-L-lysine-coated slides prepared as described previously ⁵¹ with a spot spacing of 400 μm , centre to centre onto a 1 cm^2 area.

Probe labeling and hybridization were performed as described previously in ²⁴. Data were collected as an image file, gridded, and converted into an Excel file using β -Vision software (Biospace mesure). The global background was subtracted for each signal. In addition, the signal intensity of each spot was normalized according to the average value of intensities of spots corresponding to exogenous control PCR products (Alien, Stratagene) on each array in order to control for variation in the efficiency of reverse transcription. Furthermore, *P. falciparum* genomic DNA was used to normalize the amount of cDNA present in the hybridization. Four independent biological experiments were performed. Statistical discrimination of genes expressed differentially in the two stages was performed on the complete normalized set of data using the freely available SAM software (<http://www-stat.stanford.edu/~tibs/SAM/>) ²⁵.

Chromatin immuno-precipitation

Potential MRE were localized in promoter sequences by using MatInspector ⁵² according to three matrix families (one from the plant and two from the vertebrate subsections). This protocol is adapted from Hecht *et al.* ⁵³. One hundred mL of trophozoite culture was subjected to protein-DNA crosslink using 1 % formaldehyde (Sigma) at 37 °C for 15 min. Nuclei were prepared as described in ¹⁶ and resuspended in RIPA buffer (30 mM Tris-HCl pH 8, 150 mM NaCl, 20 mM MgCl₂, 1 mM EDTA, 1 mM DTT, 0.5 %, Triton X-100, 1 % Nonidet-P40, 1 mM PMSF, 1X Complete TM mixture tablet from Roche Molecular Biochemicals). Samples were sonicated (BioBlock Scientific, VibraCell

72405) 3 times at 30 W for 10 sec, allowing 30 sec cooling between sonifications, and centrifuged for 5 min at 13000 rpm. The supernatant, containing soluble chromatin, was incubated overnight at 4 °C with 2×10^7 magnetic tosylactivated Dynabeads (M-280, Dynal Biotech) coated with 2.5 µg of either preimmune or anti-PfMyb1 serum. Beads were washed twice with 1 mL RIPA buffer supplemented with 0.01 % SDS. Immuno-precipitated chromatin was eluted using TE (Tris 10 mM pH 8, EDTA 1 mM) supplemented with 1 % SDS. Chromatin was reverse cross-linked during 6 hr at 65 °C and subjected to proteinase K (20 µg) treatment for 2 hr at 37 °C. DNA was extracted with phenol-chloroform-isoamyl alcohol and precipitated with ethanol and sodium acetate. DNA from input and CHIP samples were resuspended in 200 µL and 20 µL TE, respectively.

The amount of template used for PCR was 2 µL of a 1:100 dilution of input DNA and 2 µL of undiluted CHIP DNA. PCR was performed in the presence of 1 µM specific primers (see supplementary material 2), 200 µM dNTP, and 2 U of Taq polymerase. After denaturation at 95 °C for 2 min, 40 cycles were performed with annealing at 50 °C, elongation at 60 °C and denaturation at 95 °C, each step for 1 min. PCR products were analyzed on 1.2 % agarose gel.

Acknowledgements

We thank Dr. Christian Doerig for the kind gift of the anti-PfPfk5 serum and Dr. Jennifer Richardson for critical reading of the manuscript. We are grateful to Pr. Dominique Mazier (INSERM U511) for continuous encouragement and support. M. G., P. R. and C. B. were financially supported by the Ministère de l'Education Nationale, France. M.G. held a fellowship from Fondation pour la Recherche Médicale (FRM). The project was supported by INSERM funds to C. V.

References

1. World Health Organization. (2000). WHO Expert Committee on Malaria - Twentieth Report. WHO.
2. Demidov, V. V. (2003). Infection stage clues to new antimalarial medicines. *Drug Discov Today* 8, 913-5.
3. Gardner, M. J., Tettelin, H., Carucci, D. J., Cummings, L. M., Aravind, L., Koonin, E. V., Shallom, S., Mason, T., Yu, K., Fujii, C., Pederson, J., Shen, K., Jing, J., Aston, C., Lai, Z., Schwartz, D. C., Pertea, M., Salzberg, S., Zhou, L., Sutton, G. G., Clayton, R., White, O., Smith, H. O., Fraser, C. M., Hoffman, S. L. & et al. (1998). Chromosome 2 sequence of the human malaria parasite *Plasmodium falciparum*. *Science* 282, 1126-32.
4. Bowman, S., Lawson, D., Basham, D., Brown, D., Chillingworth, T., Churcher, C. M., Craig, A., Davies, R. M., Devlin, K., Feltwell, T., Gentles, S., Gwilliam, R., Hamlin, N., Harris, D., Holroyd, S., Hornsby, T., Horrocks, P., Jagels, K., Jassal, B., Kyes, S., McLean, J., Moule, S., Mungall, K., Murphy, L., Barrell, B. G. & et al. (1999). The complete nucleotide sequence of chromosome 3 of *Plasmodium falciparum* [see comments]. *Nature* 400, 532-8.
5. Gardner, M. J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R. W., Carlton, J. M., Pain, A., Nelson, K. E., Bowman, S., Paulsen, I. T., James, K., Eisen, J. A., Rutherford, K., Salzberg, S. L., Craig, A., Kyes, S., Chan, M. S., Nene, V., Shallom, S. J., Suh, B., Peterson, J., Angiuoli, S., Pertea, M., Allen, J., Selengut, J., Haft, D., Mather, M. W., Vaidya, A. B., Martin, D. M., Fairlamb, A. H., Fraunholz, M. J., Roos, D. S., Ralph, S. A., McFadden, G. I., Cummings, L. M., Subramanian, G. M., Mungall, C., Venter, J. C., Carucci, D. J., Hoffman, S. L., Newbold, C., Davis, R. W., Fraser, C. M. & Barrell, B. (2002). Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* 419, 498-511.
6. Horrocks, P., Dechering, K. & Lanzer, M. (1998). Control of gene expression in *Plasmodium falciparum*. *Mol Biochem Parasitol* 95, 171-81.
7. Levitt, A. (1993). RNA processing in malarial parasites. *Parasitol Today* 9, 465-8.

8. Cary, C., Lamont, D., Dalton, J. P. & Doerig, C. (1994). Plasmodium falciparum chromatin: nucleosomal organisation and histone-like proteins. *Parasitol Res* 80, 255-8.
9. Li, W. B., Bzik, D. J., Gu, H. M., Tanaka, M., Fox, B. A. & Inselburg, J. (1989). An enlarged largest subunit of *Plasmodium falciparum* RNA polymerase II defines conserved and variable RNA polymerase domains. *Nucleic Acids Res* 17, 9621-36.
10. Li, W. B., Bzik, D. J., Tanaka, M., Gu, H. M., Fox, B. A. & Inselburg, J. (1991). Characterization of the gene encoding the largest subunit of *Plasmodium falciparum* RNA polymerase III. *Mol Biochem Parasitol* 46, 229-39.
11. Fox, B. A., Li, W. B., Tanaka, M., Inselburg, J. & Bzik, D. J. (1993). Molecular characterization of the largest subunit of Plasmodium falciparum RNA polymerase I. *Mol Biochem Parasitol* 61, 37-48.
12. McAndrew, M. B., Read, M., Sims, P. F. & Hyde, J. E. (1993). Characterisation of the gene encoding an unusually divergent TATA- binding protein (TBP) from the extremely A+T-rich human malaria parasite Plasmodium falciparum. *Gene* 124, 165-71.
13. Hirtzlin, J., Farber, P. M. & Franklin, R. M. (1994). Isolation of a novel *Plasmodium falciparum* gene encoding a protein homologous to the Tat-binding protein family. *Eur J Biochem* 226, 673-80.
14. Horrocks, P. & Kilbey, B. J. (1996). Physical and functional mapping of the transcriptional start sites of Plasmodium falciparum proliferating cell nuclear antigen. *Mol Biochem Parasitol* 82, 207-15.
15. Lanzer, M., de Bruin, D. & Ravetch, J. V. (1992). Transcription mapping of a 100 kb locus of Plasmodium falciparum identifies an intergenic region in which transcription terminates and reinitiates. *Embo J* 11, 1949-55.
16. Osta, M., Gannoun-Zaki, L., Bonnefoy, S., Roy, C. & Vial, H. J. (2002). A 24 bp cis-acting element essential for the transcriptional activity of Plasmodium falciparum CDP-diacylglycerol synthase gene promoter. *Mol Biochem Parasitol* 121, 87-98.
17. Miller, L. H., Baruch, D. I., Marsh, K. & Doumbo, O. K. (2002). The pathogenic basis of malaria. *Nature* 415, 673-9.

18. Calderwood, M. S., Gannoun-Zaki, L., Wellem, T. E. & Deitsch, K. W. (2003). Plasmodium falciparum var genes are regulated by two regions with separate promoters, one upstream of the coding region and a second within the intron. *J Biol Chem* 278, 34125-32.
19. Dechering, K. J., Kaan, A. M., Mbacham, W., Wirth, D. F., Eling, W., Konings, R. N. & Stunnenberg, H. G. (1999). Isolation and functional characterization of two distinct sexual-stage-specific promoters of the human malaria parasite *Plasmodium falciparum*. *Mol Cell Biol* 19, 967-78.
20. Kanei-Ishii, C., Nomura, T., Ogata, K., Sarai, A., Yasukawa, T., Tashiro, S., Takahashi, T., Tanaka, Y. & Ishii, S. (1996). Structure and function of the proteins encoded by the *myb* gene family. *Curr Top Microbiol Immunol* 211, 89-98.
21. Lipsick, J. S. (1996). One billion years of Myb. *Oncogene* 13, 223-35.
22. Boschet, C., Gissot, M., Briquet, S., Hamid, Z., Claudel-Renard, C. & Vaquero, C. (2004). Characterization of PfMyb1 transcription factor during erythrocytic development of 3D7 and F12 Plasmodium falciparum clones. *Mol Biochem Parasitol* 138, 159-163.
23. Ness, S. A., Marknell, A. & Graf, T. (1989). The v-myb oncogene product binds to and activates the promyelocyte-specific *mim-1* gene. *Cell* 59, 1115-25.
24. Gissot, M., Refour, P., Briquet, S., Boschet, C., Coupe, S., Mazier, D. & Vaquero, C. (2004). Transcriptome of 3D7 and its gametocyte-less derivative F12 Plasmodium falciparum clones during erythrocytic development using a gene-specific microarray assigned to gene regulation, cell cycle and transcription factors. *Gene* 341, 267-77.
25. Tusher, V. G., Tibshirani, R. & Chu, G. (2001). Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A* 98, 5116-21.
26. Bozdech, Z., Llinas, M., Pulliam, B. L., Wong, E. D., Zhu, J. & DeRisi, J. L. (2003). The Transcriptome of the Intraerythrocytic Developmental Cycle of Plasmodium falciparum. *PLoS Biol* 1, 5.
27. Le Roch, K. G., Johnson, J. R., Florens, L., Zhou, Y., Santrosyan, A., Grainger, M., Yan, S. F., Williamson, K. C., Holder, A. A., Carucci, D. J., Yates, J. R., 3rd & Winzeler, E. A.

- (2004). Global analysis of transcript and protein levels across the *Plasmodium falciparum* life cycle. *Genome Res* 14, 2308-18.
28. Malhotra, P., Dasaradhi, P. V., Kumar, A., Mohammed, A., Agrawal, N., Bhatnagar, R. K. & Chauhan, V. S. (2002). Double-stranded RNA-mediated gene silencing of cysteine proteases (falcipain-1 and -2) of *Plasmodium falciparum*. *Mol Microbiol* 45, 1245-54.
29. McRobert, L. & McConkey, G. A. (2002). RNA interference (RNAi) inhibits growth of *Plasmodium falciparum*. *Mol Biochem Parasitol* 119, 273-8.
30. Noonpakdee, W., Pothikasikorn, J., Nimitsantiwong, W. & Wilairat, P. (2003). Inhibition of *Plasmodium falciparum* proliferation in vitro by antisense oligodeoxynucleotides against malarial topoisomerase II. *Biochem Biophys Res Commun* 302, 659-64.
31. Gunasekera, A. M., Patankar, S., Schug, J., Eisen, G., Kissinger, J., Roos, D. & Wirth, D. F. (2004). Widespread distribution of antisense transcripts in the *Plasmodium falciparum* genome. *Mol Biochem Parasitol* 136, 35-42.
32. Patankar, S., Munasinghe, A., Shoaibi, A., Cummings, L. M. & Wirth, D. F. (2001). Serial analysis of gene expression in *Plasmodium falciparum* reveals the global expression profile of erythrocytic stages and the presence of anti-sense transcripts in the malarial parasite. *Mol Biol Cell* 12, 3114-25.
33. Kyes, S., Christodoulou, Z., Pinches, R. & Newbold, C. (2002). Stage-specific merozoite surface protein 2 antisense transcripts in *Plasmodium falciparum*. *Mol Biochem Parasitol* 123, 79-83.
34. Ross-Macdonald, P. B., Graeser, R., Kappes, B., Franklin, R. & Williamson, D. H. (1994). Isolation and expression of a gene specifying a cdc2-like protein kinase from the human malaria parasite *Plasmodium falciparum*. *Eur J Biochem* 220, 693-701.
35. Morgan, D. O. (1997). Cyclin-dependent kinases: engines, clocks, and microprocessors. *Annu Rev Cell Dev Biol* 13, 261-91.
36. Graeser, R., Wernli, B., Franklin, R. M. & Kappes, B. (1996). *Plasmodium falciparum* protein kinase 5 and the malarial nuclear division cycles. *Mol Biochem Parasitol* 82, 37-49.

37. Graeser, R., Franklin, R. M. & Kappes, B. (1996). Mechanisms of activation of the cdc2-related kinase PfPK5 from *Plasmodium falciparum*. *Mol Biochem Parasitol* 79, 125-7.
38. Arnot, D. E. & Gull, K. (1998). The *Plasmodium* cell-cycle: facts and questions. *Ann Trop Med Parasitol* 92, 361-5.
39. Read, M., Sherwin, T., Holloway, S. P., Gull, K. & Hyde, J. E. (1993). Microtubular organization visualized by immunofluorescence microscopy during erythrocytic schizogony in *Plasmodium falciparum* and investigation of post-translational modifications of parasite tubulin. *Parasitology* 106 (Pt 3), 223-32.
40. Kelman, Z. (1997). PCNA: structure, functions and interactions. *Oncogene* 14, 629-40.
41. Kilbey, B. J., Fraser, I., McAleese, S., Goman, M. & Ridley, R. G. (1993). Molecular characterisation and stage-specific expression of proliferating cell nuclear antigen (PCNA) from the malarial parasite, *Plasmodium falciparum*. *Nucleic Acids Res* 21, 239-43.
42. Grall, M., Srivastava, I. K., Schmidt, M., Garcia, A. M., Mael, J. & Perrin, L. H. (1992). *Plasmodium falciparum*: identification and purification of the phosphoglycerate kinase of the malaria parasite. *Exp Parasitol* 75, 10-8.
43. Hovring, I., Bostad, A., Ording, E., Myrset, A. H. & Gabrielsen, O. S. (1994). DNA-binding domain and recognition sequence of the yeast BAS1 protein, a divergent member of the Myb family of transcription factors. *J Biol Chem* 269, 17663-9.
44. Bannister, A. J. & Miska, E. A. (2000). Regulation of gene expression by transcription factor acetylation. *Cell Mol Life Sci* 57, 1184-92.
45. Whitmarsh, A. J. & Davis, R. J. (2000). Regulation of transcription factor function by phosphorylation. *Cell Mol Life Sci* 57, 1172-83.
46. Trager, W. & Jensen, J. B. (1976). Human malaria parasites in continuous culture. *Science* 193, 673-5.
47. Lambros, C. & Vanderberg, J. P. (1979). Synchronization of *Plasmodium falciparum* erythrocytic stages in culture. *J Parasitol* 65, 418-20.

48. Desjardins, R. E., Canfield, C. J., Haynes, J. D. & Chulay, J. D. (1979). Quantitative assessment of antimalarial activity in vitro by a semiautomated microdilution technique. *Antimicrob Agents Chemother* 16, 710-8.
49. Makler, M. T., Lee, L. G. & Recktenwald, D. (1987). Thiazole orange: a new dye for Plasmodium species analysis. *Cytometry* 8, 568-70.
50. Merckx, A., Le Roch, K., Nivez, M. P., Dorin, D., Alano, P., Guiterrez, G. J., Nebreda, A. R., Goldring, D., Whittle, C., Patterson, S., Chakrabarti, D. & Doerig, C. (2003). Identification and initial characterization of three novel cyclin-related proteins of the human malaria parasite Plasmodium falciparum. *J Biol Chem*.
51. Schena, M., Shalon, D., Davis, R. W. & Brown, P. O. (1995). Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 270, 467-70.
52. Quandt, K., Frech, K., Karas, H., Wingender, E. & Werner, T. (1995). MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. *Nucleic Acids Res* 23, 4878-84.
53. Hecht, A., Strahl-Bolsinger, S. & Grunstein, M. (1999). Mapping DNA interaction sites of chromosomal proteins. Crosslinking studies in yeast. *Methods Mol Biol* 119, 469-79.

Abbreviations

The abbreviations used are: asRNA: anti-sense RNA; cDNA: DNA complementary to RNA; CHIP: chromatin immuno-precipitation; dsRNA: double stranded RNA; DNase: deoxyribonuclease; dNTP: deoxyribonucleoside triphosphate; MRE: Myb regulatory element; PCR: polymerase chain reaction; qPCR: real-time quantitative PCR; RNase: ribonuclease ; RT: reverse transcription; sRNA: sense RNA

Figure Legends

Figure 1. Parasite growth and death evaluation during trophozoite to schizont transition. (A). Evaluation of parasite growth 48 hr after addition of dsRNA to the media. Parasitaemia was either counted on GIEMSA stained blood smears (1 to 3) or monitored by ^3H -Hypoxanthine uptake (1 and 4). Parasitaemia was evaluated in untreated culture (1), unrelated dsRNA (2) and *pfmyb1* dsRNA treated culture (3 and 4). Asterisks indicate statistically significant values (p-value < 0.0001). Parasitaemia was evaluated in 9 independent experiments on GIEMSA stained blood smears and 2 independent experiments as regards ^3H -hypoxanthine uptake assay. **(B).** Parasitaemia of highly synchronized cultures was determined by counting at late trophozoite (1 and 2) and at late schizont stage (3 and 4) in untreated culture (1 and 3) and *pfmyb1* dsRNA treated culture (2 and 4). Asterisks are as in (A). The experiment was repeated 4 times in triplicate. The results are expressed as a percentage of the control cultures.

Figure 2. Analysis and quantification of *pfmyb1* mRNA expression. (A). Amplicons of *pfmyb1*, *18s* ribosomal RNA and MAL13P1.76 (indicated on the top) were obtained after semi-quantitative RT-PCR on DNase treated total RNA prepared from either untreated culture (control), irrelevant *gp120* dsRNA (unrelated dsRNA) or *pfmyb1* dsRNA (*pfmyb1* dsRNA) treated culture. Two independent experiments are presented. Markers are indicated on the left side of the figure. **(B).** Graphical representation of the *pfmyb1* amplified fragment intensities after normalization with *18s* ribosomal RNA. Semi-quantitative (1 to 3) and real time quantitative RT-PCR (4 to 6) were performed on samples from either untreated culture (1 and 4), or cultures treated with irrelevant (2 and 5) or *pfmyb1* dsRNA (3 and 6). The results are expressed as a percentage of the control transcript expression in 3 independent experiments.

Figure 3. Analysis and quantification of PfMyb1 protein expression. (A). Serial immunoprecipitation and Western blotting of *P. falciparum* PfMyb1. The protein extracts of trophozoite

parasites were first subjected to immuno-precipitation using beads coated with an anti-PfMyb1 serum followed by Western blot analysis with the same anti-PfMyb1 serum (lower panel). Concomitantly a Western blot analysis with an anti-PfHSP70 serum was performed (upper panel). The protein markers are indicated on the right of the figure. Samples from either untreated culture (control), or cultures treated with irrelevant (unrelated dsRNA) or *pfmyb1* dsRNA (*pfmyb1* dsRNA) were studied. **(B)**. Graphical representation of the PfMyb1 protein intensities after normalization with the intensities of the PfHSP70 protein : Lane 1, untreated culture; lane 3, irrelevant dsRNA treated culture; lane 2, *pfmyb1* dsRNA (*pfmyb1* dsRNA) treated culture. One representative experiment among three is presented in Figures 3A. The results are expressed as for Figure 2.

Figure 4. Graphical output of SAM software. Genes are plotted where the x-axis represents expected distribution of each PCR product intensity and the y-axis observed distribution of each PCR product intensity. Genes identified by SAM as up-regulated in *pfmyb1* dsRNA treated samples or as down-regulated in *pfmyb1* dsRNA treated samples are represented by triangles or circles, respectively, and are as follows: 1, PfPk5 (MAL13P1.279); 2, PF14_0366; 3, PFC0385C; 4, PFE0420c; 5, PfPGK (PFI1105w); 6, PfPk2 (PFL1885c); 7, PfTBP (PFL2345c); 8, PfPCNA1 (PFL1285c); 9, PP1-like (PF14_0224); 10, PfHistone H3 (MAL6P1.248); and 11, PfHistone H2A (MAL6P1.249).

Figure 5. Comparison of microarray and RT-qPCR data. Microarray data are indicated by open bars. RT-qPCR data are indicated by black bars. Relative transcriptional level of each gene (accession number indicated on bottom) is expressed as a \log_2 ratio of *pfmyb1* dsRNA treated samples over untreated samples.

Figure 6. PfPk5 expression in *pfmyb1* dsRNA treated samples. Trophozoite protein extracts were subjected to 10 % SDS-PAGE followed by Western blotting with either anti-PfHSP70 serum (upper panel) or anti-PfPk5 serum (lower panel) as indicated on the right side of the figure. The protein markers are indicated on the left side of the figure. Samples of untreated culture (control) and *pfmyb1*

dsRNA treated culture (*pfmyb1* dsRNA) were investigated. Figure 6 presents one representative experiment among three.

Figure 7. Chromatin immuno-precipitation (ChIP) using an anti-PfMyb1 serum. After DNA/PfMyb1 cross linking within the parasite nuclei and immuno-precipitation of DNA using anti-PfMyb1 antibody, PCR was carried out with input DNA (odd numbers) and immuno-precipitated DNA (even numbers) using primers amplifying promoters of the following genes: MAL13P1_279 (lane 1 and 2), PFI1105w (lane 3 and 4), PFL1885c (lane 5 and 6), PFL2345c (lane 7 and 8), PFL1285c (lane 9 and 10), MAL6P1.248 and MAL6P1.249 (lane 11 and 12), PF14_084 (lane 13 and 14), PF07_073 (lane 15 and 16), PF11_0096 (lane 17 and 18), PF08_0054 (lane 19 and 20) and PF10_0159 (lane 21 and 22). Amplicons were fractionated on 1.2 % agarose gel. Genes identified in previous experiments as displaying up-regulated (▲), down-regulated (▼) or unaltered expression (-) are indicated. The presence (+) or absence (-) of Myb related elements (MRE) in the promoter of the gene is indicated. Figure 7 portrays one representative experiment among three.

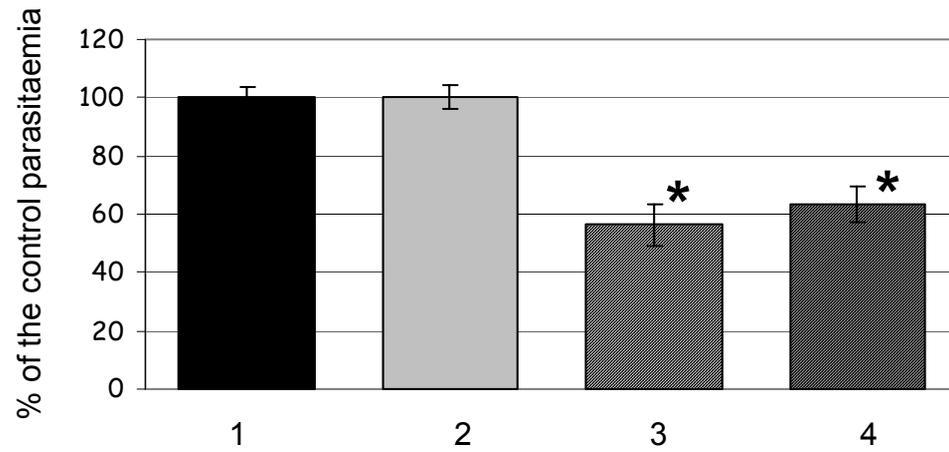
Supplementary data 1: microarray data used for SAM comparison of untreated samples and *pfmyb1* dsRNA treated samples.

Supplementary data 2: microarray data used for SAM comparison of untreated samples and unrelated dsRNA treated samples.

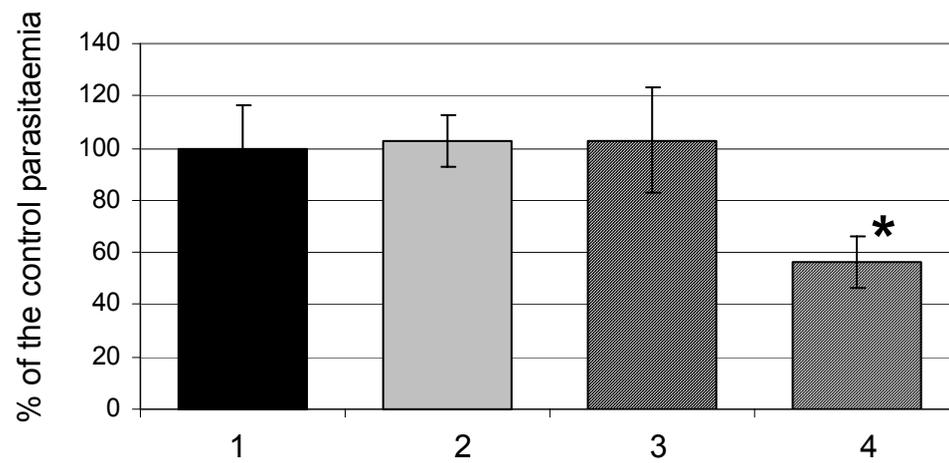
Supplementary material 1: primers used for qPCR experiments.

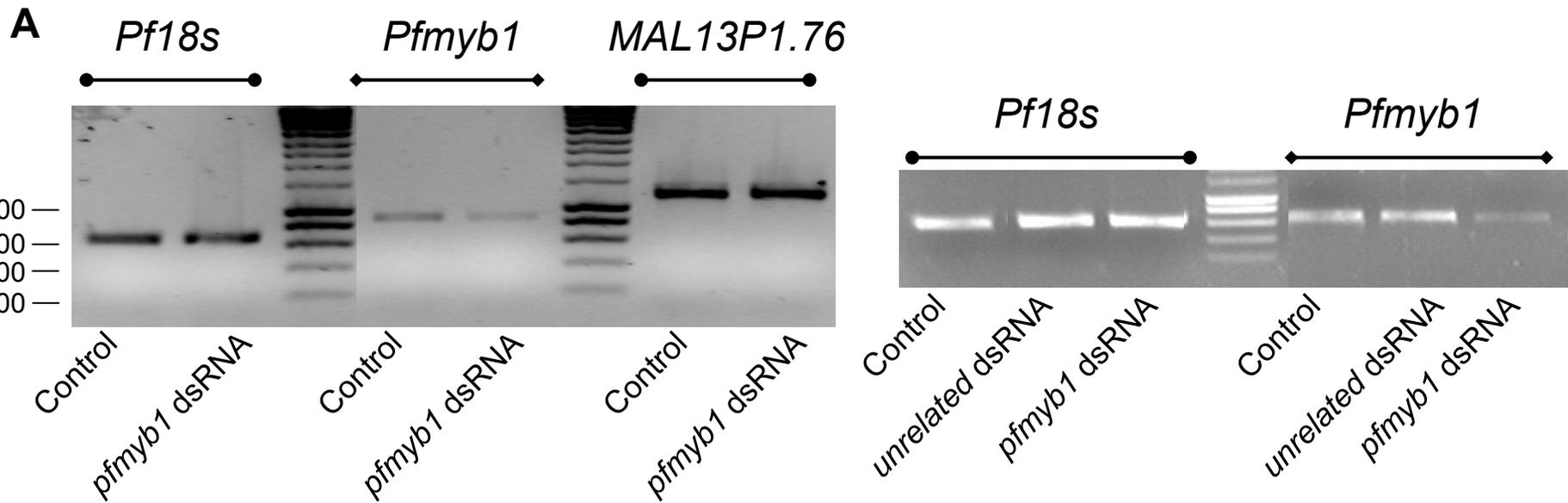
Supplementary material 2: primers used for ChIP experiments.

A

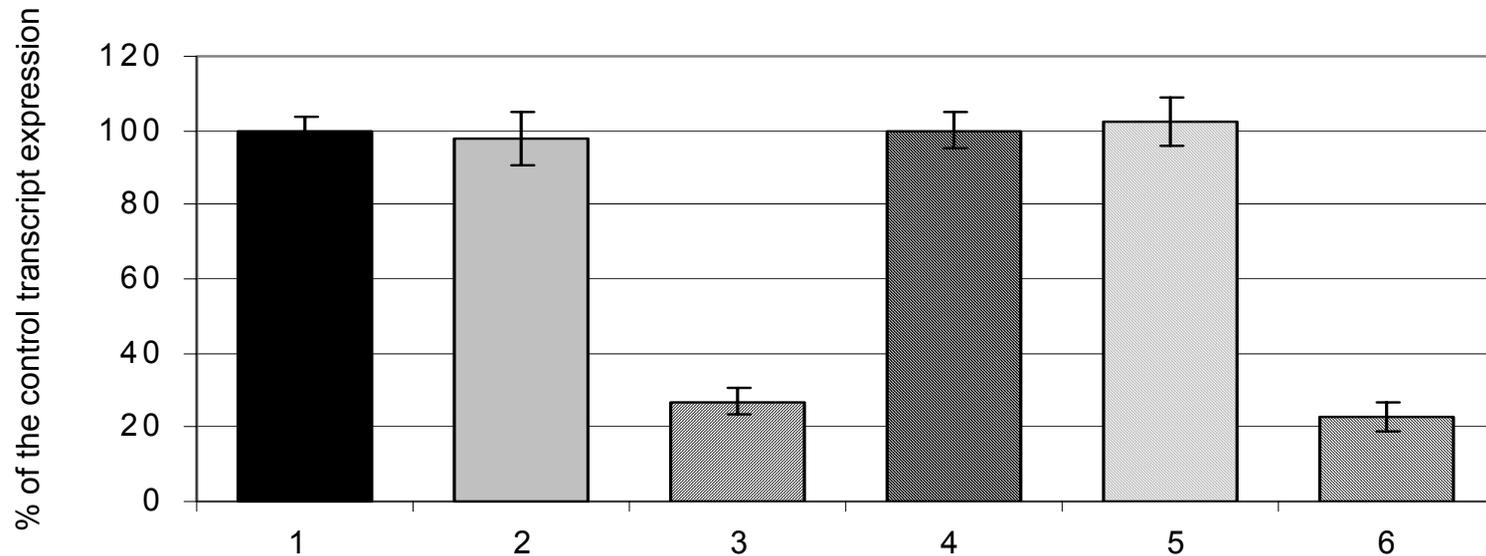


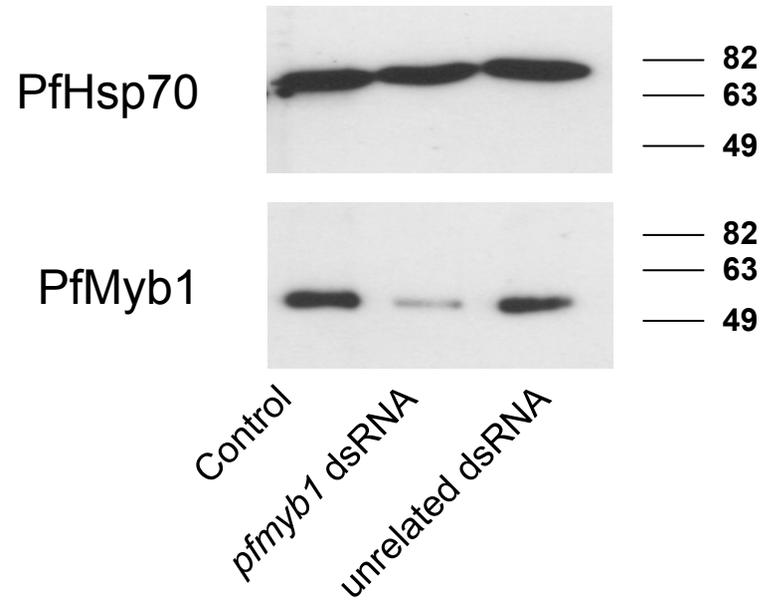
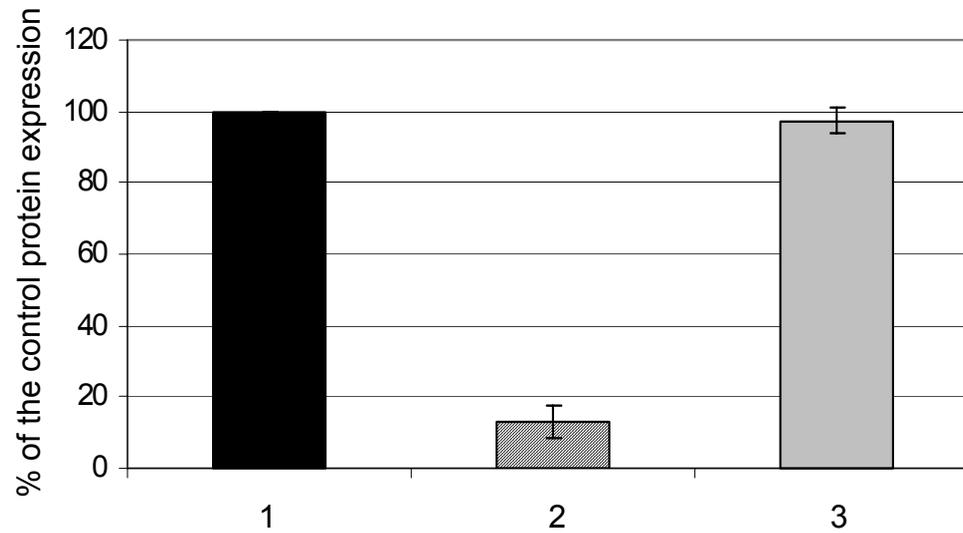
B

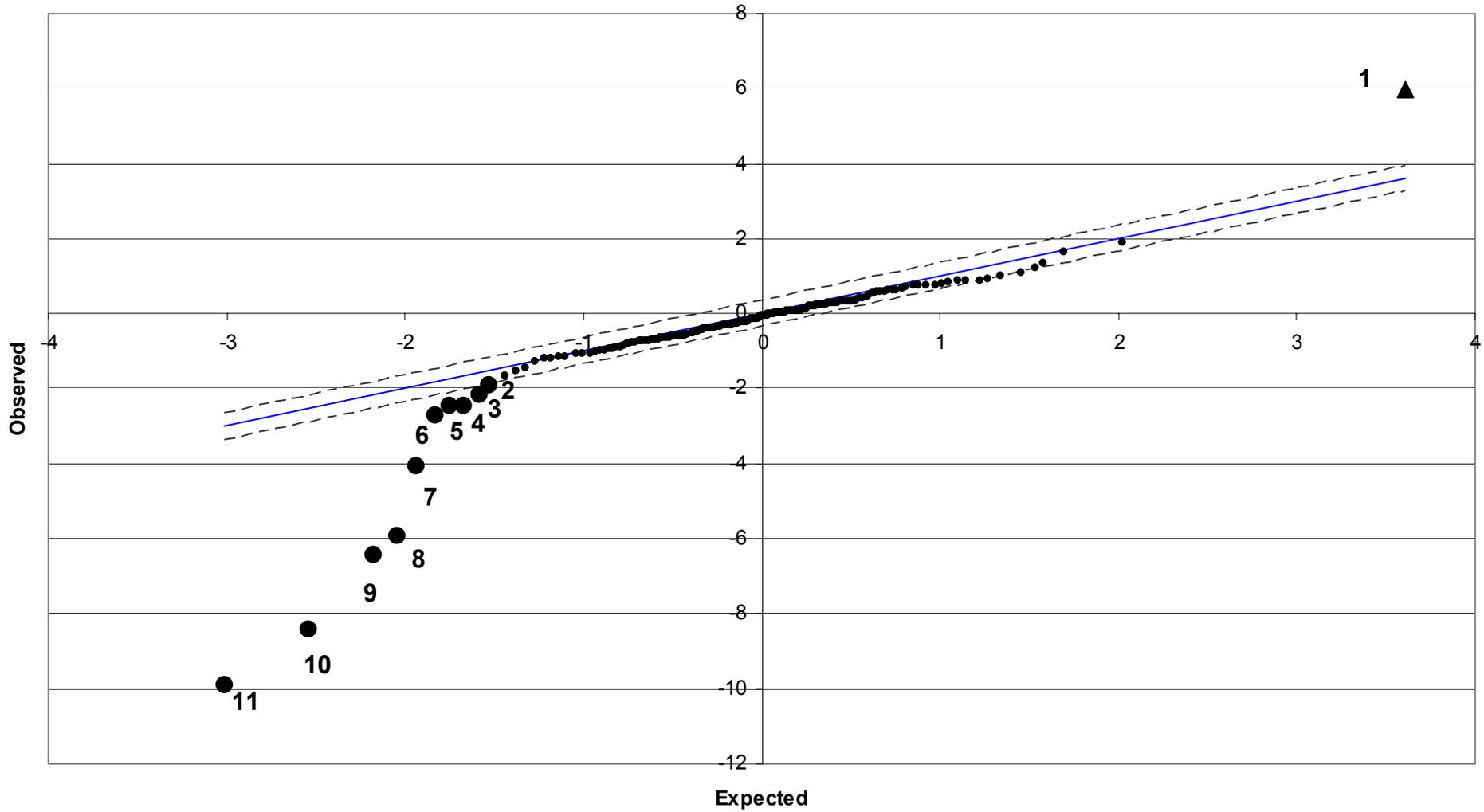


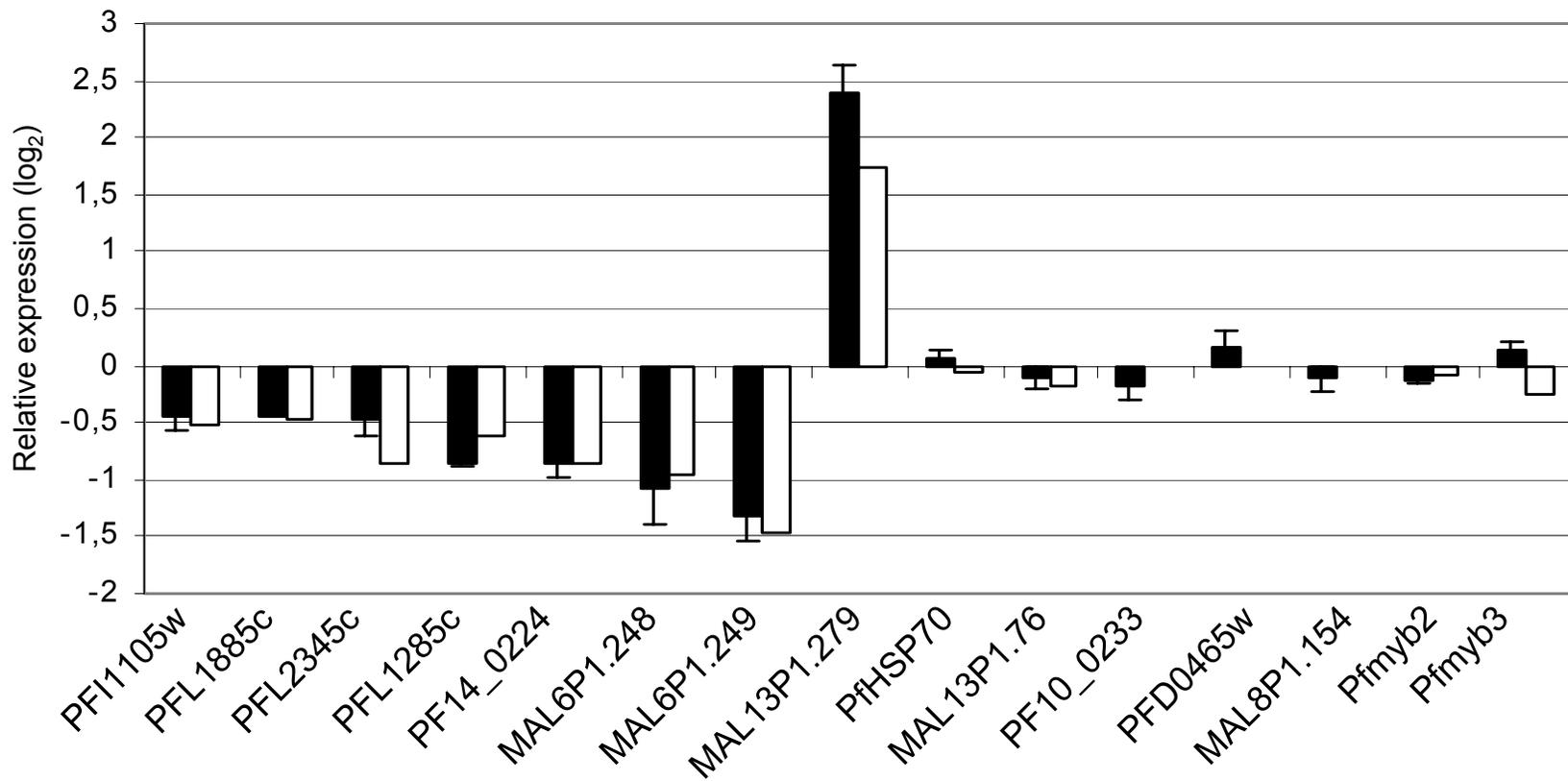


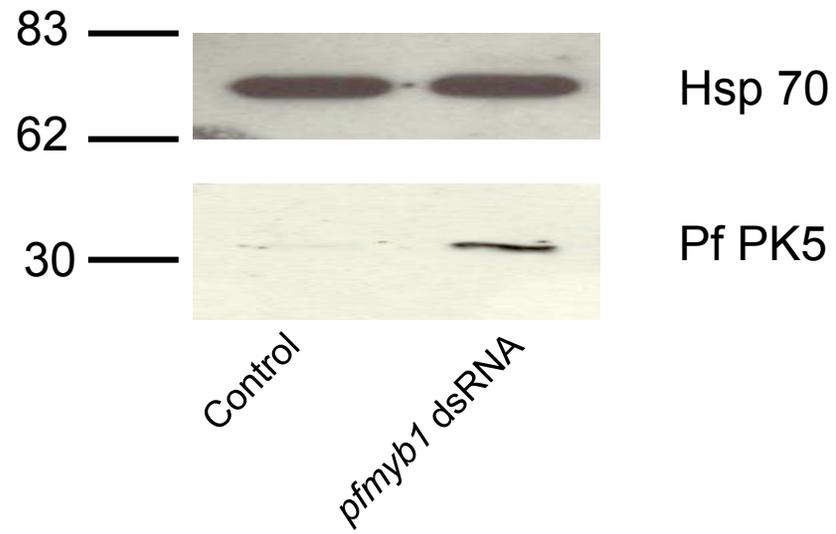
B



A**B**







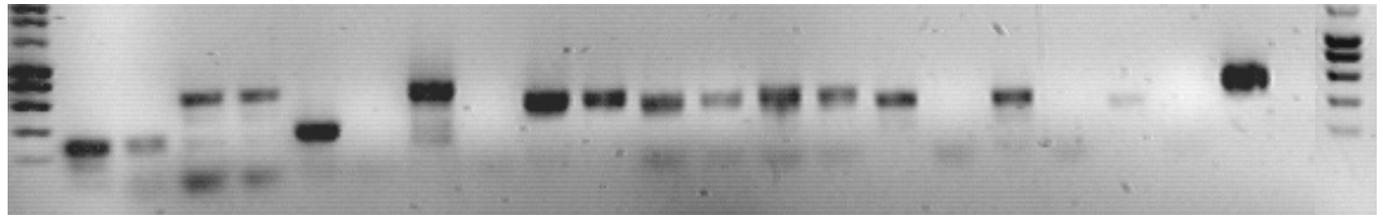
Altered expression in
pfmyb1 dsRNA treated
samples

Presence of a MRE

▲ ▼ ▼ ▼ ▼ ▼ - - - -
+ - + + + + + - + + +

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22

1000 —
600 —
400 —
200 —



DISCUSSION

Le cycle de *Plasmodium* décrit une succession d'étapes de prolifération et de différenciation. A chaque stade de développement, le parasite possède une double capacité: celle de répondre aux contraintes de l'environnement dans lequel il se trouve et celle de préparer les modifications qui lui permettront de répondre aux contraintes dans lesquelles il va se trouver au prochain stade de son cycle de vie. Cette habilité, résultant d'une coévolution entre le parasite et ses hôtes, fait de *P. falciparum* un organisme particulièrement performant pour envahir et se développer dans plusieurs tissus différents d'un vertébré et d'un invertébré.

La phase érythrocytaire de ce cycle complexe est, par bien des aspects, cruciale. Les formes érythrocytaires sont responsables de la pathologie et de la transmission du paludisme. Ces stades sont aussi ceux dont la manipulation est la plus aisée. Leur étude peut servir de base à la connaissance des mécanismes contrôlant la prolifération et l'initiation de la différenciation chez *P. falciparum*.

Au cours du cycle érythrocytaire, chaque étape de développement et de différenciation peut être caractérisée par la présence de transcrits et de protéines dont l'expression serait préférentielle à ces différents stades [138, 171]. Pour autant, le lien entre la définition des stades du point de vue morphologique et du profil d'expression caractéristique des gènes est à relativiser. En effet, une expérience de transcriptome, dans laquelle l'expression relative des transcrits du clone HB3 de *P. falciparum* a été mesurée heure par heure au cours du cycle érythrocytaire, a montré qu'il pourrait exister des vagues d'expression de transcrits qui ne sont pas restreintes aux stades définis morphologiquement [171]. Le contrôle fin de l'expression transcriptionnelle des gènes, chez *P. falciparum*, apparaît alors comme une évidence. Celle-ci est corroborée par l'analyse de quelques promoteurs et la découverte de leur structure bipartite canonique. De même, l'identification dans le génome de phases ouvertes de lecture codant pour de nombreux acteurs potentiels de la régulation transcriptionnelle des gènes, plaide pour un rôle crucial de ce contrôle dans le cycle de vie de *P. falciparum*. La diversité et la complexité des voies disponibles à *P. falciparum* afin d'assurer la régulation transcriptionnelle de ses gènes va à l'encontre des conclusions sur une prétendue pauvreté en acteurs de la régulation transcriptionnelle [58, 61], tirées lors des premières analyses bioinformatiques de la séquence complète du génome de *P. falciparum*. En effet, en plus des 53 phases ouvertes de lectures identifiées par *Coulson et collaborateurs* [58] (Annexe 1.2), notre analyse du génome révèle près de 110 nouvelles phases ouvertes de lectures (voir Annexe 1.1) auxquelles on peut assigner une fonction potentielle dans la régulation transcriptionnelle des gènes.

Cependant, ces différents constats n'ont pas été suivis par des études moléculaires plus approfondies des acteurs (les séquences de liaison des FT sur l'ADN et les FT eux-mêmes) et de leur implication dans le contrôle de l'initiation de la transcription chez *P. falciparum*. De plus, le rôle de ces protéines, dont la fonction reste encore putative, et leur implication à chaque stade du cycle érythrocytaire, n'ont pas fait l'objet d'investigations.

Au cours de ce travail de thèse, nous avons tenté d'appréhender les mécanismes impliqués dans le contrôle transcriptionnel de l'expression des gènes au cours du cycle érythrocytaire. Pour cela, nous avons étudié certains acteurs de cette régulation qui auraient un rôle dans les événements clefs du cycle érythrocytaire: la prolifération asexuée et la différenciation sexuée. Différents aspects des régulations génétiques conduisant soit à l'initiation de la différenciation, soit au maintien de la prolifération, ont fait l'objet de notre travail.

➤ ***L'initiation de la différenciation sexuée.***

Comme nous l'avons vu dans la partie Introduction, l'induction de la différenciation sexuée peut être induit par de nombreux facteurs environnementaux [27]. Parmi eux, l'implication d'un facteur autocrine a été suggérée, permettant la régulation de la balance entre les deux destins (prolifération ou différenciation) [31]. La susceptibilité de *P. falciparum* à ces facteurs diffusibles implique une capacité à relayer ce signal à l'intérieur du parasite afin de permettre une traduction phénotypique de ce signal d'initiation de la différenciation. Cette fonction est généralement assurée par des kinases dont l'implication dans la différenciation des gamétocytes mâles chez le moustique a déjà été décrite en réponse à un stimulus environnemental [38]. Ce signal doit aussi avoir une manifestation en terme de profil d'expression des transcrits et des protéines. Cette régulation semble être réalisée au moins en partie de manière transcriptionnelle, comme en témoigne le contrôle transcriptionnel de l'expression des deux premiers transcrits marquant la différenciation sexuée, *pfs16* et *pfg27* [245, 246]. Ceci implique la participation de facteurs de transcription dans la régulation de ces événements. De plus, la séquence du promoteur du gène *pfg27* n'apparaissant pas différente dans plusieurs clones de *P. falciparum* capables ou incapables de produire de gamétocytes [246], et l'absence de transcription de celui-ci dans les clones non producteurs, comme le clone F12, pourrait être dépendante de la disponibilité de certains FT. Même si la régulation de la présence des protéines lors du développement des gamétocytes semble être en partie dépendante de la régulation post-transcriptionnelle [174], les FT et leur fonction dans la

régulation de l'expression des gènes pourraient être impliqués dans l'initiation de la différenciation sexuée, débutée dans l'hôte vertébré.

Sur les bases de cette analyse nous avons utilisé la technique des puces à ADN afin d'étudier le transcriptome au cours du cycle érythrocytaire de deux clones de *P. falciparum*: 3D7, producteur de gamétocytes et F12, dérivant de 3D7 et incapable de produire des gamétocytes. Une puce à ADN contenant environ 200 fragments de PCR dont 150 sont représentatifs de gènes de *P. falciparum*, potentiellement impliqués dans la régulation transcriptionnelle, traductionnelle, dans la transduction du signal et le cycle cellulaire, a été construite dans notre laboratoire.

Le choix de restreindre l'analyse à ce nombre de gènes a été motivé par la pertinence particulière de ces voies de régulation dans le contrôle du phénomène étudié. La flexibilité de la technique et la simplicité d'analyse des données ont aussi été des arguments en faveur de l'utilisation de cette approche ciblée. L'emploi de la radioactivité a été aussi choisi car la limite de détection théorique de ce marqueur permettait une plus grande fiabilité des données pour des gènes dont l'expression risquait d'être faible comme pour les FT annotés par nos soins (voir l'article 4 en annexe 2). Cette étude a permis la validation biologique de la technique des puces à ADN couplée à l'utilisation de la radioactivité comme marqueur différentiel des ADNc. Cependant, ces données sur les profils d'expression des transcrits ne présument pas totalement du profil d'expression des protéines correspondantes. En effet, les modifications post-transcriptionnelles et post-traductionnelles peuvent modifier la présence et l'activité des protéines. Les données issues du transcriptome sont donc une première étape qui devra être suivie de l'analyse moléculaire des protéines.

Dans cette étude, près de 128 gènes (sur 153 gènes présents sur la puce à ADN) de *P. falciparum*, possédaient un profil d'expression similaire au cours du cycle érythrocytaire (annexe 3), ce qui n'apparaît pas étonnant au regard de la proximité génétique des deux clones (l'un dérivant d'une culture continue de l'autre). Parmi ces gènes figurent une dizaine de FT, dont la fonction semble importante lors du cycle asexué érythrocytaire car leur profil d'expression est conservé dans le clone HB3 [171], pourtant sans proximité génétique avec les deux précédents.

Huit autres phases ouvertes de lecture ont été identifiées comme ayant des profils d'expression différents. Parmi eux, une kinase et deux protéines ayant une fonction putative dans la régulation transcriptionnelle, ont particulièrement retenu notre attention.

- Le transcrit codant pour le gène *pfpk4* (MAL6P1.146) est exprimé de manière similaire tout au cours du cycle érythrocytaire du clone 3D7, tandis que le pic d'expression se trouve au stade trophozoïte pour le clone F12. Cette protéine a été annotée comme une enzyme proche des eIF-2 α kinases grâce à sa similitude avec la kinase HRI (heam regulated inhibitory kinase) [151]. En effet, l'activité de la protéine recombinante PfPk4 semble inhibée par des concentrations croissantes d'hémine [151] à l'instar de la kinase HRI. La distribution de la protéine est différente au cours du cycle érythrocytaire, avec en particulier deux formes de la protéines: une de 90 kDa présente dans les membranes du parasite et de l'érythrocyte et une de 80 kDa apparaissant restreinte au cytosol du parasite. Cependant, au moment de la segmentation du schizonte et lorsque les mérozoïtes sont formés, la protéine colocalise avec une protéine caractéristique des rhoptries, PfAMA1. Cette distribution implique que la protéine PfPk4 est présente et pourrait avoir un rôle lors des premières heures après l'invasion de l'érythrocyte. La production de gamétocyte paraît influencée par l'âge des érythrocytes [30]. Or, la concentration en hémine, est faible dans les jeunes érythrocytes et augmente avec l'âge des érythrocytes [247]. Il est donc possible que la présence de la protéine PfPk4 dans le jeune anneau puisse avoir un rôle en tant que capteur de l'état de l'environnement de l'érythrocyte, grâce à sa capacité d'intégrer les signaux produits par des facteurs environnementaux (comme la concentration en hémine libre). Cette fonction a d'ailleurs déjà été décrite pour des enzymes proches des eIF-2 α kinases de type HRI, chez *S. pombe*. En effet, dans cet organisme, ces protéines jouent un rôle dans la réponse à un choc thermique ou à la présence d'arsenic, donc au stress provenant de l'environnement [248]. Ces protéines agissent alors, en réponse au stress, en phosphorylant la protéine eIF-2 α , régulant ainsi son activité d'initiation de la traduction [249]. Chez *S. cerevisiae*, ce mécanisme constitue un moyen de régulation sélectif permettant le contrôle de la traduction d'un FT de la famille des bZip (GCN4) et ainsi de l'initiation de la transcription de ses cibles [250]. Chez *P. falciparum*, la phosphorylation du facteur eIF-2 α , par la protéine PfPk4, pourrait donc conduire à l'induction spécifique de la traduction d'un gène, même dans un contexte de réduction générale de la traduction, à l'instar d'autres eucaryotes [251]. L'induction spécifique de la traduction d'un ou plusieurs gènes pourrait amorcer ou réprimer une cascade d'événements conduisant à l'induction de la gamétocytogénèse. Ce mécanisme de sensibilité à l'environnement interne et aux stimuli externes à l'érythrocyte, pourrait permettre de faire le lien entre ces signaux de « stress » et l'induction de la gamétocytogénèse. Le dérèglement de ce contrôle, dans le cas d'une surexpression de PfPk4 dans le clone F12, pourrait alors

empêcher la production de gamétocyte. Cette hypothèse donnerait à la kinase PfPk4 un rôle de plateforme de réception des signaux de stress, dès l'invasion de l'érythrocyte, qu'elle pourrait alors transmettre au facteur eIF-2 α . Un rôle qui semble déjà joué par ce type de kinases chez d'autres organismes eucaryotes [250, 252].

L'équilibre, résultant de la somme des signaux inducteurs et répresseurs, contrôlant l'activité de PfPk4 sur eIF-2 α , pourrait être la clef du déclenchement de l'induction de la gamétocytogénèse. Afin de vérifier cette hypothèse, l'état de phosphorylation de eIF-2 α et/ou l'état d'activation de PfPk4 devraient être mesurés dans différentes conditions connues pour induire la gamétocytogénèse dans les deux clones de *P. falciparum*, 3D7 et F12. De plus, des expériences visant à moduler l'activité de la protéine PfPk4 (inhibiteurs spécifiques) pourraient être menées afin de connaître les effets de ces variations sur l'initiation de la gamétocytogénèse.

Par ailleurs, il est intéressant de noter que la plupart des gènes, dont l'expression semble contrôlée de manière post-transcriptionnelle, aurait un rôle durant le développement sexué [174]. Cette donnée souligne l'importance de la régulation de la traduction dans la poursuite du processus de gamétocytogénèse.

Deux autres transcrits, pouvant être impliqués dans la régulation de l'initiation de la transcription, ont été identifiés. Leurs profils d'expression sont différents dans les clones 3D7 et F12.

- Le gène *pfb7* (PFI0930c) code pour une protéine possédant un motif de type NAP (Nucleosome Assembly Protein). Chez les eucaryotes, ces protéines sont impliquées dans le transport des histones du lieu de leur synthèse jusqu'au noyau [253]. L'interaction de certaines de ces protéines, avec la machinerie de remodelage de la chromatine, indique une possible participation à la régulation de l'expression des gènes [254]. Chez *S. cerevisiae*, la perturbation du locus codant pour la protéine Nap1 entraîne la dérégulation de l'expression de 10 % des gènes dont plus de 35 % se trouvent regroupés physiquement dans le génome [255]. Cette expérience montre que ces protéines pourraient être impliquées dans le maintien d'un état dynamique de la chromatine, permettant ainsi la régulation de l'expression de certains loci. De plus, ces protéines semblent jouer un rôle dans la régulation du cycle cellulaire et plus particulièrement dans le déroulement de la mitose comme l'indique sa capacité d'interaction avec les cyclines [256]. Ces dernières ayant des représentants dans le génome de *P. falciparum* [48], cette fonction pourrait être conservée dans cet organisme. Cette famille de

protéines pourrait alors servir de lien entre le cycle cellulaire et la localisation des histones. De plus, la localisation de la protéine PfB7 semble être nucléaire [257].

Chez *P. falciparum*, le transcrit *pfb7* est de taille différente au stade érythrocytaire asexué ou gamétocyte à l'instar de *P. berghei* [199]. Chez le parasite murin, cette différence est expliquée par un site d'initiation de la transcription distinct pour ces deux stades, associé à un épissage alternatif, conduisant à la formation de régions 5'UTR plus longues au stade gamétocyte [199]. Cependant, la séquence de la phase ouverte de lecture semble identique pour les deux transcrits. Ces données impliquent la possibilité de l'action de deux mécanismes régulant la présence de cette protéine dans les deux stades: d'une part la régulation de l'initiation de la transcription grâce à deux promoteurs différents et d'autre part une régulation post-transcriptionnelle au niveau des régions 5'UTR. Ces différents contrôles, lors du processus de gamétocytogénèse, ainsi que la fonction putative de cette protéine dans la régulation de la transcription et du cycle cellulaire, indiquent que PfB7 jouerait un rôle dans l'incapacité du clone F12 à produire des gamétocytes.

Il serait intéressant de clarifier les raisons pour lesquelles la cinétique d'expression des transcrits *pfb7* est différente dans les deux clones analysés. L'évaluation de l'implication du transcrit spécifique des gamétocytes, dont la mesure de l'expression relative est indissociable de celle du transcrit asexué (du fait du produit de PCR utilisé), permettrait de mieux comprendre son rôle dans le phénotype du clone F12. Dans le même temps, déterminer les séquences dont dépend l'activité du promoteur sexué et asexué du gène *pfb7* pourrait nous permettre d'identifier des acteurs spécifiques de la transcription au stade sexué.

- Le gène *constans* (PF14_0383) code pour une protéine possédant deux motifs en doigt de zinc de type B-box. Ces motifs composant cette protéine sont particulièrement proches du consensus défini par l'alignement de plusieurs protéines qui sont responsables de la régulation transcriptionnelle de la floraison chez les plantes [97]. Il est intéressant de noter que des acides aminés situés dans les domaines B-box, dont la mutation abolit la fonction de ces protéines [258], sont conservés dans la séquence de la protéine de *Plasmodium*. Chez *A. thaliana*, parmi les protéines responsables de la régulation transcriptionnelle de la floraison, la protéine CONSTANS (CO), a été tout particulièrement étudiée. CO est capable de promouvoir l'expression de groupes de gènes [98]. Cependant, les modalités d'action de CO restent encore indéterminées. Cette protéine ne se liant pas à l'ADN, son action dépend vraisemblablement d'interactions protéines-protéines, dont la responsabilité incomberait aux domaines B-box [98]. A l'instar de ces protéines, PF14_0383 pourrait avoir un rôle dans la régulation transcriptionnelle. Cette fonction reste évidemment à démontrer, même si l'analyse

bioinformatique de sa séquence révèle des similitudes avec les protéines de la famille de CO. Des expériences visant à augmenter ou à empêcher l'expression de cette protéine pourrait nous instruire sur son rôle dans le processus de gamétocytogénèse.

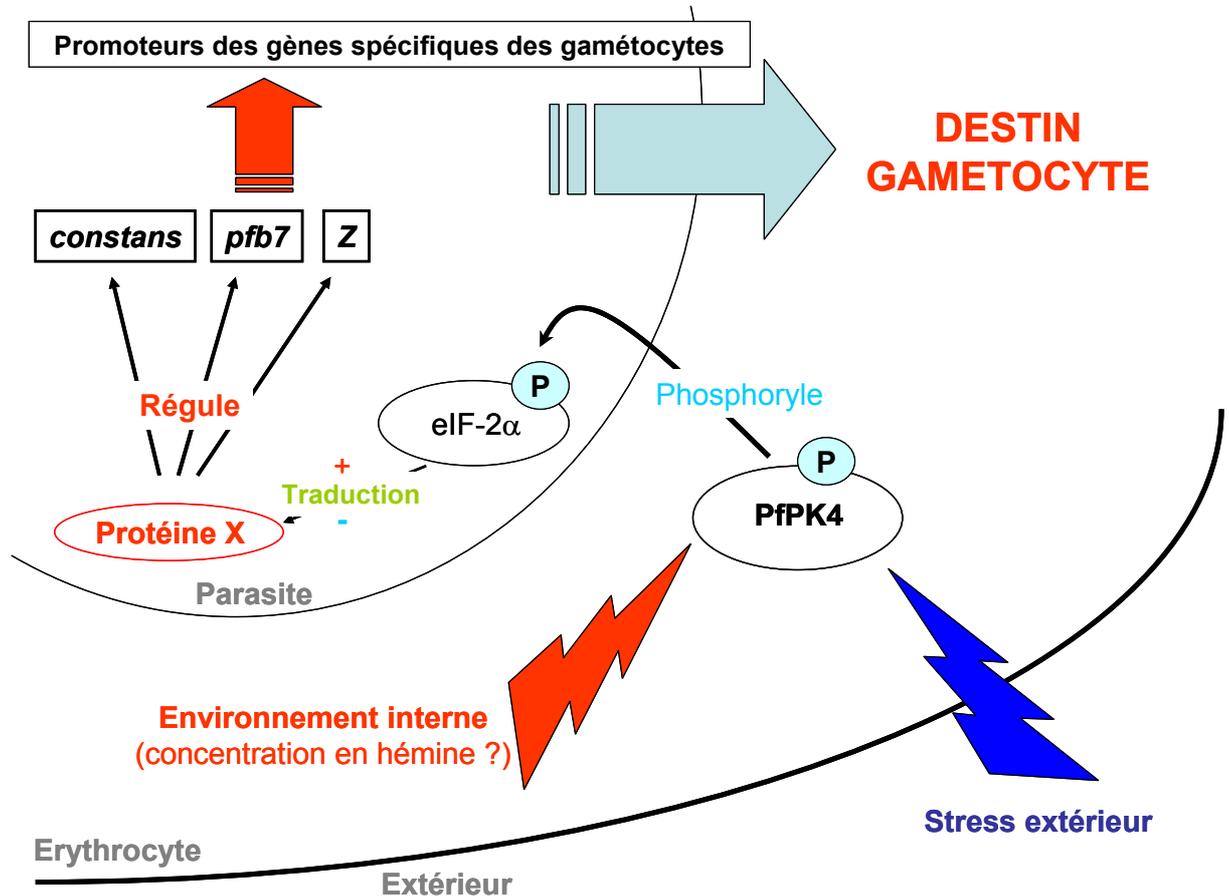


Figure 19. Hypothèse impliquant les gènes identifiés dans la régulation de la gamétocytogénèse.
 X et Z représentent des protéines inconnues.

Ces données sont le premier pas vers la découverte de gènes dont la fonction serait impliquée dans la régulation de l'initiation de la gamétocytogénèse. Elles seraient idéalement complétées par le transcriptome de ces mêmes clones en utilisant des puces à ADN comprenant la totalité des phases ouvertes de lecture. Ces expériences, qui ont été amorcées grâce à la participation du Consortium Plateforme Post-génomique Pasteur *Plasmodium* (P4), compléteront les données que nous avons produites. Les résultats sont en cours d'analyse en collaboration avec Peter David et Emmanuel Bishof. Les premiers résultats de ces expériences révèlent que de nombreux gènes ont une expression relative réduite dans le clone F12 comparativement au clone 3D7, en plus des gènes *pfg27* et *pfs16* déjà identifiés. Ces gènes

dont la plupart semblent posséder un pic d'expression au stade gamétoocyte [138], sont autant de gènes potentiellement impliqués dans la gamétoctogénèse. Par ailleurs, les promoteurs de ces gènes seraient particulièrement intéressants à étudier afin de découvrir des motifs conservés responsables de l'expression des gènes qu'ils dirigent. Enfin, l'étude de l'ensemble des données du transcriptome, issu de la puce pan-génomique, nous permettra de grouper les gènes selon leurs profils d'expression et de tenter d'identifier dans les promoteurs correspondants des motifs conservés et qui seraient impliqués dans leur régulation transcriptionnelle.

En conclusion, ces données issues de l'étude comparative du transcriptome ouvrent des perspectives en terme d'intégration du signal de « stress » et de régulation transcriptionnelle de la gamétoctogénèse (Figure 19). Ces expériences permettent d'envisager l'implication possible des facteurs identifiés dans ce processus. Toutefois, leur rôle dans la régulation transcriptionnelle de la gamétoctogénèse, reste encore à être démontré.

➤ *La prolifération asexuée.*

Nous avons ensuite entrepris d'étudier un acteur de la régulation transcriptionnelle: un des facteurs de transcription qui avaient été annotés dans l'équipe. Cette étude a alors débuté par la réalisation de la caractérisation moléculaire d'une protéine possédant des domaines caractéristiques de la famille des facteurs de transcription à domaines tryptophane. L'étude du profil d'expression du transcrit *pfmyb1* a conduit à identifier son pic d'expression au stade trophozoïte pour les deux clones 3D7 et F12, de même que pour le clone HB3 [171]. Nous avons démontré, dans des expériences complémentaires (p.94), que la cinétique d'expression de la protéine PfMyb1 est similaire à celle de son transcrit dans les deux clones étudiés. Ces données, en plus de la conservation de PfMyb1 dans les différentes espèces de *Plasmodium*, semblent souligner l'importance de cette protéine pour le cycle érythrocytaire asexué.

De plus, nous avons pu montrer que cette protéine, présente dans les extraits nucléaires, était capable de se lier aux séquences MRE consensus. En effet, la cinétique d'expression de la protéine PfMyb1 pour le clone 3D7 (Figure 3, résultats complémentaires, p.96) est proche de celle obtenue par l'analyse des bandes retardées, corroborant l'hypothèse d'une participation de PfMyb1 à ce complexe protéine-ADN. Cette dernière a été confirmée grâce à l'emploi d'un anticorps anti-Myb1. Ces expériences montrent que la protéine PfMyb1 possède toutes

les caractéristiques d'un vrai facteur de transcription. Les MRE au sein du génome de *P. falciparum* ont une distribution préférentielle dans les régions intergéniques par rapport à une séquence de même composition en AT. Cependant, ces MRE sont présents aussi dans les phases ouvertes de lecture. *In vivo*, les FT se fixent préférentiellement sur les sites intergéniques, même si les déterminants de cette sélectivité sont encore inconnus [259]. Il est donc possible que la protéine PfMyb1 possède aussi cette spécificité.

L'apparition de plusieurs bandes (une majoritaire et deux minoritaires) lors des expériences de retard sur gel suggère la présence de partenaires lors de la liaison au MRE (Figure 2, article 2, p. 89). Dans la littérature, de nombreux partenaires pour les protéines Myb ont été décrits. Parmi eux, les FT possédant un domaine bZip et d'autres membres de la famille Myb sont présents dans le génome de *P. falciparum*. Il est intéressant de noter que la coopération avec les FT à domaine bZip peut nécessiter l'existence d'un site de liaison à ce facteur dont la position par rapport à celle du MRE est déterminante pour leur interaction [260]. Cependant, l'interaction semble possible sans la présence d'ADN [260].

Ces partenaires putatifs ont d'autant plus importance que la cinétique d'interaction avec les extraits nucléaires des clones 3D7 et F12 est différente. Les cinétiques d'expression de la protéine PfMyb1 étant similaires dans les deux clones, la disparition de l'interaction avec le MRE dans les stades anneau et schizonte du clone F12 pourrait être causée par l'absence d'un ou plusieurs partenaires déterminant pour la formation du complexe protéine-MRE ou bien par des modifications post-traductionnelles modifiant l'interaction du FT avec l'ADN. Ces partenaires seraient présents dans le clone 3D7 et absents dans le clone F12. Cette différence est d'autant plus importante qu'elle pourrait être une des causes du phénotype du clone F12. Dans ce cadre, la recherche de partenaires de PfMyb1 prend d'autant plus d'importance qu'ils pourraient être impliqués dans le processus de gamétocytogénèse. Ces partenaires, dont la fonction se situe sans doute dans la régulation de la transcription, pourraient expliquer l'abolition de l'expression de certains gènes, comme *pfg27*. Les techniques d'immuno-précipitation couplées à la spectrométrie de masse seraient particulièrement intéressantes afin de conduire cette analyse des partenaires de PfMyb1. Par ailleurs, l'action de modifications post-traductionnelles ne peut être exclue, ces dernières pouvant réguler la liaison de PfMyb1 à l'ADN.

En conclusion, cette étude a permis de mettre en évidence, pour la première fois chez *P. falciparum*, la présence d'un facteur de transcription spécifique de séquence et fonctionnel au cours du cycle érythrocytaire.

Afin de continuer l'étude de ce facteur de transcription, nous avons tenté, dans l'article 3 (p.98), d'interférer avec l'expression de son transcrit et de sa protéine en utilisant des ARN double brin spécifiques d'une partie du gène *pfmyb1*. L'incubation dans le milieu de culture d'ARNdb de *pfmyb1* entraîne une mort prématurée des parasites lors de la transition du trophozoïte au schizonte. Cette mort serait en partie la conséquence de la baisse de représentativité du transcrit et de la protéine PfMyb1 dans les cultures traitées par l'ARNdb *pfmyb1*. Ces données, en plus de la conservation du profil d'expression de *pfmyb1* dans plusieurs clones de *P. falciparum*, ajoutées à la conservation de la séquence en aa de PfMyb1 chez *Plasmodium spp*, sont autant d'arguments en faveur d'un rôle crucial de PfMyb1 pour la survie lors du cycle érythrocytaire de *P. falciparum*.

Le mécanisme par lequel ces ARNdb agissent est encore largement inconnu, mais les ARNdb ont été utilisés efficacement dans un certain nombre d'expériences chez *P. falciparum* [235, 236, 261]. La controverse autour de plusieurs études indiquant une voie d'action par ARN interférence [236, 262], est née de l'incapacité actuelle à trouver, dans le génome de *P. falciparum*, les enzymes clefs de ce mécanisme [61]. Cependant, il est possible, pour les raisons que nous avons invoquées en introduction, que ces activités existent mais n'aient pas encore été mises à jour. La découverte chez *P. falciparum* d'une grande occurrence de transcrits anti-sens [167, 169, 232] a relancé le débat sur une éventuelle participation de ces anti-sens à la régulation transcriptionnelle. Si tel était le cas, l'existence d'une machinerie capable de prendre en charge ces transcrits anti-sens et les ARNdb résultant de la présence des deux brins d'un même transcrit, serait alors une hypothèse recevable. D'autre part, la méthylation de l'ADN et des histones a été décrite comme un des moyens possibles d'extinction de l'expression de gènes en présence d'ARNdb [263]. Ce mécanisme emploie des enzymes dont la présence a été identifiée dans le génome de *P. falciparum*, tel que des histones méthylases et des cytosine-C5 méthylases [264]. Il serait alors intéressant de vérifier l'état de méthylation du promoteur de *pfmyb1* dans le contexte d'exposition de la culture aux ARNdb.

Par ailleurs, il existe de nombreux mécanismes par lesquels un ARN anti-sens peut agir, comme nous l'avons montré en introduction [226]. La baisse d'expression observée pourrait donc être due à un effet anti-sens. Les oligonucléotides anti-sens ont d'ailleurs été largement utilisés chez *P. falciparum* et ont montré leur efficacité sur le parasite [265]. De plus, l'introduction d'un vecteur exprimant une partie du transcrit anti-sens du gène *clag9*, impliqué dans la cytoadhérence [266], entraîne la baisse de l'expression du transcrit et de la protéine

Clag9 [267]. Cette inhibition partielle de l'expression a pour conséquence l'apparition d'un phénotype de cytoadhérence réduite [267]. La disparition du plasmide, après avoir levé la pression de sélection, permet un retour au phénotype sauvage [267]. L'état actuel des connaissances sur ce sujet n'autorise pas la compréhension des mécanismes responsables de l'inhibition de l'expression des transcrits et de la protéine, mais l'outil reste utile, et pour l'instant le seul disponible, dans le contexte de gènes dont l'expression est essentielle pour le cycle érythrocytaire asexué.

Dans ce cadre, nous avons choisi cette approche expérimentale afin d'étudier les conséquences de l'inhibition partielle de l'expression du transcrit et de sa protéine. Nous avons alors utilisé la puce à ADN ciblée. Celle-ci a permis d'identifier 8 gènes dont l'expression était altérée dans un contexte de sous expression de la protéine PfMyb1. Afin de s'assurer de la spécificité de l'effet des ARNdb, l'expression de plusieurs gènes a été comparée dans les extraits de cultures contrôles et de cultures traitées. Les gènes *pfmyb2* et *pfmyb3*, de la famille des facteurs de transcription à domaines tryptophane, ainsi que 3 gènes dont la séquence en acide nucléique comportaient une similitude faible mais détectable à la séquence utilisée pour générer l'ARNdb *pfmyb1*, ont été utilisés comme témoins. L'expression de ces gènes étant similaire dans les deux conditions, il semble que l'action de ces ARNdb soit spécifique de séquence.

Parmi ces 8 gènes, la surexpression du messenger *pfpk5* et la sous-expression du transcrit *pcna1* pourraient permettre d'expliquer la mort des parasites lors de la transition trophozoïte schizonte. En effet, les protéines produites par ces deux gènes seraient responsables de la régulation du cycle cellulaire. La protéine PfPk5, dont la surexpression est vérifiée au niveau protéique, paraît impliquée dans la régulation de la progression du cycle et de la synthèse d'ADN chez *P. falciparum* [49], à l'instar de la protéine PCNA chez les eucaryotes [268].

Les 131 gènes dont l'expression est similaire dans les deux conditions, sont autant de marqueurs de la synchronisation de la culture, permettant ainsi de penser que l'altération de l'expression des 8 gènes dans les cultures traitées est bien la conséquence de la baisse de disponibilité de la protéine PfMyb1 et non pas le produit d'un décalage de croissance entre les deux cultures. L'expression de ces 8 gènes serait donc sous la dépendance directe ou indirecte du facteur de transcription PfMyb1. Il est d'ailleurs intéressant de noter que parmi ces gènes, les gènes *histone h2A* et *h3*, partagent le même promoteur bidirectionnel.

Afin d'identifier les gènes dont l'expression serait directement régulée par PfMyb1, nous avons réalisé des expériences d'immuno-précipitation de la chromatine. Parmi les 8 gènes dont l'expression est altérée en présence d'ARNdb de *pfmyb1* dans la culture, 6 de leurs

promoteurs semblent interagir avec la protéine PfMyb1 et pourraient ainsi être régulés directement par PfMyb1. Cette donnée indique clairement qu'*ex vivo* la protéine PfMyb1 joue bien le rôle d'un facteur de transcription et en se liant à l'ADN dans le noyau du parasite, participe à la régulation transcriptionnelle de certains gènes.

Les promoteurs de deux gènes, dont l'expression était altérée dans un contexte de baisse de disponibilité de PfMyb1, n'ont pu être mis en évidence dans ces expériences. Dans ces deux cas, l'interaction PfMyb1-promoteur serait trop faible pour en permettre la détection par la technique employée. De même, leur régulation pourrait être assurée de manière indirecte par PfMyb1, contrôlant lui-même un autre facteur de transcription qui serait responsable du contrôle des promoteurs de ces deux gènes. En ce qui concerne l'expression des FT présents sur la puce, elle ne semble pas soumise à une régulation par PfMyb1.

Les 5 promoteurs interagissant avec PfMyb1 possèdent tous un MRE à l'exception notable du promoteur du gène *pgk*. La capacité de PfMyb1 de se lier au MRE a déjà été décrite dans l'article 2 (p.89). De plus, des protéines capables de lier les MRE ont été plus récemment identifiées dans les extraits nucléaires de *P. falciparum* [205]. L'analyse bioinformatique de la séquence de ce promoteur du gène *pgk* n'a pas pu nous permettre d'identifier des séquences correspondant au consensus du MRE en dépit de la présence de PfMyb1 à son promoteur. Cette donnée impliquerait que PfMyb1 possède d'autres moyens de se lier à l'ADN que par le biais d'un MRE. Cette caractéristique a déjà été décrite pour une protéine de la famille Myb, Bas1p de *S. cerevisiae* [107], mais aussi pour d'autres protéines de ce même organisme contenant des domaines de liaison à l'ADN de type Myb [269-271]. De plus, la liaison de PfMyb1 au promoteur serait dépendante de la présence d'autres facteurs [260]. Les connaissances apportées par des expériences d'identification par immuno-précipitation de la chromatine, dans tout le génome de *S. cerevisiae*, des sites de liaison de la protéine de type Myb, devraient être applicables à PfMyb1. Dans cet article, les auteurs rapportent que Rap1 se fixe quasi-exclusivement à des promoteurs contenant son site de fixation, mais peut se lier à des loci ne possédant pas ce site [259]. Dans ce cas, la coopération avec d'autres FT constituerait une des raisons expliquant cette observation [259, 272]. Cette flexibilité quant à la liaison à un promoteur ne possédant pas de site de liaison spécifique pourrait être aussi une caractéristique de PfMyb1.

Le fait que PfMyb1 régule directement le gène *pfpk5* et les 5 autres gènes implique qu'il possède la capacité d'engendrer soit l'activation, soit la répression de la transcription. Cette donnée souligne, encore une fois, l'importance des protéines interagissant avec PfMyb1.

Celles-ci orienteraient positivement ou négativement l'activité transcriptionnelle de PfMyb1. De même, une modification post-traductionnelle de PfMyb1 pourrait permettre de réguler son activité, vers une activation ou une répression de la transcription.

En conclusion, cette étude nous a permis de caractériser un facteur de transcription spécifique de séquence chez *P. falciparum*. Nous avons pu aussi comprendre l'importance de son rôle lors du cycle érythrocytaire, et notamment dans la régulation de gènes clés pour le cycle cellulaire (Figure 20).

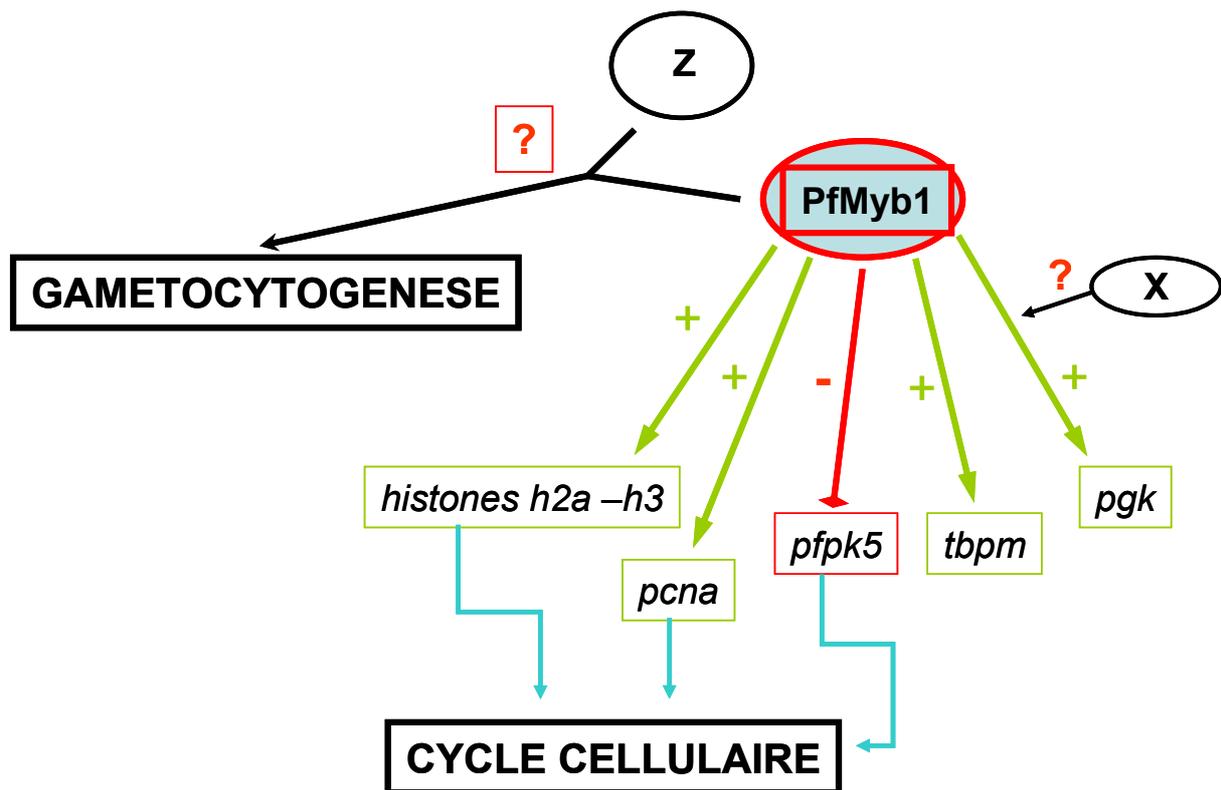


Figure 20. Ebauche du réseau de régulation de PfMyb1 et de son implication dans les événements clés du cycle érythrocytaire. X et Z représentent des protéines inconnues. Les flèches vertes avec un signe (+) désignent une régulation positive de la transcription du gène par PfMyb1. La flèche rouge avec un signe (-) désigne une régulation négative de la transcription du gène par PfMyb1.

Ces études soulignent aussi que la régulation des gènes chez *P. falciparum* possède des spécificités, à l'image de la liaison de PfMyb1 à l'ADN dont les déterminants restent encore flous. De plus, ce travail révèle la première ébauche de l'élaboration du réseau de régulation chez *P. falciparum*. Un travail identique pourra être mené en utilisant des puces à ADN représentant tout le génome de *P. falciparum*, permettant ainsi d'accumuler plus de données

Discussion

sur les promoteurs des gènes régulés directement par PfMyb1. Cette expérience permettra peut être de mieux comprendre les déterminants de la liaison de PfMyb1 à l'ADN.

Enfin, l'ensemble de ce travail constitue les premières étapes vers la compréhension des modes de régulation transcriptionnelle des événements cruciaux du cycle érythrocytaire de *P. falciparum*. Il permet d'envisager ce contrôle dans sa complexité et sa spécificité, et ouvre de nouvelles perspectives pour l'appréhension du rôle de la régulation transcriptionnelle dans les événements clefs du cycle de vie de *Plasmodium falciparum*. L'accumulation de ces connaissances sur la biologie de *P. falciparum* est un pas essentiel dans la lutte contre le paludisme.

BIBLIOGRAPHIE

1. Hume, J.C., E.J. Lyons, and K.P. Day, *Human migration, mosquitoes and the evolution of Plasmodium falciparum*. Trends Parasitol, 2003. **19**(3): p. 144-9.
2. Dobson, M.J., *The malariology centenary*. Parassitologia, 1999. **41**(1-3): p. 21-32.
3. Fantini, B., *The concept of specificity and the Italian contribution to the discovery of the malaria transmission cycle*. Parassitologia, 1999. **41**(1-3): p. 39-47.
4. Hay, S.I., et al., *The global distribution and population at risk of malaria: past, present, and future*. Lancet Infect Dis, 2004. **4**(6): p. 327-36.
5. WHO, *WHO Expert Committee on Malaria - Twentieth Report*. 2000, World Health Organization: Geneva.
6. Hay, S.I., et al., *Annual Plasmodium falciparum entomological inoculation rates (EIR) across Africa: literature survey, Internet access and review*. Trans R Soc Trop Med Hyg, 2000. **94**(2): p. 113-27.
7. WHO, *The world health report 2002: reducing risks, promoting healthy life*. 2003, World Health Organization: Geneva.
8. Luzzatto, L., *Genetics of red cells and susceptibility to malaria*. Blood, 1979. **54**(5): p. 961-76.
9. Sachs, J. and P. Malaney, *The economic and social burden of malaria*. Nature, 2002. **415**(6872): p. 680-5.
10. Breman, J.G., M.S. Alilio, and A. Mills, *Conquering the intolerable burden of malaria: what's new, what's needed: a summary*. Am J Trop Med Hyg, 2004. **71**(2 Suppl): p. 1-15.
11. Snow, R.W., et al., *Estimating mortality, morbidity and disability due to malaria among Africa's non-pregnant population*. Bull World Health Organ, 1999. **77**(8): p. 624-40.
12. Sylla, A., et al., *[Self-destructive behavior in African adolescents]*. Med Trop (Mars), 2003. **63**(4-5): p. 539-43.
13. Mackintosh, C.L., J.G. Beeson, and K. Marsh, *Clinical features and pathogenesis of severe malaria*. Trends Parasitol, 2004. **20**(12): p. 597-603.
14. Marsh, K., et al., *Indicators of life-threatening malaria in African children*. N Engl J Med, 1995. **332**(21): p. 1399-404.
15. Dugbartey, A.T., F.J. Spellacy, and M.T. Dugbartey, *Somatosensory discrimination deficits following pediatric cerebral malaria*. Am J Trop Med Hyg, 1998. **59**(3): p. 393-6.
16. http://www.dpd.cdc.gov/dpdx/HTML/Image_Library.htm, DPDx : parasite library.
17. Frevort, U., *Sneaking in through the back entrance: the biology of malaria liver stages*. Trends Parasitol, 2004. **20**(9): p. 417-24.
18. Mota, M.M., et al., *Migration of Plasmodium sporozoites through cells before infection*. Science, 2001. **291**(5501): p. 141-4.
19. Bannister, L.H., et al., *A brief illustrated guide to the ultrastructure of plasmodium falciparum asexual blood stages*. Parasitol Today, 2000. **16**(10): p. 427-33.
20. Chitnis, C.E. and M.J. Blackman, *Host cell invasion by malaria parasites*. Parasitol Today, 2000. **16**(10): p. 411-5.
21. Haldar, K., *Intracellular trafficking in Plasmodium-infected erythrocytes*. Curr Opin Microbiol, 1998. **1**(4): p. 466-71.
22. Cooke, B.M., et al., *Protein trafficking in Plasmodium falciparum-infected red blood cells*. Trends Parasitol, 2004. **20**(12): p. 581-9.
23. Wickham, M.E., J.G. Culvenor, and A.F. Cowman, *Selective inhibition of a two-step egress of malaria parasites from the host erythrocyte*. J Biol Chem, 2003. **278**(39): p. 37658-63.

24. Eichner, M., et al., *Genesis, sequestration and survival of Plasmodium falciparum gametocytes: parameter estimates from fitting a model to malariatherapy data*. Trans R Soc Trop Med Hyg, 2001. **95**(5): p. 497-501.
25. Bruce, M.C., et al., *Commitment of the malaria parasite Plasmodium falciparum to sexual and asexual development*. Parasitology, 1990. **100 Pt 2**: p. 191-200.
26. Silvestrini, F., P. Alano, and J.L. Williams, *Commitment to the production of male and female gametocytes in the human malaria parasite Plasmodium falciparum*. Parasitology, 2000. **121 Pt 5**: p. 465-71.
27. Dyer, M. and K.P. Day, *Commitment to gametocytogenesis in Plasmodium falciparum*. Parasitol Today, 2000. **16**(3): p. 102-7.
28. Dyer, M. and K. Day, *Expression of Plasmodium falciparum trimeric G proteins and their involvement in switching to sexual development*. Mol Biochem Parasitol, 2000. **110**(2): p. 437-48.
29. Inselburg, J., *Stage-specific inhibitory effect of cyclic AMP on asexual maturation and gametocyte formation of Plasmodium falciparum*. J Parasitol, 1983. **69**(3): p. 592-7.
30. Trager, W. and G.S. Gill, *Enhanced gametocyte formation in young erythrocytes by Plasmodium falciparum in vitro*. J Protozool, 1992. **39**(3): p. 429-32.
31. Dyer, M. and K.P. Day, *Regulation of the rate of asexual growth and commitment to sexual development by diffusible factors from in vitro cultures of Plasmodium falciparum*. Am J Trop Med Hyg, 2003. **68**(4): p. 403-9.
32. Schneweis, S., W.A. Maier, and H.M. Seitz, *Haemolysis of infected erythrocytes--a trigger for formation of Plasmodium falciparum gametocytes?* Parasitol Res, 1991. **77**(5): p. 458-60.
33. Lobo, C.A., R.N. Konings, and N. Kumar, *Expression of early gametocyte-stage antigens Pfg27 and Pfs16 in synchronized gametocytes and non-gametocyte producing clones of Plasmodium falciparum*. Mol Biochem Parasitol, 1994. **68**(1): p. 151-4.
34. Lobo, C.A., et al., *Disruption of the Pfg27 locus by homologous recombination leads to loss of the sexual phenotype in P. falciparum*. Mol Cell, 1999. **3**(6): p. 793-8.
35. Kongkasuriyachai, D., H. Fujioka, and N. Kumar, *Functional analysis of Plasmodium falciparum parasitophorous vacuole membrane protein (Pfs16) during gametocytogenesis and gametogenesis by targeted gene disruption*. Mol Biochem Parasitol, 2004. **133**(2): p. 275-85.
36. Alano, P., et al., *Plasmodium falciparum: parasites defective in early stages of gametocytogenesis*. Exp Parasitol, 1995. **81**(2): p. 227-35.
37. Day, K.P., et al., *Genes necessary for expression of a virulence determinant and for transmission of Plasmodium falciparum are located on a 0.3-megabase region of chromosome 9*. Proc Natl Acad Sci U S A, 1993. **90**(17): p. 8292-6.
38. Billker, O., et al., *Calcium and a calcium-dependent protein kinase regulate gamete formation and mosquito transmission in a malaria parasite*. Cell, 2004. **117**(4): p. 503-14.
39. Rodriguez, M.H. and L. Hernandez-Hernandez Fde, *Insect-malaria parasites interactions: the salivary gland*. Insect Biochem Mol Biol, 2004. **34**(7): p. 615-24.
40. Kaiser, K., et al., *Differential transcriptome profiling identifies Plasmodium genes encoding pre-erythrocytic stage-specific proteins*. Mol Microbiol, 2004. **51**(5): p. 1221-32.
41. Graeser, R., R.M. Franklin, and B. Kappes, *Mechanisms of activation of the cdc2-related kinase PfPK5 from Plasmodium falciparum*. Mol Biochem Parasitol, 1996. **79**(1): p. 125-7.
42. Doerig, C., et al., *The cell cycle in protozoan parasites*. Prog Cell Cycle Res, 2000. **4**: p. 163-83.

43. Leete, T.H. and H. Rubin, *Malaria and the cell cycle*. Parasitol Today, 1996. **12**(11): p. 442-4.
44. Arnot, D.E. and K. Gull, *The Plasmodium cell-cycle: facts and questions*. Ann Trop Med Parasitol, 1998. **92**(4): p. 361-5.
45. Schrevel, J., G. Asfaux-Foucher, and J.M. Bafort, [*Ultrastructural study of multiple mitoses during sporogony of Plasmodium b. berghei*]. J Ultrastruct Res, 1977. **59**(3): p. 332-50.
46. Read, M., et al., *Microtubular organization visualized by immunofluorescence microscopy during erythrocytic schizogony in Plasmodium falciparum and investigation of post-translational modifications of parasite tubulin*. Parasitology, 1993. **106 (Pt 3)**: p. 223-32.
47. Doerig, C., J. Endicott, and D. Chakrabarti, *Cyclin-dependent kinase homologues of Plasmodium falciparum*. Int J Parasitol, 2002. **32**(13): p. 1575-85.
48. Merckx, A., et al., *Identification and initial characterization of three novel cyclin-related proteins of the human malaria parasite Plasmodium falciparum*. J Biol Chem, 2003.
49. Graeser, R., et al., *Plasmodium falciparum protein kinase 5 and the malarial nuclear division cycles*. Mol Biochem Parasitol, 1996. **82**(1): p. 37-49.
50. Hoffman, S.L., et al., *Funding for malaria genome sequencing*. Nature, 1997. **387**(6634): p. 647.
51. Gardner, M.J., et al., *Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum*. Science, 1998. **282**(5391): p. 1126-32.
52. Bowman, S., et al., *The complete nucleotide sequence of chromosome 3 of Plasmodium falciparum [see comments]*. Nature, 1999. **400**(6744): p. 532-8.
53. Gardner, M.J., et al., *Genome sequence of the human malaria parasite Plasmodium falciparum*. Nature, 2002. **419**(6906): p. 498-511.
54. Abrahamsen, M.S., et al., *Complete genome sequence of the apicomplexan, Cryptosporidium parvum*. Science, 2004. **304**(5669): p. 441-5.
55. Waters, A.P., *The ribosomal RNA genes of Plasmodium*. Adv Parasitol, 1994. **34**: p. 33-79.
56. Velichutina, I.V., et al., *Chimeric rRNAs containing the GTPase centers of the developmentally regulated ribosomal rRNAs of Plasmodium falciparum are functionally distinct*. Rna, 1998. **4**(5): p. 594-602.
57. O'Donnell, R.A., et al., *A genetic screen for improved plasmid segregation reveals a role for Rep20 in the interaction of Plasmodium falciparum chromosomes*. Embo J, 2002. **21**(5): p. 1231-9.
58. Coulson, R.M., N. Hall, and C.A. Ouzounis, *Comparative genomics of transcriptional control in the human malaria parasite Plasmodium falciparum*. Genome Res, 2004. **14**(8): p. 1548-54.
59. Desai, S.A., S.M. Bezrukov, and J. Zimmerberg, *A voltage-dependent channel involved in nutrient uptake by red blood cells infected with the malaria parasite*. Nature, 2000. **406**(6799): p. 1001-5.
60. Mercereau-Puijalon, O., J.C. Barale, and E. Bischoff, *Three multigene families in Plasmodium parasites: facts and questions*. Int J Parasitol, 2002. **32**(11): p. 1323-44.
61. Aravind, L., et al., *Plasmodium biology: genomic gleanings*. Cell, 2003. **115**(7): p. 771-85.
62. Bastien, O., et al., *Analysis of the compositional biases in Plasmodium falciparum genome and proteome using Arabidopsis thaliana as a reference*. Gene, 2004. **336**(2): p. 163-73.

63. Rost, B., *Twilight zone of protein sequence alignments*. Protein Eng, 1999. **12**(2): p. 85-94.
64. Sharma, A., et al., *Structure of a gametocyte protein essential for sexual development in Plasmodium falciparum*. Nat Struct Biol, 2003. **10**(3): p. 197-203.
65. <http://www.sanger.ac.uk/Software/Pfam/>, Sanger Institute.
66. Bahl, A., et al., *PlasmoDB: the Plasmodium genome resource. A database integrating experimental and computational data*. Nucleic Acids Res, 2003. **31**(1): p. 212-5.
67. http://myhits.isb-sib.ch/cgi-bin/motif_scan, Expasy.
68. Corden, J.L., *Tails of RNA polymerase II*. Trends Biochem Sci, 1990. **15**(10): p. 383-7.
69. Carlson, M., *Genetics of transcriptional regulation in yeast: connections to the RNA polymerase II CTD*. Annu Rev Cell Dev Biol, 1997. **13**: p. 1-23.
70. Stiller, J.W. and B.D. Hall, *Evolution of the RNA polymerase II C-terminal domain*. Proc Natl Acad Sci U S A, 2002. **99**(9): p. 6091-6.
71. Evers, R., et al., *Trypanosoma brucei contains two RNA polymerase II largest subunit genes with an altered C-terminal domain*. Cell, 1989. **56**(4): p. 585-97.
72. Hirtzlin, J., P.M. Farber, and R.M. Franklin, *Isolation of a novel Plasmodium falciparum gene encoding a protein homologous to the Tat-binding protein family*. Eur J Biochem, 1994. **226**(2): p. 673-80.
73. McAndrew, M.B., et al., *Characterisation of the gene encoding an unusually divergent TATA-binding protein (TBP) from the extremely A+T-rich human malaria parasite Plasmodium falciparum*. Gene, 1993. **124**(2): p. 165-71.
74. Levine, M. and R. Tjian, *Transcription regulation and animal diversity*. Nature, 2003. **424**(6945): p. 147-51.
75. Coulson, R.M. and C.A. Ouzounis, *The phylogenetic diversity of eukaryotic transcription*. Nucleic Acids Res, 2003. **31**(2): p. 653-60.
76. Hansen, J.C., *Conformational dynamics of the chromatin fiber in solution: determinants, mechanisms, and functions*. Annu Rev Biophys Biomol Struct, 2002. **31**: p. 361-92.
77. Harvey, A.C. and J.A. Downs, *What functions do linker histones provide?* Mol Microbiol, 2004. **53**(3): p. 771-5.
78. Ausio, J., *Are linker histones (histone H1) dispensable for survival?* Bioessays, 2000. **22**(10): p. 873-7.
79. Henikoff, S., T. Furuyama, and K. Ahmad, *Histone variants, nucleosome assembly and epigenetic inheritance*. Trends Genet, 2004. **20**(7): p. 320-6.
80. He, H. and N. Lehming, *Global effects of histone modifications*. Brief Funct Genomic Proteomic, 2003. **2**(3): p. 234-43.
81. Strahl, B.D. and C.D. Allis, *The language of covalent histone modifications*. Nature, 2000. **403**(6765): p. 41-5.
82. Fan, Q., L. An, and L. Cui, *Plasmodium falciparum histone acetyltransferase, a yeast GCN5 homologue involved in chromatin remodeling*. Eukaryot Cell, 2004. **3**(2): p. 264-76.
83. Roth, S.Y., J.M. Denu, and C.D. Allis, *Histone acetyltransferases*. Annu Rev Biochem, 2001. **70**: p. 81-120.
84. Fan, Q., L. An, and L. Cui, *PfADA2, a Plasmodium falciparum homologue of the transcriptional coactivator ADA2 and its in vivo association with the histone acetyltransferase PfGCN5*. Gene, 2004. **336**(2): p. 251-61.
85. Blander, G. and L. Guarente, *The Sir2 family of protein deacetylases*. Annu Rev Biochem, 2004. **73**: p. 417-35.

86. Lutterbach, B., et al., *The MYND motif is required for repression of basal transcription from the multidrug resistance 1 promoter by the t(8;21) fusion protein.* Mol Cell Biol, 1998. **18**(6): p. 3604-11.
87. Becker, P.B. and W. Horz, *ATP-dependent nucleosome remodeling.* Annu Rev Biochem, 2002. **71**: p. 247-73.
88. Matys, V., et al., *TRANSFAC: transcriptional regulation, from patterns to profiles.* Nucleic Acids Res, 2003. **31**(1): p. 374-8.
89. Wood, V., et al., *The genome sequence of Schizosaccharomyces pombe.* Nature, 2002. **415**(6874): p. 871-80.
90. Hurst, H.C., *Transcription factors 1: bZIP proteins.* Protein Profile, 1995. **2**(2): p. 101-68.
91. Ellenberger, T.E., et al., *The GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted alpha helices: crystal structure of the protein-DNA complex.* Cell, 1992. **71**(7): p. 1223-37.
92. Klug, A., *Zinc finger peptides for the regulation of gene expression.* J Mol Biol, 1999. **293**(2): p. 215-8.
93. Klug, A. and J.W. Schwabe, *Protein motifs 5. Zinc fingers.* Faseb J, 1995. **9**(8): p. 597-604.
94. Pedone, P.V., et al., *The single Cys2-His2 zinc finger domain of the GAGA protein flanked by basic residues is sufficient for high-affinity specific DNA binding.* Proc Natl Acad Sci U S A, 1996. **93**(7): p. 2822-6.
95. Iuchi, S., *Three classes of C2H2 zinc finger proteins.* Cell Mol Life Sci, 2001. **58**(4): p. 625-35.
96. Torok, M. and L.D. Etkin, *Two B or not two B? Overview of the rapidly expanding B-box family of proteins.* Differentiation, 2001. **67**(3): p. 63-71.
97. Robson, F., et al., *Functional importance of conserved domains in the flowering-time gene CONSTANS demonstrated by analysis of mutant alleles and transgenic plants.* Plant J, 2001. **28**(6): p. 619-31.
98. Samach, A., et al., *Distinct roles of CONSTANS target genes in reproductive development of Arabidopsis.* Science, 2000. **288**(5471): p. 1613-6.
99. Putilina, T., P. Wong, and S. Gentleman, *The DHHC domain: a new highly conserved cysteine-rich motif.* Mol Cell Biochem, 1999. **195**(1-2): p. 219-26.
100. Harrison, S.C. and A.K. Aggarwal, *DNA recognition by proteins with the helix-turn-helix motif.* Annu Rev Biochem, 1990. **59**: p. 933-69.
101. Introna, M., et al., *The myb oncogene family of transcription factors: potent regulators of hematopoietic cell proliferation and differentiation.* Semin Cancer Biol, 1994. **5**(2): p. 113-24.
102. Howe, K.M., C.F. Reakes, and R.J. Watson, *Characterization of the sequence-specific interaction of mouse c-myb protein with DNA.* Embo J, 1990. **9**(1): p. 161-9.
103. Tanikawa, J., et al., *Recognition of specific DNA sequences by the c-myb protooncogene product: role of three repeat units in the DNA-binding domain.* Proc Natl Acad Sci U S A, 1993. **90**(20): p. 9320-4.
104. Frampton, J., et al., *Proposed structure for the DNA-binding domain of the Myb oncoprotein based on model building and mutational analysis.* Protein Eng, 1991. **4**(8): p. 891-901.
105. Lipsick, J.S., *One billion years of Myb.* Oncogene, 1996. **13**(2): p. 223-35.
106. Morrow, B.E., Q. Ju, and J.R. Warner, *A bipartite DNA-binding domain in yeast Reb1p.* Mol Cell Biol, 1993. **13**(2): p. 1173-82.

107. Hovring, I., et al., *DNA-binding domain and recognition sequence of the yeast BAS1 protein, a divergent member of the Myb family of transcription factors*. J Biol Chem, 1994. **269**(26): p. 17663-9.
108. Pinson, B., et al., *Highly conserved features of DNA binding between two divergent members of the myb family of transcription factors*. Nucleic Acids Res, 2001. **29**(2): p. 527-35.
109. Stober-Grasser, U., et al., *The Myb DNA-binding domain is highly conserved in Dictyostelium discoideum*. Oncogene, 1992. **7**(3): p. 589-96.
110. Ganter, B., S.T. Chao, and J.S. Lipsick, *Transcriptional activation by the myb proteins requires a specific local promoter structure*. FEBS Lett, 1999. **460**(3): p. 401-10.
111. Ness, S.A., et al., *Myb and NF-M: combinatorial activators of myeloid genes in heterologous cell types*. Genes Dev, 1993. **7**(5): p. 749-59.
112. Mo, Y., et al., *Crystal structure of a ternary SAP-1/SRF/c-fos SRE DNA complex*. J Mol Biol, 2001. **314**(3): p. 495-506.
113. Messenguy, F. and E. Dubois, *Role of MADS box proteins and their cofactors in combinatorial control of gene expression and cell development*. Gene, 2003. **316**: p. 1-21.
114. Shore, P. and A.D. Sharrocks, *The MADS-box family of transcription factors*. Eur J Biochem, 1995. **229**(1): p. 1-13.
115. Maity, S.N. and B. de Crombrughe, *Role of the CCAAT-binding protein CBF/NF-Y in transcription*. Trends Biochem Sci, 1998. **23**(5): p. 174-8.
116. Mantovani, R., *The molecular biology of the CCAAT-binding factor NF-Y*. Gene, 1999. **239**(1): p. 15-27.
117. Wistow, G., *Cold shock and DNA binding*. Nature, 1990. **344**(6269): p. 823-4.
118. Matsumoto, K. and A.P. Wolffe, *Gene regulation by Y-box proteins: coupling control of transcription and translation*. Trends Cell Biol, 1998. **8**(8): p. 318-23.
119. Zou, Y. and K.R. Chien, *EF1A/YB-1 is a component of cardiac HF-1A binding activity and positively regulates transcription of the myosin light-chain 2v gene*. Mol Cell Biol, 1995. **15**(6): p. 2972-82.
120. Karlson, D. and R. Imai, *Conservation of the cold shock domain protein family in plants*. Plant Physiol, 2003. **131**(1): p. 12-5.
121. Grasser, K.D., *Chromatin-associated HMGA and HMGB proteins: versatile co-regulators of DNA-dependent processes*. Plant Mol Biol, 2003. **53**(3): p. 281-95.
122. Webster, C.I., L.C. Packman, and J.C. Gray, *HMG-1 enhances HMG-I/Y binding to an A/T-rich enhancer element from the pea plastocyanin gene*. Eur J Biochem, 2001. **268**(11): p. 3154-62.
123. Thomas, J.O. and A.A. Travers, *HMGI and 2, and related 'architectural' DNA-binding proteins*. Trends Biochem Sci, 2001. **26**(3): p. 167-74.
124. Travers, A.A., *Priming the nucleosome: a role for HMGB proteins?* EMBO Rep, 2003. **4**(2): p. 131-6.
125. Lee, Y., et al., *Jumonji, a nuclear protein that is necessary for normal heart development*. Circ Res, 2000. **86**(9): p. 932-8.
126. Balciunas, D. and H. Ronne, *Evidence of domain swapping within the jumonji family of transcription factors*. Trends Biochem Sci, 2000. **25**(6): p. 274-6.
127. Kim, T.G., et al., *JUMONJI, a critical factor for cardiac development, functions as a transcriptional repressor*. J Biol Chem, 2003. **278**(43): p. 42247-55.
128. Woychik, N.A. and M. Hampsey, *The RNA polymerase II machinery: structure illuminates function*. Cell, 2002. **108**(4): p. 453-63.
129. Lewis, B.A. and D. Reinberg, *The mediator coactivator complex: functional and physical roles in transcriptional regulation*. J Cell Sci, 2003. **116**(Pt 18): p. 3667-75.

130. Baek, H.J., et al., *Requirement of TRAP/mediator for both activator-independent and activator-dependent transcription in conjunction with TFIID-associated TAF(II)s*. Mol Cell Biol, 2002. **22**(8): p. 2842-52.
131. Yudkovsky, N., J.A. Ranish, and S. Hahn, *A transcription reinitiation intermediate that is stabilized by activator*. Nature, 2000. **408**(6809): p. 225-9.
132. Bourbon, H.M., et al., *A unified nomenclature for protein subunits of mediator complexes linking transcriptional regulators to RNA polymerase II*. Mol Cell, 2004. **14**(5): p. 553-7.
133. Linder, T. and C.M. Gustafsson, *The Suhl1/MED31 protein is an ancient component of Schizosaccharomyces pombe and Saccharomyces cerevisiae mediator*. J Biol Chem, 2004.
134. Li, F.Q., H. Ueda, and S. Hirose, *Mediators of activation of fushi tarazu gene transcription by BmFTZ-F1*. Mol Cell Biol, 1994. **14**(5): p. 3013-21.
135. Takemaru, K., et al., *Yeast coactivator MBF1 mediates GCN4-dependent transcriptional activation*. Mol Cell Biol, 1998. **18**(9): p. 4971-6.
136. Tsuda, K., et al., *Three Arabidopsis MBF1 homologs with distinct expression profiles play roles as transcriptional co-activators*. Plant Cell Physiol, 2004. **45**(2): p. 225-31.
137. Millership, J.J., et al., *Differential expression and interaction of transcription co-activator MBF1 with TATA-binding protein (TBP) in the apicomplexan Cryptosporidium parvum*. Microbiology, 2004. **150**(Pt 5): p. 1207-13.
138. Le Roch, K.G., et al., *Discovery of gene function by expression profiling of the malaria parasite life cycle*. Science, 2003. **301**(5639): p. 1503-8.
139. Collart, M.A. and K. Struhl, *NOT1(CDC39), NOT2(CDC36), NOT3, and NOT4 encode a global-negative regulator of transcription that differentially affects TATA-element utilization*. Genes Dev, 1994. **8**(5): p. 525-37.
140. Collart, M.A., *Global control of gene expression in yeast by the Ccr4-Not complex*. Gene, 2003. **313**: p. 1-16.
141. Guhaniyogi, J. and G. Brewer, *Regulation of mRNA stability in mammalian cells*. Gene, 2001. **265**(1-2): p. 11-23.
142. Levitt, A., *RNA processing in malarial parasites*. Parasitol Today, 1993. **9**(12): p. 465-8.
143. Korner, C.G. and E. Wahle, *Poly(A) tail shortening by a mammalian poly(A)-specific 3'-exoribonuclease*. J Biol Chem, 1997. **272**(16): p. 10448-56.
144. Cui, L., Q. Fan, and J. Li, *The malaria parasite Plasmodium falciparum encodes members of the Puf RNA-binding protein family with conserved RNA binding activity*. Nucleic Acids Res, 2002. **30**(21): p. 4607-17.
145. Spassov, D.S. and R. Jurecic, *The PUF family of RNA-binding proteins: does evolutionarily conserved structure equal conserved function?* IUBMB Life, 2003. **55**(7): p. 359-66.
146. Lai, W.S., et al., *Evidence that tristetraproline binds to AU-rich elements and promotes the deadenylation and destabilization of tumor necrosis factor alpha mRNA*. Mol Cell Biol, 1999. **19**(6): p. 4311-23.
147. Ross, J., *mRNA stability in mammalian cells*. Microbiol Rev, 1995. **59**(3): p. 423-50.
148. Sachs, A.B. and G. Varani, *Eukaryotic translation initiation: there are (at least) two sides to every story*. Nat Struct Biol, 2000. **7**(5): p. 356-61.
149. Mazumder, B., V. Seshadri, and P.L. Fox, *Translational control by the 3'-UTR: the ends specify the means*. Trends Biochem Sci, 2003. **28**(2): p. 91-8.
150. Loyevsky, M., et al., *Expression of a recombinant IRP-like Plasmodium falciparum protein that specifically binds putative plasmodial IREs*

- An IRP-like protein from Plasmodium falciparum binds to a mammalian iron-responsive element.* Mol Biochem Parasitol, 2003. **126**(2): p. 231-8.
151. Mohrle, J.J., et al., *Molecular cloning, characterization and localization of PfPK4, an eIF-2alpha kinase-related enzyme from the malarial parasite Plasmodium falciparum.* Biochem J, 1997. **328 (Pt 2)**: p. 677-87.
152. Chen, J.J. and I.M. London, *Regulation of protein synthesis by heme-regulated eIF-2 alpha kinase.* Trends Biochem Sci, 1995. **20**(3): p. 105-8.
153. Ferrell, K., et al., *Regulatory subunit interactions of the 26S proteasome, a complex problem.* Trends Biochem Sci, 2000. **25**(2): p. 83-8.
154. Desterro, J.M., M.S. Rodriguez, and R.T. Hay, *Regulation of transcription factors by protein degradation.* Cell Mol Life Sci, 2000. **57**(8-9): p. 1207-19.
155. Li, G.D., et al., *Molecular cloning of a gene encoding a 20S proteasome beta subunit from Plasmodium falciparum.* Int J Parasitol, 2000. **30**(6): p. 729-33.
156. Horrocks, P. and C.I. Newbold, *Intraerythrocytic polyubiquitin expression in Plasmodium falciparum is subjected to developmental and heat-shock control.* Mol Biochem Parasitol, 2000. **105**(1): p. 115-25.
157. Gantt, S.M., et al., *Proteasome inhibitors block development of Plasmodium spp.* Antimicrob Agents Chemother, 1998. **42**(10): p. 2731-8.
158. Pickart, C.M., *Ubiquitin in chains.* Trends Biochem Sci, 2000. **25**(11): p. 544-8.
159. Bach, I. and H.P. Ostendorff, *Orchestrating nuclear functions: ubiquitin sets the rhythm.* Trends Biochem Sci, 2003. **28**(4): p. 189-95.
160. Muller, S., A. Ledl, and D. Schmidt, *SUMO: a regulator of gene expression and genome integrity.* Oncogene, 2004. **23**(11): p. 1998-2008.
161. Ward, P., et al., *Protein kinases of the human malaria parasite Plasmodium falciparum: the kinome of a divergent eukaryote.* BMC Genomics, 2004. **5**(1): p. 79.
162. Whitmarsh, A.J. and R.J. Davis, *Regulation of transcription factor function by phosphorylation.* Cell Mol Life Sci, 2000. **57**(8-9): p. 1172-83.
163. Palancade, B. and O. Bensaude, *Investigating RNA polymerase II carboxyl-terminal domain (CTD) phosphorylation.* Eur J Biochem, 2003. **270**(19): p. 3859-70.
164. Bannister, A.J. and E.A. Miska, *Regulation of gene expression by transcription factor acetylation.* Cell Mol Life Sci, 2000. **57**(8-9): p. 1184-92.
165. Horrocks, P., K. Dechering, and M. Lanzer, *Control of gene expression in Plasmodium falciparum.* Mol Biochem Parasitol, 1998. **95**(2): p. 171-81.
166. Hayward, R.E., et al., *Shotgun DNA microarrays and stage-specific gene expression in Plasmodium falciparum malaria.* Mol Microbiol, 2000. **35**(1): p. 6-14.
167. Patankar, S., et al., *Serial Analysis of Gene Expression in Plasmodium falciparum Reveals the Global Expression Profile of Erythrocytic Stages and the Presence of Anti-Sense Transcripts in the Malarial Parasite.* Mol Biol Cell, 2001. **12**(10): p. 3114-25.
168. Mamoun, C.B., et al., *Co-ordinated programme of gene expression during asexual intraerythrocytic development of the human malaria parasite Plasmodium falciparum revealed by microarray analysis.* Mol Microbiol, 2001. **39**(1): p. 26-36.
169. Le Roch, K.G., et al., *Monitoring the chromosome 2 intraerythrocytic transcriptome of Plasmodium falciparum using oligonucleotide arrays.* Am J Trop Med Hyg, 2002. **67**(3): p. 233-43.
170. Bozdech, Z., et al., *Expression profiling of the schizont and trophozoite stages of Plasmodium falciparum with a long-oligonucleotide microarray.* Genome Biol, 2003. **4**(2): p. R9.
171. Bozdech, Z., et al., *The Transcriptome of the Intraerythrocytic Developmental Cycle of Plasmodium falciparum.* PLoS Biol, 2003. **1**(1): p. 5.

172. Triglia, T., et al., *Apical membrane antigen 1 plays a central role in erythrocyte invasion by Plasmodium species*. Mol Microbiol, 2000. **38**(4): p. 706-18.
173. Florens, L., et al., *A proteomic view of the Plasmodium falciparum life cycle*. Nature, 2002. **419**(6906): p. 520-6.
174. Le Roch, K.G., et al., *Global analysis of transcript and protein levels across the Plasmodium falciparum life cycle*. Genome Res, 2004. **14**(11): p. 2308-18.
175. Gunasekera, A.M., et al., *Drug-induced alterations in gene expression of the asexual blood forms of Plasmodium falciparum*. Mol Microbiol, 2003. **50**(4): p. 1229-39.
176. Fang, J., et al., *Ambient glucose concentration and gene expression in Plasmodium falciparum*. Mol Biochem Parasitol, 2004. **133**(1): p. 125-9.
177. Fang, J., M. Sullivan, and T.F. McCutchan, *The effects of glucose concentration on the reciprocal regulation of rRNA promoters in Plasmodium falciparum*. J Biol Chem, 2004. **279**(1): p. 720-5.
178. Lanzer, M., D. de Bruin, and J.V. Ravetch, *Transcription mapping of a 100 kb locus of Plasmodium falciparum identifies an intergenic region in which transcription terminates and reinitiates*. Embo J, 1992. **11**(5): p. 1949-55.
179. Su, X.Z. and T.E. Wellems, *Sequence, transcript characterization and polymorphisms of a Plasmodium falciparum gene belonging to the heat-shock protein (HSP) 90 family*. Gene, 1994. **151**(1-2): p. 225-30.
180. Lanzer, M., D. de Bruin, and J.V. Ravetch, *A sequence element associated with the Plasmodium falciparum KAHRP gene is the site of developmentally regulated protein-DNA interactions*. Nucleic Acids Res, 1992. **20**(12): p. 3051-6.
181. Wu, Y., et al., *Transfection of Plasmodium falciparum within human red blood cells*. Proc Natl Acad Sci U S A, 1995. **92**(4): p. 973-7.
182. Crabb, B.S. and A.F. Cowman, *Characterization of promoters and stable transfection by homologous and nonhomologous recombination in Plasmodium falciparum*. Proc Natl Acad Sci U S A, 1996. **93**(14): p. 7289-94.
183. Osta, M., et al., *A 24 bp cis-acting element essential for the transcriptional activity of Plasmodium falciparum CDP-diacylglycerol synthase gene promoter*. Mol Biochem Parasitol, 2002. **121**(1): p. 87-98.
184. Horrocks, P. and M. Lanzer, *Mutational analysis identifies a five base pair cis-acting sequence essential for GBP130 promoter activity in Plasmodium falciparum*. Mol Biochem Parasitol, 1999. **99**(1): p. 77-87.
185. Militello, K.T., et al., *Identification of regulatory elements in the Plasmodium falciparum genome*. Mol Biochem Parasitol, 2004. **134**(1): p. 75-88.
186. Dechering, K.J., et al., *Developmentally regulated expression of pfs16, a marker for sexual differentiation of the human malaria parasite Plasmodium falciparum*. Mol Biochem Parasitol, 1997. **89**(2): p. 235-44.
187. Robert, V., et al., *Kinetics and efficiency of Plasmodium falciparum development in the midguts of Anopheles gambiae, An. funestus and An. nili*. Ann Trop Med Parasitol, 1998. **92**(1): p. 115-8.
188. Dechering, K.J., et al., *Isolation and functional characterization of two distinct sexual-stage-specific promoters of the human malaria parasite Plasmodium falciparum*. Mol Cell Biol, 1999. **19**(2): p. 967-78.
189. Fernandez-Becerra, C., et al., *Plasmodium falciparum: new vector with bi-directional promoter activity to stably express transgenes*. Exp Parasitol, 2003. **103**(1-2): p. 88-91.
190. Deitsch, K.W., M.S. Calderwood, and T.E. Wellems, *Malaria. Cooperative silencing elements in var genes*. Nature, 2001. **412**(6850): p. 875-6.

191. Calderwood, M.S., et al., *Plasmodium falciparum* var genes are regulated by two regions with separate promoters, one upstream of the coding region and a second within the intron. *J Biol Chem*, 2003. **278**(36): p. 34125-32.
192. Horrocks, P. and M. Lanzer, *Differences in nucleosome organization over episomally located plasmids coincides with aberrant promoter activity in P. falciparum*. *Parasitology International*, 1999. **1999**(48): p. 55-61.
193. Horrocks, P., et al., *Stage-specific promoter activity from stably maintained episomes in Plasmodium falciparum*. *Int J Parasitol*, 2002. **32**(10): p. 1203-6.
194. Nakaar, V., et al., *Upstream elements required for expression of nucleoside triphosphate hydrolase genes of Toxoplasma gondii*. *Mol Biochem Parasitol*, 1998. **92**(2): p. 229-39.
195. Soldati, D. and J.C. Boothroyd, *A selector of transcription initiation in the protozoan parasite Toxoplasma gondii*. *Mol Cell Biol*, 1995. **15**(1): p. 87-93.
196. Bohne, W., A. Wirsing, and U. Gross, *Bradyzoite-specific gene expression in Toxoplasma gondii requires minimal genomic elements*. *Mol Biochem Parasitol*, 1997. **85**(1): p. 89-98.
197. Kutach, A.K. and J.T. Kadonaga, *The downstream promoter element DPE appears to be as widely used as the TATA box in Drosophila core promoters*. *Mol Cell Biol*, 2000. **20**(13): p. 4754-64.
198. Chow, C.S. and D.F. Wirth, *Linker scanning mutagenesis of the Plasmodium gallinaceum sexual stage specific gene pgs28 reveals a novel downstream cis-control element*. *Mol Biochem Parasitol*, 2003. **129**(2): p. 199-208.
199. Pace, T., et al., *Developmental regulation of a Plasmodium gene involves the generation of stage-specific 5' untranslated sequences*. *Mol Biochem Parasitol*, 1998. **97**(1-2): p. 45-53.
200. Dechering, K.J., et al., *Distinct frequency-distributions of homopolymeric DNA tracts in different genomes*. *Nucleic Acids Res*, 1998. **26**(17): p. 4056-62.
201. Perez-Martin, J. and V. de Lorenzo, *Clues and consequences of DNA bending in transcription*. *Annu Rev Microbiol*, 1997. **51**: p. 593-628.
202. Porter, M.E., *The DNA polymerase delta promoter from Plasmodium falciparum contains an unusually long 5' untranslated region and intrinsic DNA curvature*. *Mol Biochem Parasitol*, 2001. **114**(2): p. 249-55.
203. Porter, M.E., *Positive and negative effects of deletions and mutations within the 5' flanking sequences of Plasmodium falciparum DNA polymerase delta*. *Mol Biochem Parasitol*, 2002. **122**(1): p. 9-19.
204. Struhl, K., *Naturally occurring poly(dA-dT) sequences are upstream promoter elements for constitutive transcription in yeast*. *Proc Natl Acad Sci U S A*, 1985. **82**(24): p. 8419-23.
205. Kumar, N., et al., *Molecular complexity of sexual development and gene regulation in Plasmodium falciparum*. *Int J Parasitol*, 2004. **34**(13-14): p. 1451-8.
206. Roberts, G.C. and C.W. Smith, *Alternative splicing: combinatorial output from the genome*. *Curr Opin Chem Biol*, 2002. **6**(3): p. 375-83.
207. Vinkenoog, R., et al., *Comparison of introns in a cdc2-homologous gene within a number of Plasmodium species*. *Mol Biochem Parasitol*, 1995. **71**(2): p. 233-41.
208. Knapp, B., et al., *Demonstration of alternative splicing of a pre-mRNA expressed in the blood stage form of Plasmodium falciparum*. *J Biol Chem*, 1991. **266**(11): p. 7148-54.
209. Muhia, D.K., et al., *Multiple splice variants encode a novel adenylyl cyclase of possible plastid origin expressed in the sexual stage of the malaria parasite Plasmodium falciparum*. *J Biol Chem*, 2003. **278**(24): p. 22014-22.

210. Singh, N., et al., *Conservation and developmental control of alternative splicing in maeb1 among malaria parasites*. J Mol Biol, 2004. **343**(3): p. 589-99.
211. Paton, M.G., et al., *Structure and expression of a post-transcriptionally regulated malaria gene encoding a surface protein from the sexual stages of Plasmodium berghei*. Mol Biochem Parasitol, 1993. **59**(2): p. 263-75.
212. Curtis, D., R. Lehmann, and P.D. Zamore, *Translational regulation in development*. Cell, 1995. **81**(2): p. 171-8.
213. Margos, G., et al., *Transgenic expression of a mosquito-stage malarial protein, Pbs21, in blood stages of transformed Plasmodium berghei and induction of an immune response upon infection*. Infect Immun, 1998. **66**(8): p. 3884-91.
214. Golightly, L.M., et al., *3' UTR elements enhance expression of Pgs28, an ookinete protein of Plasmodium gallinaceum*. Mol Biochem Parasitol, 2000. **105**(1): p. 61-70.
215. Cann, H., et al., *3' UTR signals necessary for expression of the Plasmodium gallinaceum ookinete protein, Pgs28, share similarities with those of yeast and plants*. Mol Biochem Parasitol, 2004. **137**(2): p. 239-45.
216. Shue, P., et al., *The 3' UTR elements of P. gallinaceum protein Pgs28 are functionally distinct from those of human cells*. Mol Biochem Parasitol, 2004. **137**(2): p. 355-9.
217. Taylor, H.M., et al., *Plasmodium falciparum homologue of the genes for Plasmodium vivax and Plasmodium yoelii adhesive proteins, which is transcribed but not translated*. Infect Immun, 2001. **69**(6): p. 3635-45.
218. Triglia, T., J.K. Thompson, and A.F. Cowman, *An EBA175 homologue which is transcribed but not translated in erythrocytic stages of Plasmodium falciparum*. Mol Biochem Parasitol, 2001. **116**(1): p. 55-63.
219. Bischoff, E., et al., *A member of the Plasmodium falciparum Pf60 multigene family codes for a nuclear protein expressed by readthrough of an internal stop codon*. Mol Microbiol, 2000. **35**(5): p. 1005-16.
220. Zhang, K. and P.K. Rathod, *Divergent regulation of dihydrofolate reductase between malaria parasite and human host*. Science, 2002. **296**(5567): p. 545-7.
221. Nirmalan, N., P.F. Sims, and J.E. Hyde, *Translational up-regulation of antifolate drug targets in the human malaria parasite Plasmodium falciparum upon challenge with inhibitors*. Mol Biochem Parasitol, 2004. **136**(1): p. 63-70.
222. Yelin, R., et al., *Widespread occurrence of antisense transcription in the human genome*. Nat Biotechnol, 2003. **21**(4): p. 379-86.
223. Liniger, M., et al., *Overlapping sense and antisense transcription units in Trypanosoma brucei*. Mol Microbiol, 2001. **40**(4): p. 869-78.
224. Knodler, L.A., et al., *Developmental gene regulation in Giardia lamblia: first evidence for an encystation-specific promoter and differential 5' mRNA processing*. Mol Microbiol, 1999. **34**(2): p. 327-40.
225. Hildebrandt, M. and W. Nellen, *Differential antisense transcription from the Dictyostelium EB4 gene locus: implications on antisense-mediated regulation of mRNA stability*. Cell, 1992. **69**(1): p. 197-204.
226. Vanhee-Brossollet, C. and C. Vaquero, *Do natural antisense transcripts make sense in eukaryotes?* Gene, 1998. **211**(1): p. 1-9.
227. Vanhee-Brossollet, C., et al., *A natural antisense RNA derived from the HIV-1 env gene encodes a protein which is recognized by circulating antibodies of HIV+ individuals*. Virology, 1995. **206**(1): p. 196-202.
228. Nellen, W. and C. Lichtenstein, *What makes an mRNA anti-sense-itive?* Trends Biochem Sci, 1993. **18**(11): p. 419-23.
229. Matzke, M., A.J. Matzke, and J.M. Kooter, *RNA: guiding gene silencing*. Science, 2001. **293**(5532): p. 1080-3.

230. Kyes, S., et al., *Stage-specific merozoite surface protein 2 antisense transcripts in Plasmodium falciparum*. Mol Biochem Parasitol, 2002. **123**(1): p. 79-83.
231. Wickham, M.E., J.K. Thompson, and A.F. Cowman, *Characterisation of the merozoite surface protein-2 promoter using stable and transient transfection in Plasmodium falciparum*. Mol Biochem Parasitol, 2003. **129**(2): p. 147-56.
232. Gunasekera, A.M., et al., *Widespread distribution of antisense transcripts in the Plasmodium falciparum genome*. Mol Biochem Parasitol, 2004. **136**(1): p. 35-42.
233. Wanidworanun, C., R.L. Nagel, and H.L. Shear, *Antisense oligonucleotides targeting malarial aldolase inhibit the asexual erythrocytic stages of Plasmodium falciparum*. Mol Biochem Parasitol, 1999. **102**(1): p. 91-101.
234. Noonpakdee, W., et al., *Inhibition of Plasmodium falciparum proliferation in vitro by antisense oligodeoxynucleotides against malarial topoisomerase II*. Biochem Biophys Res Commun, 2003. **302**(4): p. 659-64.
235. Malhotra, P., et al., *Double-stranded RNA-mediated gene silencing of cysteine proteases (falcipain-1 and -2) of Plasmodium falciparum*. Mol Microbiol, 2002. **45**(5): p. 1245-54.
236. McRobert, L. and G.A. McConkey, *RNA interference (RNAi) inhibits growth of Plasmodium falciparum*. Mol Biochem Parasitol, 2002. **119**(2): p. 273-8.
237. Srinivasan, N. and A. Krupa, *A genomic perspective of protein kinases in Plasmodium falciparum*. Proteins, 2004.
238. Fidock, D.A., et al., *Plasmodium falciparum liver stage antigen-1 is well conserved and contains potent B and T cell determinants*. J Immunol, 1994. **153**(1): p. 190-204.
239. Wesseling, J.G., et al., *Stage-specific expression and genomic organization of the actin genes of the malaria parasite Plasmodium falciparum*. Mol Biochem Parasitol, 1989. **35**(2): p. 167-76.
240. Lobo, C.A. and N. Kumar, *Differential transcription of histone genes in asexual and sexual stages of Plasmodium falciparum*. Int J Parasitol, 1999. **29**(9): p. 1447-9.
241. Salin, H., et al., *A novel sensitive microarray approach for differential screening using probes labelled with two different radioelements*. Nucleic Acids Res, 2002. **30**(4): p. e17.
242. Bustin, S.A., *Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays*. J Mol Endocrinol, 2000. **25**(2): p. 169-93.
243. Livak, K.J. and T.D. Schmittgen, *Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method*. Methods, 2001. **25**(4): p. 402-8.
244. Hecht, A., S. Strahl-Bolsinger, and M. Grunstein, *Mapping DNA interaction sites of chromosomal proteins. Crosslinking studies in yeast*. Methods Mol Biol, 1999. **119**: p. 469-79.
245. Pologe, L.G., *Aberrant transcription and the failure of Plasmodium falciparum to differentiate into gametocytes*. Mol Biochem Parasitol, 1994. **68**(1): p. 35-43.
246. Alano, P., F. Silvestrini, and L. Roca, *Structure and polymorphism of the upstream region of the pfg27/25 gene, transcriptionally regulated in gametocytogenesis of Plasmodium falciparum*. Mol Biochem Parasitol, 1996. **79**(2): p. 207-17.
247. Shaklai, N., et al., *Accumulation and drainage of hemin in the red cell membrane*. Biochim Biophys Acta, 1985. **821**(2): p. 355-66.
248. Zhan, K., et al., *Phosphorylation of eukaryotic initiation factor 2 by heme-regulated inhibitor kinase-related protein kinases in Schizosaccharomyces pombe is important for resistance to environmental stresses*. Mol Cell Biol, 2002. **22**(20): p. 7134-46.
249. Clemens, M.J., *Initiation factor eIF2 alpha phosphorylation in stress responses and apoptosis*. Prog Mol Subcell Biol, 2001. **27**: p. 57-89.

250. Hinnebusch, A.G., *Translational regulation of yeast GCN4. A window on factors that control initiator-trna binding to the ribosome.* J Biol Chem, 1997. **272**(35): p. 21661-4.
251. Wek, R.C., *eIF-2 kinases: regulators of general and gene-specific translation initiation.* Trends Biochem Sci, 1994. **19**(11): p. 491-6.
252. Sarkar, A., et al., *Lead exposure and heat shock inhibit cell proliferation in human HeLa and K562 cells by inducing expression and activity of the heme-regulated eIF-2alpha kinase.* J Biochem Mol Biol Biophys, 2002. **6**(6): p. 391-6.
253. Akey, C.W. and K. Luger, *Histone chaperones and nucleosome assembly.* Curr Opin Struct Biol, 2003. **13**(1): p. 6-14.
254. Mello, J.A. and G. Almouzni, *The ins and outs of nucleosome assembly.* Curr Opin Genet Dev, 2001. **11**(2): p. 136-41.
255. Ohkuni, K., K. Shirahige, and A. Kikuchi, *Genome-wide expression analysis of NAP1 in Saccharomyces cerevisiae.* Biochem Biophys Res Commun, 2003. **306**(1): p. 5-9.
256. Kellogg, D.R., et al., *Members of the NAP/SET family of proteins interact specifically with B-type cyclins.* J Cell Biol, 1995. **130**(3): p. 661-73.
257. Birago, C., et al., *A chromatin-associated protein is encoded in a genomic region highly conserved in the Plasmodium genus.* Mol Biochem Parasitol, 1996. **80**(2): p. 193-202.
258. Koornneef, M., C.J. Hanhart, and J.H. van der Veen, *A genetic and physiological analysis of late flowering mutants in Arabidopsis thaliana.* Mol Gen Genet, 1991. **229**(1): p. 57-66.
259. Lieb, J.D., et al., *Promoter-specific binding of Rap1 revealed by genome-wide maps of protein-DNA association.* Nat Genet, 2001. **28**(4): p. 327-34.
260. Tahirov, T.H., et al., *Mechanism of c-Myb-C/EBP beta cooperation from separated sites on a promoter.* Cell, 2002. **108**(1): p. 57-70.
261. Li, H., et al., *Isolation and functional characterization of a dynamin-like gene from Plasmodium falciparum.* Biochem Biophys Res Commun, 2004. **320**(3): p. 664-71.
262. Mohammed, A., et al., *In vivo gene silencing in Plasmodium berghei--a mouse malaria model.* Biochem Biophys Res Commun, 2003. **309**(3): p. 506-11.
263. Matzke, M., et al., *Genetic analysis of RNA-mediated transcriptional gene silencing.* Biochim Biophys Acta, 2004. **1677**(1-3): p. 129-41.
264. Schramke, V. and R. Allshire, *Those interfering little RNAs! Silencing and eliminating chromatin.* Curr Opin Genet Dev, 2004. **14**(2): p. 174-80.
265. Ikemoto, N., et al., *Antisense oligodeoxynucleotides: useful tool for search and assessment of new targets for anti-malarial drugs.* Nucleic Acids Symp Ser, 1999(42): p. 89-90.
266. Trenholme, K.R., et al., *clag9: A cytoadherence gene in Plasmodium falciparum essential for binding of parasitized erythrocytes to CD36.* Proc Natl Acad Sci U S A, 2000. **97**(8): p. 4029-33.
267. Gardiner, D.L., et al., *Inhibition of Plasmodium falciparum clag9 gene function by antisense RNA.* Mol Biochem Parasitol, 2000. **110**(1): p. 33-41.
268. Kelman, Z., *PCNA: structure, functions and interactions.* Oncogene, 1997. **14**(6): p. 629-40.
269. Ohi, R., et al., *Myb-related Schizosaccharomyces pombe cdc5p is structurally and functionally conserved in eukaryotes.* Mol Cell Biol, 1998. **18**(7): p. 4097-108.
270. Ju, Q.D., B.E. Morrow, and J.R. Warner, *REB1, a yeast DNA-binding protein with many targets, is essential for growth and bears some resemblance to the oncogene myb.* Mol Cell Biol, 1990. **10**(10): p. 5226-34.

Bibliographie

271. König, P., et al., *The crystal structure of the DNA-binding domain of yeast RAP1 in complex with telomeric DNA*. Cell, 1996. **85**(1): p. 125-36.
272. Moretti, P., et al., *Evidence that a complex of SIR proteins interacts with the silencer and telomere-binding protein RAP1*. Genes Dev, 1994. **8**(19): p. 2257-69.

ANNEXES

1. Numéros d'accèsion des protéines liées à la régulation de l'expression des gènes chez *P. falciparum*.

1.1 Protéines identifiées par nos soins.

| Numero d'accèsion | Superclasse ou fonction | Domaine caractéristique |
|-----------------------------|-------------------------|------------------------------|
| MAL13P1.61 | Architecture Bêta | cold shock |
| PFA0470c | Architecture Bêta | cold-shock protein, putative |
| PFL0145c | Architecture Bêta | HMG |
| MAL8P1.72 | Architecture Bêta | HMG |
| MAL13P1.290 | Architecture Bêta | HMG |
| PF10_0024 | Architecture Bêta | MADS box |
| PFC0345w | Domaines Basiques | B-box |
| PFE0895c | Domaines Basiques | B-box |
| MAL13P1.37 | Domaines Basiques | B-box |
| PFA0550w | Domaines Basiques | b-Zip |
| PFB0615c | Domaines Basiques | b-Zip |
| PFB0690w | Domaines Basiques | b-Zip |
| PFC0235w | Domaines Basiques | b-Zip |
| PFC0365w | Domaines Basiques | b-Zip |
| PFC0760c | Domaines Basiques | b-Zip |
| PFC0960c | Domaines Basiques | b-Zip |
| PFD0330w | Domaines Basiques | b-Zip |
| PFF0140c | Domaines Basiques | b-Zip |
| PFF0375c | Domaines Basiques | b-Zip |
| PFF1285w | Domaines Basiques | b-Zip |
| PF07_0038 | Domaines Basiques | b-Zip |
| MAL7P1.87 | Domaines Basiques | b-Zip |
| PFI0175w | Domaines Basiques | b-Zip |
| PF11_0268 | Domaines Basiques | b-Zip |
| PF11_0342 | Domaines Basiques | b-Zip |
| PF11_0392 | Domaines Basiques | b-Zip |
| PF11_0433 | Domaines Basiques | b-Zip |
| PFL1930w | Domaines Basiques | b-Zip |
| PFL2185w | Domaines Basiques | b-Zip |
| PF13_0047 | Domaines Basiques | b-Zip |
| MAL13P1.72 | Domaines Basiques | b-Zip |
| MAL13P1.97 | Domaines Basiques | b-Zip |
| MAL13P1.140 | Domaines Basiques | b-Zip |
| MAL13P1.188 | Domaines Basiques | b-Zip |
| MAL13P1.295 | Domaines Basiques | b-Zip |
| PFC0690c | Domaines Basiques | C2H2 |
| PFD0375w | Domaines Basiques | C2H2 |
| PFD0595w | Domaines Basiques | C2H2 |
| PFD0765w | Domaines Basiques | C2H2 |
| PFI0590c | Domaines Basiques | C2H2 |
| PFI1255w | Domaines Basiques | C2H2 |
| PF10_0091 | Domaines Basiques | C2H2 |
| PFL0455c | Domaines Basiques | C2H2 |
| PFL2075c | Domaines Basiques | C2H2 |
| PF13_0278 | Domaines Basiques | C2H2 |

| | | |
|-----------------------------|-------------------------------|----------------------------------|
| PF14_0479 | Domaines Basiques | C2H2 |
| PF14_0559 | Domaines Basiques | C2H2 |
| PF14_0612 | Domaines Basiques | C2H2 |
| PF14_0643 | Domaines Basiques | C2H2 |
| PF14_0657 | Domaines Basiques | C2H2 |
| PF14_0707 | Domaines Basiques | C2H2 |
| PF14_0383 | Domaines Basiques | Constans |
| PFB0140w | Domaines Basiques | DHHC-type zinc finger protein |
| PFB0725c | Domaines Basiques | DHHC-type zinc finger protein |
| PFE1415w | Domaines Basiques | DHHC-type zinc finger protein |
| PFF0485c | Domaines Basiques | DHHC-type zinc finger protein |
| MAL7P1.68 | Domaines Basiques | DHHC-type zinc finger protein |
| PFI0665w | Domaines Basiques | DHHC-type zinc finger protein |
| PFI1580c | Domaines Basiques | DHHC-type zinc finger protein |
| PF10_0273 | Domaines Basiques | DHHC-type zinc finger protein |
| PF11_0167 | Domaines Basiques | DHHC-type zinc finger protein |
| PF11_0217 | Domaines Basiques | DHHC-type zinc finger protein |
| MAL13P1.117 | Domaines Basiques | DHHC-type zinc finger protein |
| MAL13P1.126 | Domaines Basiques | DHHC-type zinc finger protein |
| PFL0465c | Domaines Basiques | krox |
| PFF0105w | Domaines Basiques | MYND finger domain protein |
| PFF0350w | Domaines Basiques | MYND finger protein |
| PF07_0124 | Domaines Basiques | MYND finger protein |
| PF13_0293 | Domaines Basiques | MYND finger protein |
| PFI1480w | Hélice-Tour-Hélice | Myb |
| PF11_0241 | Hélice-Tour-Hélice | Myb |
| PF10_0327 | Hélice-Tour-Hélice | Myb |
| PFL0815w | Hélice-Tour-Hélice | Myb |
| PFL1215c | Hélice-Tour-Hélice | Myb |
| PFE0420c | Hélice-Tour-Hélice | Myb |
| PFL1345c | Hélice-Tour-Hélice | Myb |
| PF13_0054 | Hélice-Tour-Hélice | Myb |
| PF13_0088 | Hélice-Tour-Hélice | Myb1 protein |
| PFL0290w | Facteur général | TFIIIB' |
| PF14_0333 | Mediator | MED10 Mediator component |
| PF14_0718 | Mediator | MED31 Mediator component |
| PF08_0037 | Mediator | MED7 Mediator component |
| PF11_0192 | Modification de la chromatine | moz acetyl transferase |
| PF11_0053 | Modification de la chromatine | PfSNF2L |
| PFF1440w | Modification de la chromatine | putative SET-domain protein |
| PFD0190w | Modification de la chromatine | SET |
| PFL0690c | Modification de la chromatine | SET |
| PF13_0293 | Modification de la chromatine | SET |
| PF14_0489 | Modification de la chromatine | sir2 |
| PF13_0152 | Modification de la chromatine | protein sir2 homologue, putative |
| PFB0730w | Remodelage de la chromatine | brahma snf2/iswi |
| PFF1185w | Remodelage de la chromatine | iswi protein homologue |
| PFF1185w | Remodelage de la chromatine | iswi protein homologue |
| PFE1140c | Remodelage de la chromatine | PhD |
| PF10_0079 | Remodelage de la chromatine | PhD |
| PF11_0429 | Remodelage de la chromatine | PhD |
| PFL0575w | Remodelage de la chromatine | PhD |

| | | |
|-----------------------------|-----------------------------|------|
| PFL1010c | Remodelage de la chromatine | PhD |
| PFL1210w | Remodelage de la chromatine | PhD |
| PFL1905w | Remodelage de la chromatine | PhD |
| MAL13P1.302 | Remodelage de la chromatine | PhD |
| PF14_0315 | Remodelage de la chromatine | PhD |
| PFF0225w | Remodelage de la chromatine | SNF2 |
| PF08_0048 | Remodelage de la chromatine | SNF2 |
| MAL8P1.65 | Remodelage de la chromatine | SNF2 |
| MAL8P1.73 | Remodelage de la chromatine | SNF2 |
| PF10_0232 | Remodelage de la chromatine | SNF2 |
| MAL13P1.216 | Remodelage de la chromatine | SNF2 |
| PF13_0308 | Remodelage de la chromatine | SNF2 |

1.2 Protéines identifiées par Coulson et collaborateurs [58].

| | |
|-------------|---|
| PFC0805w | RNA POLYMERASE II::Rpb-1 |
| PFB0715w | RNA POLYMERASE II::Rpb-2 |
| PF11130c | RNA POLYMERASE II::Rpb-3 |
| PFB0245c | RNA POLYMERASE II::Rpb-4 |
| PF13_0341 | RNA POLYMERASE I/II/III::Rpb-5 [ABC27] |
| PFC0155c | RNA POLYMERASE I/II/III::Rpb-6 [ABC23] |
| PF10_0269 | RNA POLYMERASE II::Rpb-7 |
| PFL0665c | RNA POLYMERASE I/II/III::Rpb-8 [ABC14.5] |
| PFA0505c | RNA POLYMERASE II::Rpb-9 |
| PF07_0027 | RNA POLYMERASE I/II/III::Rpb-10 [ABC10-beta] |
| PF13_0023 | RNA POLYMERASE II::Rpb-11 |
| MAL13P1_213 | RNA POLYMERASE I/II/III::Rpb-12 [ABC10-alpha] |
| PFE0465c | RNA POLYMERASE I::A190 |
| PF11_0358 | RNA POLYMERASE I::A135 |
| PFD0360w | RNA POLYMERASE I::A12.2 |
| PF11_0445 | RNA POLYMERASE I/III::AC40 |
| PF14_0150 | RNA POLYMERASE I/III::AC19 |
| PF13_0150 | RNA POLYMERASE III::C160 |
| PFL0330c | RNA POLYMERASE III::C128 |
| PF14_0207 | RNA POLYMERASE III::C34 |
| PF11_0058 | RNA POLYMERASE III::C25 |
| PFB0290c | RNA POLYMERASE III::C11 |
| PF11_0264 | MITOCHONDRIAL RNA POLYMERASE |
| PF14_0695 | APICOPLAST RNA POLYMERASE::alpha subunit |
| PF13_0040 | APICOPLAST RNA POLYMERASE::alpha subunit |
| PF10_0367 | APICOPLAST RNA POLYMERASE::beta, beta' & beta'' subunit |
| PF11_0315 | APICOPLAST RNA POLYMERASE::beta, beta' & beta'' subunit |
| PFE0635c | APICOPLAST RNA POLYMERASE::beta, beta' & beta'' subunit |
| PF10330c | APICOPLAST RNA POLYMERASE::beta, beta' & beta'' subunit |
| PFA0525w | TFIIB::TFIIB |
| PFE0415w | TFIIB::TFIIB-like |
| PFE0305w | TFIID::TBP |
| PF14_0267 | TFIID::TBP-like factor |
| MAL7P1_86 | TFIIE::TFIIE-alpha |
| PF10_0369 | TFIIH::XPB |
| PF11650w | TFIIH::XPD |
| PFL2125c | TFIIH::p52 |
| MAL13P1_76 | TFIIH::p44 |
| PF13_0279 | TFIIH::p34 |
| PFE0610c | TFIIH::MAT1 {PF00097} |
| PF13_0022 | TFIIH::cyclin K |
| PF14_0469 | TFIIIB::70/90 kDa subunits |
| PFL0520c | TFIIIC::63 kDa subunit |
| PF10_0293 | Transcription-elongation::SPT4 |
| MAL6P1_111 | Transcription-elongation::SPT5 |
| PF14_0059 | Transcription-elongation::SPT6 |
| PF07_0057 | Transcription elongation factor TFIIS |
| PFE0670w | Transcription elongation factor CDC68 |
| PFA0355w | CCR4-NOT complex::CCR4 |
| PFE0980c | CCR4-NOT complex::CCR4 |
| PFC0650c | CCR4-NOT complex::CCR4 |
| MAL8P1_104 | CCR4-NOT complex::CAF1 |
| PF11_0049 | CCR4-NOT complex::NOT1 |
| PF14_0170 | CCR4-NOT complex::NOT1 |
| PF11_0297 | CCR4-NOT complex::NOT2 |
| PF10_0062 | CCR4-NOT complex::NOT3/NOT5 |
| PFL1705w | CCR4-NOT complex::NOT4 |
| PF11_0477 | CCAAT BOX-BINDING COMPLEX::CBF-A/NF-Y-B {PF0080} |
| PF13_0043 | CCAAT BOX-BINDING COMPLEX::CBF-B/NF-Y-A |
| PF14_0374 | CCAAT BOX-BINDING COMPLEX::CBF-C/NF-Y-C {PF0080} |
| PF08_0034 | Histone acetyltransferase GCN5 {PF00439} |
| PF10_0143 | Transcriptional adaptor ADA2 {PF00249, PF00560} |
| PF11_0053 | Nucleosome remodelling::SNF2L |
| MAL6P1_183 | Nucleosome remodelling::ISWI {PF00628} |
| PF11260c | Histone deacetylase RPD3 |
| PF10_0078 | Histone deacetylase |
| PF14_0690 | Histone deacetylase |
| MAL7P1_37 | SIN3-REPRESSING COMPLEX::p18 |
| MAL6P1_269 | Transcription co-activator ALY |
| PF11_0293 | Transcription co-activator MBF1 {PF01381} |
| PFC0175w | Transcription co-activator TRIAD2 |
| PF14_0489 | Sir2-like protein |
| PF07_0083 | TBP-binding protein |
| PF14_0393 | Single-strand recognition protein |
| MAL6P1_142 | DNA methyltransferase 1-associated protein 1 DNMAP1 |
| PFB0875c | Chromatin-binding protein (SKI/SNW family) SkIP |
| PF14_0241 | General transcription factor BTF3 |
| PFL2390c | Pleiotropic regulatory protein THO2 |
| PF14_0718 | RNA polymerase II-associated factor SOH1 |
| PF14_0649 | Transcription factor C1 |
| MAL8P1_131 | Transcription factor GAS41 {PF03366} |
| PF07_0123 | Transcriptional regulator SPO8 |
| PF13_0293 | SET-domain transcriptional regulator |
| PFE0090w | Histone transcription regulator |
| PF11470c | Histone transcription regulator |
| PF14_0492 | Calcium-dependent transcriptional repressor DREAM |
| MAL8P1_111 | Transcriptional repressor RPH1 |

2. Article 4 : Vers une utilisation des bio-puces à ADN pour l'étude d'isolats plasmodiaux de terrain.

P. Rerour, M. Gissot, A. Siau, D. Mazier et C. Vaquero.

Accepté à la publication dans la *Revue de Médecine Tropicale*.

VERS UNE UTILISATION DES BIO-PUCES À ADN POUR L'ÉTUDE D'ISOLATS PLASMODIAUX DE TERRAIN

P. REFOUR, M. GISSOT, A. SIAU, D. MAZIER, C. VAQUERO

Med Trop 2004; **64** : 387-393

RÉSUMÉ • L'infection à *Plasmodium falciparum* est l'une des premières causes de mortalité dans les zones d'endémie plasmodiale. Face à l'augmentation des résistances du parasite aux thérapies, l'utilisation d'outils permettant l'analyse à grande échelle de la biologie de *P. falciparum* devient de plus en plus urgente. L'exploration du génome et du profil d'expression de l'ensemble des gènes de *P. falciparum* est désormais possible grâce à la technique des bio-puces. Cette technologie reste cependant limitée à l'étude d'échantillons dont les quantités d'ARN totaux sont de l'ordre de 8 à 50 µg. En effet, afin d'assurer la robustesse statistique de ce type d'expériences, au moins une réplique technique et une réplique biologique sont nécessaires. Cette contrainte technique exclut de fait des échantillons provenant de souches de terrain. De nombreuses méthodes ont été développées afin d'adapter ces techniques aux faibles quantités d'ARN total. L'utilisation de techniques d'amplification et des marqueurs différents des fluorochromes couramment utilisés, ayant pour vocation d'augmenter la sensibilité de détection, permet de diminuer la quantité de matériel nécessaire à chaque expérience. Cependant, les techniques d'amplification semblent introduire un biais dans les résultats générés. De même, l'emploi de nouveaux marqueurs très performants est encore limité du fait de contingences techniques. A l'heure actuelle, le marquage radioactif semble le mieux répondre aux exigences de l'utilisation de faibles quantités d'ARN totaux. Utilisable en hybridation compétitive, ce type d'approche combine tous les critères de criblage des bio-puces et a fait la preuve de son efficacité sur des bio-puces thématiques de *P. falciparum* développées dans notre laboratoire.

MOTS-CLÉS • Paludisme - Biopuce - Isolats de terrain - Radio-isotope.

MICROARRAY ANALYSIS OF MALARIA FIELD SAMPLES, AN OVERVIEW OF RECENT ADVANCES

ABSTRACT • *Plasmodium falciparum* is still a major cause of mortality in malaria endemic areas. Due to the wide spread drug resistance in the vast majority of the endemic countries, the use of clinical sample large-scale analysis approaches is needed. In addition to gene sequence data, postgenomic methods including microarray-based transcript profiling are now available. DNA microarrays authorize the analysis of *P. falciparum* RNA and DNA composition on a whole-genome scale using large amount of total RNA (8 µg to 50µg). In many cases, only small quantities of biological materials are available from *ex vivo* samples, leading to an insufficient yield of RNA for conventional microarray investigation. To overcome this difficulty, RNA amplification and signal enhancement have been employed, as well as alternative methods using labelled cDNA and/or RNA probes avoiding all amplification steps. However, these procedures, using fluorescent markers, remain complex, expensive, and sometimes biased, especially when using amplification. This review examines the various approaches to *Plasmodium* clinical sample transcript-profiling that could be suitable for DNA microarray analysis. Highlighting recent progress in alternate approaches dedicated to study of quite small biological materials we focused our scope on an approach involving two differentially radio-labelled probes ($^3\text{[H]-dATP}$ and $^{35}\text{[S]-dATP}$) simultaneously hybridised on a microarray glass slide. Based on major improvements of this technology for gene expression analysis from very low amounts of total RNA, such approach could be one of the most adapted to the studies of clinical samples that are major interest in malaria research.

KEY WORDS • Malaria - Microarray - Field samples - Radio-isotope.

• Travail de l'Unité INSERM U511, Immuno-Biologie Cellulaire et Moléculaire des Infections Parasitaires (P.R., Docteur de l'université Paris VI; M.G., A.S., Doctorants de l'université Paris VI; D.M., Docteur en médecine, Professeur de l'université Paris VI, Directeur de recherche et de l'unité INSERM U511; C.V., Docteur de l'université Paris VII, Directeur de recherche INSERM U511), CHU Pitié-salpêtrière, Université Pierre et Marie Curie Paris VI, Paris, France.

• Correspondance : P. REFOUR, INSERM U511, Immuno-Biologie Cellulaire et Moléculaire des Infections Parasitaires, CHU Pitié-salpêtrière, 91 Bd de l'hôpital, 75643 Paris cedex 13, France • Fax : 33 (0)1 45 83 88 58.

• Courriel : refour @ext.jussieu.fr ; vaquero @ext.jussieu.fr •

• Article reçu le 1/12/2003, définitivement accepté le 18/05/2004.

La complétion du séquençage et l'annotation graduelle du génome de *Plasmodium falciparum* ouvrent des perspectives de recherche nouvelles pour mieux comprendre la complexité biologique du parasite. Cependant, élucider les mécanismes biologiques impliqués dans la résistance, la virulence et la pathogenèse de *Plasmodium falciparum* nécessite d'étendre les investigations menées sur les souches adaptées de laboratoire, à l'étude des isolats plasmodiaux de terrain. L'exploitation des données issues du séquençage du clone 3D7 a permis de prédire approximativement 5 300 cadres

ouverts de lecture (1). Dès lors, les études de génomique fonctionnelle impliquent l'utilisation de techniques d'investigation à haut débit. Parmi ces approches, les puces à ADN ou bio-puces, sont particulièrement propices à ce type d'analyse à grande échelle. Il apparaît cependant nécessaire d'adapter des techniques utilisées communément dans le laboratoire au matériel de terrain.

La technique des bio-puces est devenue en quelques années une technique incontournable pour l'analyse simultanée des variations de l'abondance des ARN messagers (ARNm) de plusieurs milliers de gènes, dans des cellules biologiquement différentes (2). Cependant, bien que l'utilisation majeure des bio-puces soit vouée à l'étude des transcriptomes, cette approche à l'origine utilisée pour le séquençage et le génotypage (3, 4), pourrait ainsi permettre d'étudier la biodiversité des génomes parasitaires (5).

Les premières puces à ADN de *P. falciparum* ont été élaborées aux Etats-Unis, par dépôt de produits de PCR à partir d'une banque d'ADN génomique aléatoire (3 648 fragments) (6) ou d'une banque d'ADN complémentaire (ADNc) (944 fragments) (7). Depuis peu, un répertoire d'environ 7 500 oligonucléotides d'une taille moyenne de 70 mers correspondant au génome entier du clone 3D7 (Operon®, malaria oligoset), et comprenant plusieurs cibles par gène pour les grands messagers, est disponible. Ce répertoire a récemment été utilisé pour, d'une part valider l'approche expérimentale par oligonucléotides longs (8), et d'autre part étudier le profil d'expression des gènes au cours du cycle érythrocytaire en préparant des échantillons correspondant à chaque heure au cours des 48 h du développement érythrocytaire du clone HB3 (9). De plus, deux bio-puces prototypes ont été mises au point par la société Affymetrix® avec des oligonucléotides de 25 mers. La première, avec 4 167 oligonucléotides correspondant à la presque totalité du Chromosome 2 de *P. falciparum*, clone 3D7, a permis de valider la technologie des oligonucléotides courts par deux approches expérimentales (5, 10). La seconde, couvrant la totalité du génome de 3D7 comprend les régions codantes, intergéniques et des séquences antisens (260 000 oligonucléotides). Cette dernière ayant fait l'objet d'une étude de sept stades du cycle érythrocytaire dont le stade gamécyte, et le stade sporozoïte (11), constitue une base de données utile à la communauté (données disponibles sur: <http://plasmodb.org/>), même si un certain nombre de discordances sont observés entre les données issues de ces deux expériences mettant en œuvre le génome complet de *P. falciparum* (9, 11). Il n'en reste pas moins que l'accessibilité de telles puces est encore très restreinte, avec un coût non négligeable.

Bien que l'utilisation de bio-puces tende à se généraliser au sein des laboratoires travaillant sur Plasmodium, leur application à l'étude du transcriptome et du génome d'échantillons prélevés en zone d'endémie, se heurte à plusieurs problèmes de faisabilité. Outre l'hétérogénéité des isolats tant du point de vue de la synchronisation des échantillons (stades anneaux 0 à 18 h après pénétration dans l'érythrocyte) que du point de vue du nombre de clones présents (génotypes différents), l'obstacle majeur d'une telle méthodologie reste la faible quantité de matériel biologique

disponible. Suivant le protocole utilisé le coût peut être très important, et si l'on veut utiliser des bio-puces pour l'étude d'isolats de terrain, les contraintes expérimentales de répliques techniques et biologiques, ainsi que le nombre d'échantillons nécessaires à une étude, imposent l'utilisation de techniques sensibles et peu coûteuses. A l'heure actuelle, la plupart des études par approche bio-puce sont basées sur l'utilisation des ADNc marqués par 2 fluorochromes différents (Cyanine 3: cy3 et Cyanine 5: cy5) et nécessite environ 8 µg à 50 µg d'ARN totaux (12) soit un minimum de 2 x 10⁸ parasites synchronisés au stade anneau. Dans cette revue, après une description du principe des bio-puces et des différentes approches expérimentales, nous insisterons sur les différents protocoles permettant d'augmenter la détection du signal et donc de diminuer la quantité de matériel biologique nécessaire à ce type d'étude.

PRINCIPE ET DIVERSITÉ DES BIO-PUCES

Le principe des bio-puces repose sur l'hybridation de deux acides nucléiques de séquences complémentaires : l'une, définie comme « cible » (soit des fragments d'amplification par PCR, soit des oligonucléotides correspondant aux gènes à étudier), est immobilisée sur un support, l'autre la « sonde » (correspondant aux ADNc produits lors de la rétrotranscription), qui une fois marquée va interagir avec la cible et après révélation, donner un signal. On peut ainsi analyser et comparer simultanément deux échantillons distincts discriminés par les deux marqueurs, où l'un est l'échantillon utilisé comme référence et l'autre l'échantillon test. Les gènes sont dits « sur-exprimés » ou « sous-exprimés » par rapport à l'échantillon référence, lorsque le rapport des valeurs obtenues entre les deux échantillons (ratio) est supérieur à une valeur seuil définie (en général x2 ou /2).

Plusieurs types de bio-puces existent à ce jour, répertoriées selon le support, la densité et la nature des dépôts. Trois types de support sont employés, les membranes de nylon, les lames de verre recouvertes d'un ligand (poly-L-lysine, aminosilane, aldéhyde, epoxy, matrice de gel polyacrylamide) et les dalles de silice (13, 14). Ces différents supports déterminent la technologie de dépôts des cibles donc le type de cible, ainsi que le type de marqueur utilisable pour la détection.

Les membranes de nylon à haute densité permettent jusqu'à 20 000 dépôts sur une surface de 8 x 12 cm², tandis que les lames de verres autorisent jusqu'à 30 000 dépôts mais sur une surface de 2 x 5,5 cm². Cependant, certaines approches tendent à combiner la densité de dépôts obtenues sur lame de verre et les supports membranes (15). Quant aux supports de silice, il est possible de concentrer jusqu'à 1 x 10⁶ cibles sur une surface 1,3 x 1,3 cm² (16).

Le matériel génétique déposé sur le support peut être de deux natures, soit des produits de PCR, soit des oligonucléotides. Il semble qu'il n'y ait aucune différence de sensibilité de détection entre l'utilisation de produits de PCR ou d'oligonucléotides longs (50 mers) (17). A notre connaissance aucune donnée publiée n'a fait état de différence de

Tableau I - Approches expérimentales utilisées pour baisser le seuil de détection.

| | Fluorescence | Amplification d'ARN | | Marquage d'ARN | | Amplification du signal | | | Différents marqueurs plus sensibles | |
|---------------------------------------|-------------------------------------|-----------------------------|---------------|-------------------|---------------------|-------------------------|------------------|--|-------------------------------------|---------------|
| | | Amplification T7 polymérase | Amplification | Marquage indirect | Haptènes de Platine | Dendrimers | TSA ^a | RCA ^b | Chimiluminescence | Radioactivité |
| Hybridation différentielle | Oui | Oui | Oui | Oui | Oui | Oui | Non | Oui | Oui | |
| Type de cibles | PCR ^c
OS ^d | PCR
OA ^e | PCR
OS | PCR
OS
OA | PCR
OS | PCR
OS | PCR
OS | PCR
OS | PCR
OS | |
| Quantité minimale d'ARN total requise | 8µg | 0,3 µg à 3µg | 2 µg | 2µg | 2,5 µg | 1,5 µg | ND ^f | < 1µg | 0,5 µg à 1µg | |
| Nombre de répliques/ 1 échantillon | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 à 10 | 1 à 5 | |
| Références | (6) | (30) | (33) | (34) | (40) | (41) | (42) | Communication personnelle P. Ravassard | Données personnelles | |

^a : Tyramide Signal Amplification ; ^b : Rolling Circle Amplification ; ^c : Produits PCR ; ^d : Oligonucléotides Sens ; ^e : Oligonucléotides Antisens ; ^f : Non déterminé.

sensibilité entre produits de PCR et oligonucléotides court (20 mers). En revanche des différences de profils d'expression ont été observées entre des expériences menées sur ces deux types de cibles (18). Enfin, les supports utilisés ainsi que les méthodes de marquage des sondes sont déterminants quant aux limites de détection (19) et à l'échelle du criblage. D'ailleurs, l'innovation majeure dans le domaine du criblage à haut débit a été la miniaturisation du *pattern* de dépôt des cibles. Ainsi, le remplacement des membranes par des lames de verre a permis de diminuer et les volumes d'hybridation et la quantité de matériels immobilisés, avec une haute densité de dépôts.

Selon la méthode de détection des signaux, et donc le type de marquage des ADNc, la limite de détection peut varier d'un facteur dix (entre fluorescence et radioactivité par exemple) définissant ainsi la quantité minimale d'ARN totaux nécessaire pour une expérience. Le nombre, la diversité des méthodes de détection et des techniques d'amplification du signal et/ou d'ARN total décrites dans la littérature, montrent que baisser le seuil de détection reste une priorité dans le développement des bio-puces. Plusieurs types de marquage ont en effet été mis au point utilisant des marqueurs fluorescents (2, 20), enzymatique [chimiluminescence (21)] et radioactifs (22, 23). Bien que les marqueurs fluorescents soient les plus utilisés, ils présentent : 1) le désavantage de nécessiter beaucoup d'ARN totaux (8 à 50 µg) si l'on compare aux autres techniques tel que le marquage par radio-isotopes ; 2) un seuil de détection élevé, ne permettant pas l'analyse des transcrits peu représentés.

La limite de détection à partir de petites quantités de matériel biologique (<5µg) peut être abaissée de plusieurs façons ; 1) amplification des ARN totaux ; 2) marquage direct et indirect des ARN totaux ou néo-synthétisés ; 3) amplification des ADNc ; 4) amplification du signal des ADNc marqués ; 5) utilisation de marqueurs plus sensibles pour la détec-

tion. Ces approches (Tableau I) sont développées dans le paragraphe suivant.

APPROCHES PERMETTANT D'ABAISSE LE SEUIL DE DÉTECTION

Amplification des ARN messagers

Plusieurs techniques d'amplification des ARN et du signal ont été développées, afin d'exploiter les échantillons de faible quantité. Parmi les méthodes d'amplification des ARN, la plus utilisée est l'amplification des ARNm par l'ARN polymérase du bactériophage T7 (24). Cette technique repose sur la synthèse (transcription *in vitro*) d'ARN « antisens » à partir d'ADNc double brin lui-même synthétisé à partir de l'échantillon d'ARN total à tester, en présence d'un oligonucléotides poly dTTP couplé en 5' au promoteur spécifique de l'ARN polymérase T7. Les transcrits antisens ainsi générés sont alors convertis en ADNc marqués par incorporation directe ou indirecte de marqueurs fluorescents. La contrainte majeure de cette amplification est qu'elle n'autorise pas l'utilisation de bio-puces avec des cibles simple brin (oligonucléotides) correspondant aux séquences sens, ce qui est le cas de la majorité des bio-puces commerciales dont le répertoire d'oligonucléotides de *P. falciparum* commercialisé par Operon® (8).

Il n'en reste pas moins que cette approche, par les biais qu'elle génère, reste controversée. En effet, bien que certaines expériences aient été menées sans comparaison avec des résultats obtenus en l'absence d'amplification (11, 25, 26), on observe lorsque cette comparaison est effectuée, une discordance dans la distribution des groupes de gènes repartis selon les ratios obtenus entre deux échantillons biologiques (27-31). Une étude statistique menée afin d'évaluer l'effet seul de l'amplification (32), a permis de montrer que

selon les classes de ratios déterminés pour discriminer les gènes différemment exprimés (par exemple x2 et /2), les ratios ne sont pas tous préservés. Et ce, même si l'approche avec amplification induit des intensités d'hybridation plus élevées qu'un marquage classique et une reproductibilité satisfaisante. De plus, pour des quantités d'ARN totaux inférieures à 0,3 µg, il est nécessaire de procéder à deux amplifications successives pour obtenir un résultat proche de celui obtenu avec une seule amplification (possible seulement à partir de plus de 0,3 µg); ce qui augmente d'autant le coût de l'expérience. Cette technique reste cependant une alternative pour des quantités de matériel de départ de 0,3 µg à 3 µg, lorsque l'on compare deux échantillons amplifiés.

Marquage direct et indirect des ARN totaux ou néo synthétisés

Deux techniques de marquage des ARN totaux ou néo synthétisés utilisés directement comme sonde pour les biopuces ont été décrites. La première repose sur le marquage des transcrits *in vitro* en incorporant le dUTP-aminoallyl au cours de l'amplification des ARNm avec l'ARN T7 polymérase, qui après couplage à des fluorochromes libres, permet d'augmenter l'efficacité de marquage de 2 à 3 fois (33), et l'utilisation de bio-puces commerciales avec des oligonucléotides sens. Cependant, les effets sur la représentativité des ARNm dus à la technique d'amplification, pourraient être du même ordre que ceux décrits ci avant. L'autre technique consiste en un marquage chimique direct des ARN totaux avec des fluorochromes couplés à des complexes (haptènes) de platine (34). Très prometteuse, cette technique peut être utilisée à partir de 2 µg d'ARN totaux. Elle ne permet cependant, ni l'utilisation de bio-puces contenant des oligonucléotides sens, ni une réplique technique car tout l'échantillon marqué est directement utilisé; elle reste de plus pour l'instant extrêmement coûteuse.

Amplification des ADNc

Il est aussi possible d'amplifier les ADNc générés après transcription inverse des ARN totaux [méthode SMART PCR (35), SPA (36)], mais la représentativité des ARNm d'un échantillon dont les ADNc ont été générés par ce type de technique, est cependant relative [Siau *et Coll* données personnelles; (36-38)].

Amplification du signal des ADNc marqués

Trois techniques sont aujourd'hui disponibles pour amplifier le signal des ADNc marqués: technologie des dendrimers (40), Tyramide Signal Amplification (TSA), (41) et Rolling Circle Amplification (RCA), (42). La technologie des dendrimers repose sur l'utilisation de macromolécules d'acides nucléiques agrégés et très fortement marquées. En effet, les parties simple brin de surface du complexe viennent s'hybrider à une séquence incluse, au préalable, dans la partie 5' des oligonucléotides poly T utilisés pour la transcription inverse des échantillons à tester. La technologie de TSA est basée sur l'accumulation d'un marqueur activé *in situ* par la tyramide, *via* les sondes marquées à la biotine conjuguées

à un système streptavidine-peroxydase. L'amplification du signal au fur et à mesure de l'accumulation du marqueur est proportionnelle à la concentration de sondes hybridées à la cible. Enfin, le système RCA utilise un ADN circulaire simple brin hybridé sur une amorce spécifique immobilisée, pour amplifier à température constante la séquence circulaire en continu. Le matériel obtenu après amplification correspond à un concatémère de l'ADN circulaire ancré par l'oligonucléotide amorce et ayant incorporé des dNTP marqués par un fluorochrome, augmentant ainsi le signal de détection. Ces techniques ne semblent, pour l'instant, pas utilisées en routine, et ce type d'approche reste limité car on ne peut effectuer de répliques à partir d'un même échantillon (Tableau I).

Utilisation de marqueurs plus sensibles pour la détection

Plusieurs marqueurs différents des fluorochromes peuvent être utilisés, et deux d'entre eux, les plus couramment utilisés seront décrits. Le marquage avec des marqueurs enzymatiques puis révélation par chimiluminescence (21) est le système le plus sensible, mais ne permet pas pour l'instant de quantifier les signaux (réponse tout ou rien). De plus, cette technique perd de son intérêt du fait des phénomènes de diffusion, de saturation et d'extinction rapide du signal. L'utilisation encore expérimentale, de microcapillaires pour focaliser l'émission de lumière, et de la mesure du temps d'extinction du signal, devrait permettre d'éliminer ces désavantages, autorisant même une détection en différentiel en utilisant différents couples d'enzymes/révélateurs (43).

Les marqueurs radioactifs semblent être les mieux adaptés aux faibles quantités de matériel biologique, car ils ont l'avantage de baisser le seuil de détection par rapport aux marqueurs fluorescents et ce en l'absence d'amplification. Il est certain que ce type de marquage engendre des contraintes d'utilisation non négligeables, mais plusieurs avancées récentes ont permis d'en éliminer un certain nombre. En effet, l'une des limitations majeures était la nécessité d'utiliser des membranes indépendantes pour comparer les échantillons tests à l'échantillon de référence, car il n'était pas possible de discriminer les 2 isotopes. Cette impossibilité imposait aussi de normaliser les résultats à l'aide d'une hybridation préalable avec un oligonucléotide marqué afin d'évaluer la qualité et la quantité des dépôts. Cependant, l'avantage des membranes est de pouvoir, après déshybridation, être réutilisées 5 à 10 fois, réduisant d'autant le coût des expériences. Récemment, la possibilité d'utiliser deux radio-isotopes avec des lames de verre a été décrite (44) permettant d'avoir recours aux lames conçues initialement pour les expériences en fluorescence. Le fait de faire usage des marqueurs radioactifs sur des systèmes miniaturisés présente plusieurs avantages, notamment celui de réduire les volumes d'hybridation (et les volumes de liquides contaminés), ce qui a pour effet de diminuer le seuil de détection. Néanmoins, avec de telles densités de cible, la résolution spatiale des appareils de détection des signaux radioactifs classiques (Phosphor imager) est limitante

(40µm), la discrimination efficace des signaux produits par deux cibles distinctes immobilisées sur lames de verre devient alors difficile.

La détection d'émissions radioactives sur des supports à haute densité n'a été rendue possible que par l'utilisation d'une caméra CCD (charge-coupled device) à haute résolution (45). De plus, ce type de détecteur effectue le comptage des désintégrations en temps réel, rendant possible la quantification de l'ensemble des signaux hétérogènes sans phénomène de saturation pour les transcrits abondants. Enfin, l'avancée majeure est la possibilité d'hybrider en différentiel deux échantillons sur un même support puisque la discrimination de 2 radio-isotopes est actuellement possible (46). En effet, la filtration des émissions radioactives est basée sur la discrimination de la taille du diamètre de l'émission des photons produits lors de l'impact des micro éléments contre une fine couche de scintillant immobilisé sur une feuille de mylar® (Dupont®). L'image générée par la totalité des photons est filtrée et deux images sont alors constituées selon la contribution de chaque radio-isotope. La sensibilité, l'utilisation d'un support de verre, la possibilité d'exploiter des bio-puces disponibles, ainsi que la possibilité d'hybrider en différentiel avec un comptage en temps réel sont des atouts majeurs qui nous ont conduits à choisir préférentiellement cette méthode pour l'exploitation de petites quantités d'ARN totaux et permettre d'envisager l'application aux prélèvements de terrain.

BIO-PUCE THÉMATIQUE DE *P. FALCIPARUM* ET HYBRIDATION DIFFÉRENTIELLE D'ADNc MARQUÉS AVEC DEUX RADIOSOTOPES

Nous avons choisi d'adapter, sur le site de la Pitié-Salpêtrière (Plateforme Post-génomique Pitié-Salpêtrière, P3S), la technique de double marquage radioactif (47) à notre problématique d'analyse du transcriptome de *P. falciparum*. Une bio-puce thématique a été élaborée, regroupant 150 produits de PCR correspondant à des séquences de gènes impliqués dans le cycle cellulaire et la transduction du signal (~ 3% des cadres ouverts de lecture du génome complet, prédits). Ces produits sont immobilisés sur lames de verre, avec une densité de 400 µm de centre à centre, soit une surface totale de 1 cm² avec 2 répliques par gène. L'hybridation différentielle de deux échantillons d'ADNc marqués avec deux isotopes différents, ³[H]-dATP et ³⁵[S]-dATP, génère avec un faible bruit de fond, des signaux sans saturation au cours de la détection par comptage linéaire en temps réel sur 24 h (et plus si nécessaire). L'acquisition des désintégrations des deux radio-isotopes est effectuée par un Microimager (45) (Biospace®, Paris) à une résolution de 15 µm pour le ³[H], et 20 µm pour le ³⁵[S]. L'image générée est ensuite filtrée à l'aide d'un algorithme de discrimination des signaux donnés respectivement par le ³[H]-dATP et le ³⁵[S]-dATP, (Dual Labelling®, Biospace®). Enfin, les données sont extraites (Arrayvision®, Imaging Research®), normalisées et analysées (Genespring®, Silicon Genetics®).

Cette technique de détection donne des résultats reproductibles, avec une sensibilité nous permettant d'exploiter des échantillons rares (=1µg). Les résultats obtenus à partir de différents marquages et préparations d'ARN extraits au même stade du cycle sont reproductibles (Gissot *et Coll*, manuscrit soumis). Enfin, différentes expériences menées sur des gammes de concentration en ARN totaux utilisées au départ dans la réaction de transcription inverse, montrent qu'il serait possible d'appliquer ce type de technique aux isolats de terrain (Refour *et Coll*, manuscrit soumis). En effet, à partir d'un microgramme d'ARN total par échantillon il nous est possible d'hybrider en différentiel cinq lames précédemment décrites et cette même quantité pourrait donc être utilisée pour hybrider des puces correspondant au génome complet de Plasmodium.

DISCUSSION

L'application des bio-puces à l'étude des isolats de terrain nécessite de prendre en compte toutes les contraintes techniques (quantités suffisantes de matériel de départ, reproductibilité, répliques, nombre d'échantillons, etc.), ainsi que les contraintes inhérentes à la dynamique des infections (souches asynchrones, populations polyclonales, polymorphisme des souches étudiées, ...). Cependant, l'analyse à haut débit est nécessaire, que les études soient génomiques : mise en évidence de la diversité du génome des souches de terrain, remaniements chromosomiques (délétions, mutations, polymorphismes, ...), identification des *loci* associés à la résistance et à la virulence du parasite, ou qu'elles soient de type post-génomique fonctionnelle : caractérisation des gènes et des voies métaboliques modulés sous pression médicamenteuse, identification de groupes de gènes caractéristiques d'isolats potentiellement impliqués dans les accès graves, anémie, neuropaludisme, etc.

Dans l'emploi de bio-puces, pour l'étude du génome mais plus encore pour l'étude du transcriptome, l'une des contraintes la plus importante est l'extraction de matériel biologique en quantité suffisante, en particulier pour la réalisation de plusieurs répliques expérimentales. En effet, les schémas expérimentaux communément utilisés (48) préconisent au moins une réplique expérimentale dite technique, et une réplique expérimentale dite biologique soit au moins 4 expériences par condition. La réalisation de ces différentes répliques est indispensable pour la validation et la robustesse de l'étude statistique des données expérimentales. Si l'on estime qu'à partir d'un prélèvement de cinq millilitres de sang total avec une parasitémie de 0,5 % (soit 1x10⁸ parasites), et une population parasitaire synchronisée au stade anneau, il est possible d'obtenir environ 5 µg d'ARN totaux, on comprend la nécessité de mettre en œuvre une approche expérimentale très sensible. La qualité et la quantité d'ARN total obtenu à partir d'un échantillon, est un paramètre déterminant pour les rendements de marquage. L'extraction d'ARN de qualité à partir de faible quantité de matériel biologique reste une étape délicate. Bien qu'un certain nombre de techniques et de nombreux kits soient disponibles, la technique utilisant le Trizol®

(Invitrogen) (49), et un charrier d'ARN lors de l'étape de précipitation, reste la plus efficace en termes de rendement et de qualité. En effet, l'efficacité de lyse au Trizol® du parasite est supérieure aux autres méthodes et évite le traitement à la saponine. Enfin, cette méthode d'extraction en phase liquide permet de maîtriser la concentration finale de l'échantillon.

Une autre contrainte est de pouvoir quantifier la concentration du matériel de départ et sa qualité, la limite actuelle se situant aux alentours de 50 ng à 100 ng grâce à un système d'électrophorèse capillaire développé par la société Agilent® (Biochip®). Il est en effet important de standardiser les rendements d'extraction en fonction de la parasitémie et selon la méthode d'extraction, afin de normaliser les quantités utilisées pour chaque expérience.

Bien que les formes circulantes de *P. falciparum* soient des formes anneaux, l'échantillon sanguin comprend le plus souvent plusieurs isolats plus ou moins synchrones. En effet, un décalage dans le cycle cellulaire est possible sans distinction morphologique à ce stade (4 h à 12 h de décalage), revenant à étudier le transcriptome d'un mélange d'anneaux jeunes et âgés. Or, les méthodes de séparation/synchronisation *ex vivo* induisent un stress à même de perturber le transcriptome du parasite. De même pour les prélèvements de faible parasitémie, la nécessité d'effectuer une maturation *ex vivo* des isolats afin d'augmenter la quantité de matériel disponible, pourrait avoir des effets non négligeables sur le transcriptome. Ces étapes de multiplication cellulaire sélectionnent les parasites capables de se développer *ex vivo* ce qui augmente la clonalité des cultures et pourrait induire un biais dans les données obtenues lors de l'étude du génome et du transcriptome.

Un grand nombre d'études de terrain pourraient bénéficier des propriétés de criblage à grande échelle des bio-puces. Et l'un des axes possibles de recherche serait d'identifier des groupes de gènes parasitaires associés soit aux accès pernicieux, soit aux accès simples. Il n'en reste pas moins qu'en plus du parasite qui évidemment participe à la gravité de l'infection, la susceptibilité de l'hôte intervient dans le développement des formes pernicieuses. De plus, il est important de souligner le fait que les parasites présents dans la voie sanguine, seul stade pouvant être étudié, puissent exprimer des gènes différents de ceux exprimés dans les parasites séquestrés dans le cerveau.

Enfin, bien que les possibles hybridations croisées entre l'homme et le parasite aient été évaluées de manière indépendante (5), il est important de tester l'impact sur le transcriptome d'une hybridation compétitive entre des ADNc issus d'ARN totaux humains et plasmodiaux. Aussi, il est nécessaire de prendre en compte le choix des cibles utilisées pour ces différentes études, car celles disponibles actuellement sont générées à partir des données issues du séquençage du clone 3D7. Il sera donc nécessaire d'évaluer le degré de divergence entre le génome des isolats de terrain et ce clone modèle (5). La définition de cibles spécifiques pourrait conduire à la caractérisation des différentes souches et espèces de *Plasmodium*, utilisables pour des études épidémiologiques.

CONCLUSION

L'utilisation de sondes radioactives pour un criblage à haut débit, qu'il concerne l'étude du génome ou du transcriptome de *P. falciparum* fait que les contraintes d'étude des souches de terrain ne seront plus d'ordre quantitatif mais plutôt d'ordre qualitatif. Il est raisonnable d'envisager que les isolats de terrain puissent bénéficier d'une étude globale du transcriptome et du génome. Néanmoins, l'étude du transcriptome sur bio-puce(s) pangénomique(s) n'est qu'une première étape de criblage, permettant de définir des familles de gènes caractéristiques de différents phénotypes (résistance, accès pernicieux, etc.). Les futures applications de terrain seront vraisemblablement orientées vers l'utilisation de bio-puces « thématiques » telles que celle que nous avons développée et contenant des familles de gènes caractéristiques de ces différents phénotypes.

Remerciements • Les auteurs tiennent à remercier l'INSERM, la région Ile-de-France, la Fondation de Recherche HMR-Aventis, et l'Université Paris VI, pour leur soutien à la Plateforme Post-génomique de la Pitié-Salpêtrière (P3S). P. Refour et M. Gissot ont été financés par le programme PAL+.

RÉFÉRENCES

- GARDNER MJ, HALL N, FUNG E *et Coll* - Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* 2002; **419** : 498-511.
- SCHENA M, SHALON D, DAVIS RW *et Coll* - Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 1995; **270** : 467-470.
- DRMANAC R, LABAT I, BRUKNER I *et Coll* - Sequencing of megabase plus DNA by hybridization : theory of the method. *Genomics* 1989; **4** : 114-128.
- KHRAPKO KR, LYSOV YU P, KHORLYN AA *et Coll* - An oligonucleotide hybridization approach to DNA sequencing. *FEBS Lett* 1989; **256** : 118-122.
- VOLKMAN SK, HARTL DL, WIRTH DF *et Coll* - Excess polymorphisms in genes for membrane proteins in *Plasmodium falciparum*. *Science* 2002; **298** : 216-218.
- HAYWARD RE, DERISI JL, ALFADHLI S *et Coll* - Shotgun DNA microarrays and stage-specific gene expression in *Plasmodium falciparum* malaria. *Mol Microbiol* 2000; **35** : 6-14.
- BEN MAMOUN C, GLUZMAN IY, HOTT C *et Coll* - Co-ordinated programme of gene expression during asexual intraerythrocytic development of the human malaria parasite *Plasmodium falciparum* revealed by microarray analysis. *Mol Microbiol* 2001; **39** : 26-36.
- BOZDECH Z, ZHU J, JOACHIMIAK MP *et Coll* - Expression profiling of the schizonte and trophozoite stages of *Plasmodium falciparum* with a long-oligonucleotide microarray. *Genome Biol* 2003; **4** : R9.
- BOZDECH Z, LLINAS M, PULLIAM BL *et Coll* - The Transcriptome of the Intraerythrocytic Developmental Cycle of *Plasmodium falciparum*. *PLoS Biol* 2003; **1** : E5.
- LE ROCH KG, ZHOU Y, BATALOV S *et Coll* - Monitoring the chromosome 2 intraerythrocytic transcriptome of *Plasmodium falciparum* using oligonucleotide arrays. *Am J Trop Med Hyg* 2002; **67** : 233-243.
- LE ROCH KG, ZHOU Y, BLAIR PL *et Coll* - Discovery of gene function by expression profiling of the malaria parasite life cycle. *Science* 2003; **301** : 1503-1508.
- RATHOD PK, GANESAN K, HAYWARD RE *et Coll* - DNA microarrays for malaria. *Trends Parasitol* 2002; **18** : 39-45.

- 13 - SOUTHERN EM, MASKOS U, ELDER JK - Analyzing and comparing nucleic acid sequences by hybridization to arrays of oligonucleotides : evaluation using experimental models. *Genomics* 1992; **13** : 1008-1017.
- 14 - LIVACHE T, ROGET A, DEJEAN E *et Coll* - Preparation of a DNA matrix via an electrochemically directed copolymerization of pyrrole and oligonucleotides bearing a pyrrole group. *Nucleic Acids Res* 1994; **22** : 2915-2921.
- 15 - BERTUCCI F, SALAS S, EYSTERIES S *et Coll* - Gene expression profiling of colon cancer by DNA microarrays and correlation with histoclinical parameters. *Oncogene* 2004; **23** : 1377-1391.
- 16 - LIPSHUTZ RJ, FODOR SP, GINGERAS TR *et Coll* - High density synthetic oligonucleotide arrays. *Nat Genet* 1999; **21** : 20-24.
- 17 - KANE MD, JATKOE TA, STUMPF CR *et Coll* - Assessment of the sensitivity and specificity of oligonucleotide (50mer) microarrays. *Nucleic Acids Res* 2000; **28** : 4552-4557.
- 18 - KUO WP, JENSSSEN TK, BUTTE AJ *et Coll* - Analysis of matched mRNA measurements from two different microarray technologies. *Bioinformatics* 2002; **18** : 405-412.
- 19 - DUGGAN DJ, BITTNER M, CHEN Y *et Coll* - Expression profiling using cDNA microarrays. *Nat Genet* 1999; **21** : 10-14.
- 20 - PEASE AC, SOLAS D, SULLIVAN EJ *et Coll* - Light-generated oligonucleotide arrays for rapid DNA sequence analysis. *Proc Natl Acad Sci USA* 1994; **91** : 5022-5026.
- 21 - RAJEEVAN MS, DIMULESCU IM, UNGER ER *et Coll* - Chemiluminescent analysis of gene expression on high-density filter arrays. *J Histochem Cytochem* 1999; **47** : 337-342.
- 22 - LENNON GG, LEHRACH H - Hybridization analyses of arrayed cDNA libraries. *Trends Genet* 1991; **7** : 314-317.
- 23 - NGUYEN C, ROCHA D, GRANJEAUD S *et Coll* - Differential gene expression in the murine thymus assayed by quantitative hybridization of arrayed cDNA clones. *Genomics* 1995; **29** : 207-216.
- 24 - VAN GELDER RN, VON ZASTROW ME, YOOL A *et Coll* - Amplified RNA synthesized from limited quantities of heterogeneous cDNA. *Proc Natl Acad Sci USA* 1990; **87** : 1663-1667.
- 25 - KAMME F, SALUNGA R, YU J *et Coll* - Single-cell microarray analysis in hippocampus CA1 : demonstration and validation of cellular heterogeneity. *J Neurosci* 2003; **23** : 3607-3615.
- 26 - MCCLINTICK JN, JEROME RE, NICHOLSON CR *et Coll* - Reproducibility of oligonucleotide arrays using small samples. *BMC Genomics* 2003; **4** : 4.
- 27 - BAUGH LR, HILL AA, BROWN EL *et Coll* - Quantitative analysis of mRNA amplification by in vitro transcription. *Nucleic Acids Res* 2001; **29** : E29.
- 28 - HU L, WANG J, BAGGERLY K *et Coll* - Obtaining reliable information from minute amounts of RNA using cDNA microarrays. *BMC Genomics* 2002; **3** : 16.
- 29 - ZHAO H, HASTIE T, WHITFIELD ML *et Coll* - Optimization and evaluation of T7 based RNA linear amplification protocols for cDNA microarray analysis. *BMC Genomics* 2002; **3** : 31.
- 30 - WANG E, MILLER LD, OHNMACHT GA *et Coll* - High-fidelity mRNA amplification for gene profiling. *Nat Biotechnol* 2000; **18** : 457-459.
- 31 - SAGHIZADEH M, BROWN DJ, TAJBAKHSJ J *et Coll* - Evaluation of techniques using amplified nucleic acid probes for gene expression profiling. *Biomol Eng* 2003; **20** : 97-106.
- 32 - NYGAARD V, LOLAND A, HOLDEN M *et Coll* - Effects of mRNA amplification on gene expression ratios in cDNA experiments estimated by analysis of variance. *BMC Genomics* 2003; **4** : 11.
- 33 - T HOEN PA, DE KORT F, VAN OMMEN GJ *et Coll* - Fluorescent labelling of cRNA for microarray applications. *Nucleic Acids Res* 2003; **31** : E20.
- 34 - GUPTA V, CHERKASSKY A, CHATIS P *et Coll* - Directly labeled mRNA produces highly precise and unbiased differential gene expression data. *Nucleic Acids Res* 2003; **31** : E13.
- 35 - FRANZ O, BRUCHHAUS II, ROEDER T - Verification of differential gene transcription using virtual northern blotting. *Nucleic Acids Res* 1999; **27** : E3.
- 36 - SMITH L, UNDERHILL P, PRITCHARD C *et Coll* - Single primer amplification (SPA) of cDNA for microarray expression analysis. *Nucleic Acids Res* 2003; **31** : E9.
- 37 - VERNON SD, UNGER ER, RAJEEVAN M *et Coll* - Reproducibility of alternative probe synthesis approaches for gene expression profiling with arrays. *J Mol Diagn* 2000; **2** : 124-127.
- 38 - ISCOVE NN, BARBARA M, GUM M *et Coll* - Representation is faithfully preserved in global cDNA amplified exponentially from sub-pico-gram quantities of mRNA. *Nat Biotechnol* 2002; **20** : 940-943.
- 39 - PUSKAS LG, ZVARA A, HACKLER L, JR. *et Coll* - RNA amplification results in reproducible microarray data with slight ratio bias. *Biotechniques* 2002; **32** : 1330-1334, 1336, 1338, 1340.
- 40 - STEARS RL, GETTS RC, GULLANS SR - A novel, sensitive detection system for high-density microarrays using dendrimer technology. *Physiol Genomics* 2000; **3** : 93-99.
- 41 - KARSTEN SL, VAN DEERLIN VM, SABATTI C *et Coll* - An evaluation of tyramide signal amplification and archived fixed and frozen tissue in microarray gene expression analysis. *Nucleic Acids Res* 2002; **30** : E4.
- 42 - NALLUR G, LUO C, FANG L *et Coll* - Signal amplification by rolling circle amplification on DNA microarrays. *Nucleic Acids Res* 2001; **29** : E118.
- 43 - CHEEK BJ, STEEL AB, TORRES MP *et Coll* - Chemiluminescence detection for hybridization assays on the flow-thru chip, a three-dimensional microchannel biochip. *Anal Chem* 2001; **73** : 5777-5783.
- 44 - WHITNEY LW, BECKER KG - Radioactive 33-P probes in hybridization to glass cDNA microarrays using neural tissues. *J Neurosci Methods* 2001; **106** : 9-13.
- 45 - LANIECE P, CHARON Y, CARDONA A *et Coll* - A new high resolution radioimager for the quantitative analysis of radiolabelled molecules in tissue section. *J Neurosci Methods* 1998; **86** : 1-5.
- 46 - SALIN H, MAITREJEAN S, MALLETT J *et Coll* - Sensitive and quantitative co-detection of two mRNA species by double radioactive *in situ* hybridization. *J Histochem Cytochem* 2000; **48** : 1587-1592.
- 47 - SALIN H, VUJASINOVIC T, MAZURIE A *et Coll* - A novel sensitive microarray approach for differential screening using probes labelled with two different radioelements. *Nucleic Acids Res* 2002; **30** : e17.
- 48 - YANG Y, SPEED T - Design issues for cDNA microarray experiments. *Nat Rev Genet* 2002; **3** : 579-588.
- 49 - KYES S, PINCHES R, NEWBOLD C - A simple RNA analysis method shows var and rif multigene family expression patterns in *Plasmodium falciparum*. *Mol Biochem Parasitol* 2000; **105** : 311-315.

3. Données et matériels supplémentaires de l'article 1.

Supplementary data 1. Comparison of 3D7 clone transcriptome data obtained from our experiments

to Scripps/GNF Malaria Array data using 3D7 clone and De Risi Plasmodium falciparum HB3 time course microarray data.

Genes that fit (+) or do not fit (-) the data are listed. Genes whose expression patterns were not available are marked "none".

"ND" means that the expression profile was discordant between sorbitol and temperature synchronised culture in Scripps/GNF Malaria Array data. Average of three experiments for each clone is shown

| gene name | R | ET | LT | ES | LS | accession number | Le Roch <i>et al.</i> | Bozdech <i>et al.</i> |
|-----------------|---|-------|-------|-------|-------|------------------|-----------------------|-----------------------|
| CHR13 41763 | 0 | -0,47 | 0,93 | 0,09 | -0,69 | MAL13P1.196 | + | + |
| ZnSnf2 | 0 | -0,9 | 0,68 | -2,12 | -0,5 | MAL13P1.216 | + | + |
| CHR13 40872 | 0 | -0,46 | -0,7 | -0,97 | 0,97 | MAL13P1.278 | + | + |
| PHd2 | 0 | -1,01 | -1,77 | -1,87 | -0,27 | MAL13P1.76 | + | + |
| CHR13 40983 | 0 | 0,27 | 0,36 | -0,06 | 0,24 | MAL13P1.78 | + | + |
| PF CDPK2 | 0 | -0,13 | 0,48 | 0,14 | -0,1 | MAL6P1.108 | + | - |
| PF PK2 | 0 | 0,95 | 0,16 | -0,44 | 0,04 | MAL6P1.108 | + | - |
| PF PK4 | 0 | -0,34 | -0,07 | -0,02 | -0,3 | MAL6P1.146 | + | + |
| CHRBLOB 36958 | 0 | 0,05 | -1,9 | -2,81 | 1,11 | MAL6P1.191 | + | + |
| cdc48 | 0 | 0,79 | 0,41 | 1,56 | 1,31 | MAL6P1.232 | + | + |
| HISTONE H3 | 0 | 0,48 | 0,34 | 0,83 | 1,44 | MAL6P1.248 | none | none |
| HISTONE H2A | 0 | -0,38 | 0,28 | 1,07 | 1,86 | MAL6P1.249 | none | + |
| CHRBLOB 68156 | 0 | 0,32 | 1,16 | 0,59 | -0,39 | MAL6P1.271 | + | + |
| PF FEST | 0 | 0,14 | 0,83 | 0,24 | 0,29 | MAL7P1.91 | + | + |
| PF BO520W | 0 | 1,25 | 0,35 | 0,21 | 0,14 | PF BO520W | - | - |
| PF BO665W | 0 | 0,24 | 0,21 | 0,56 | 1,29 | PF BO665W | + | + |
| PF BO815W | 0 | -0,24 | -0,66 | 0,8 | 1,75 | PF BO815W | + | + |
| PF STARP | 0 | -0,16 | -0,29 | -1,2 | -2,65 | PF07_0006 | + | + |
| PF CDC ATP | 0 | -0,65 | 0,46 | -0,35 | -1,02 | PF07_0047 | + | + |
| C2143 | 0 | -0,12 | 0,67 | 1,86 | -2,81 | PF07_0073 | + | + |
| CHRBLOB BU60D08 | 0 | -0,12 | -0,84 | -1,22 | -0,19 | PF08_0044 | nd | + |
| PF MYB3 | 0 | 0,01 | -0,26 | 0,83 | -0,16 | PF10_0143 | + | + |
| GBP130 | 0 | 0,08 | 1,33 | 0,45 | -0,12 | PF10_0159 | + | + |
| PF PUGBP130 | 0 | -0,61 | 0,98 | -0,13 | -0,92 | PF10_0159 | + | + |
| ADP-RIBO FAC1 | 0 | -0,14 | 0,8 | -0,16 | 0,1 | PF10_0203 | - | - |
| PFMYB2 | 0 | -0,52 | -1,21 | -0,73 | -0,53 | PF10_0327 | + | + |

| | | | | | | | | |
|-------------|---|-------|-------|-------|-------|-----------|------|------|
| HISTONE H4 | 0 | 0,04 | 0,63 | 1,43 | 2,21 | PF11_0061 | + | - |
| HISTONE H2B | 0 | 1,03 | 1,09 | 2,08 | 2,44 | PF11_0062 | + | - |
| CHR11 S189 | 0 | 0,33 | 0,38 | 0,37 | 0,23 | PF11_0096 | + | + |
| PF ERB | 0 | -0,77 | 0,35 | 0,93 | 0,85 | PF11_0098 | + | + |
| PF ERC | 0 | 0,67 | 0,63 | 1,09 | 0,78 | PF11_0098 | + | + |
| PF S40 | 0 | 0,36 | 0,35 | 1,1 | 0,82 | PF11_0098 | + | + |
| PF MAP2 | 0 | -1,27 | -1,87 | -1,94 | -1,77 | PF11_0147 | nd | + |
| PF RAN-1 | 0 | -0,15 | 0,09 | -0,15 | -0,09 | PF11_0183 | + | + |
| CHR11 S51 | 0 | -0,37 | 0,7 | 1,08 | 1,2 | PF11_0220 | + | + |
| PF CK1 | 0 | -0,37 | -1,59 | 0,4 | 0,42 | PF11_0377 | + | + |
| CHR11 S127 | 0 | -0,96 | -1,5 | -0,13 | 1,45 | PF11_0464 | + | + |
| CYCLIN ANIA | 0 | -0,65 | -0,24 | -1,86 | -1,01 | PF13_0022 | none | - |
| PFMYB1 | 0 | -0,27 | -0,69 | -1,09 | -0,55 | PF13_0088 | - | - |
| RNA POLBP | 0 | -0,17 | -2,13 | -3,3 | -0,01 | PF13_0123 | none | none |
| RNA POLIII | 0 | -0,08 | -0,46 | -0,45 | -0,35 | PF13_0150 | + | + |
| PF PK6 | 0 | -0,34 | 1,43 | 2,08 | -1,02 | PF13_0206 | + | none |
| CHR13 32969 | 0 | 0,59 | -0,99 | 2,11 | 1,11 | PF13_0211 | + | + |
| CHR13 21780 | 0 | -0,16 | -0,26 | -0,03 | -0,1 | PF13_0258 | - | + |
| PF PCNA | 0 | -0,6 | 1,06 | 2,22 | 2,04 | PF13_0328 | + | + |
| PF RNR2 | 0 | 1,03 | 1,22 | 2,03 | 1,48 | PF14_0053 | + | + |
| PF ACTII | 0 | -0,31 | -0,39 | -0,67 | -0,91 | PF14_0124 | + | + |
| PP1-LIKE | 0 | -0,53 | 0,37 | -0,19 | -0,32 | PF14_0224 | - | - |
| D11CY3 | 0 | -0,52 | -0,96 | -1 | 0,28 | PF14_0316 | + | + |
| PF TOPO 2 | 0 | 0,13 | -0,22 | -0,72 | -0,01 | PF14_0316 | + | + |
| CALMODULIN | 0 | 0,78 | 1,07 | 1,42 | 1,58 | PF14_0323 | + | + |
| CHR14 132 | 0 | 0,12 | -0,38 | -1,25 | -0,89 | PF14_0346 | - | - |
| PF RNR1 | 0 | 0,58 | 0,14 | 1,08 | 1,31 | PF14_0352 | + | + |
| DNA PRIM | 0 | -0,3 | 0,42 | 0,37 | -0,27 | PF14_0366 | + | + |
| KRU | 0 | -0,2 | 0,02 | 1,2 | -0,12 | PF14_0383 | nd | + |
| CHR9 9716 | 0 | 0,3 | 0,47 | 0,52 | 0,35 | PF14_0392 | + | + |
| CHR14 128 B | 0 | 0,32 | -0,25 | -1 | -0,55 | PF14_0423 | + | + |
| CALCINEURIN | 0 | -0,11 | 0,34 | -0,03 | -0,51 | PF14_0492 | - | - |
| PF CYC-1 | 0 | 0,06 | -0,22 | -1,04 | 0,28 | PF14_0605 | - | + |
| KHARP | 0 | 0,37 | -1,34 | -1,35 | -2 | PFB0100c | + | + |
| RPB4-LIKE | 0 | 0,47 | -0,26 | -1 | 0,29 | PFB0245c | + | + |
| PF RAD2 | 0 | -0,57 | 0,28 | -1,12 | -2,01 | PFB0265c | + | - |

| | | | | | | | | |
|------------------|---|-------|-------|-------|-------|----------|----|------|
| TFIIS | 0 | 0,23 | -0,23 | -1,53 | -1,55 | PFB0290c | + | + |
| YOU2 | 0 | 1,32 | -0,08 | 0,04 | 0,22 | PFB0620w | + | + |
| PF MBP | 0 | -0,59 | -0,1 | 1,29 | 0,02 | PFB0725c | + | none |
| DNA HELICASE | 0 | -1,17 | -0,97 | -0,24 | -2,12 | PFB0730w | + | - |
| PF CPK | 0 | 0,05 | -0,5 | -0,21 | 2,61 | PFB0815w | + | + |
| SPOU | 0 | 0,06 | 0,96 | -0,29 | -0,23 | PFB0855c | + | + |
| RNA HELICASE | 0 | 1,14 | -1,38 | -1,58 | -1,41 | PFB0860c | + | + |
| PF SNRNP | 0 | 0,14 | 0,67 | 0,05 | 0,18 | PFB0865w | + | - |
| RSUA | 0 | 0,13 | 1,14 | 0,51 | -0,79 | PFB0890c | + | + |
| PFC0060C | 0 | -0,55 | -1,68 | -0,95 | 0,02 | PFC0060c | + | + |
| RNA POL1 | 0 | 0,37 | 0,34 | -1,88 | -1,35 | PFC0155c | + | + |
| PFC0195W | 0 | 0,01 | 0,67 | -0,13 | -0,62 | PFC0195w | + | + |
| EIF-4A | 0 | -0,04 | 0,69 | -0,51 | -0,61 | PFC0445w | nd | none |
| PFC0485W | 0 | 0,07 | -0,57 | -2,31 | -0,82 | PFC0485w | nd | + |
| PFC0510W | 0 | 1,67 | 1,86 | 3,41 | 2,05 | PFC0510w | + | + |
| PFC0525C | 0 | -0,17 | -0,62 | 0,68 | -0,51 | PFC0525c | + | + |
| PFC0635C | 0 | 0,1 | -0,68 | -0,53 | 0,39 | PFC0635c | + | + |
| PFC0740C | 0 | -0,48 | -0,54 | -0,83 | -0,04 | PFC0740c | + | + |
| PFC0755C | 0 | -0,6 | 0,01 | 1,34 | 1,22 | PFC0755C | + | + |
| PFC0805W | 0 | 1,03 | 0,46 | 0,21 | 0,29 | PFC0805w | + | + |
| PFC0825C | 0 | 0,62 | 0,91 | 0,29 | 0,43 | PFC0825C | nd | - |
| PFC0915W | 0 | -0,2 | 0,88 | 0,75 | -0,84 | PFC0915w | + | + |
| PFC0920W | 0 | -0,41 | 0,92 | 0,17 | -0,51 | PFC0920w | + | + |
| HOX 1 | 0 | -0,07 | -0,32 | 0,4 | 0,46 | PFD0340c | + | + |
| PF MT C | 0 | -0,68 | -2,12 | -1,5 | 0,38 | PFD0370w | - | - |
| PF POLA | 0 | -0,83 | 0,71 | 1,7 | 2,56 | PFD0590c | + | + |
| PF CRK2 | 0 | -0,31 | -0,31 | -0,6 | -1,1 | PFD0740w | + | + |
| SAR1 | 0 | -0,71 | 0,42 | 0,52 | 0,27 | PFD0810w | + | + |
| PF CRK1 | 0 | -0,48 | -0,85 | 0,74 | -1,06 | PFD0865c | + | + |
| HOX 2 | 0 | 0,36 | -0,46 | 0,24 | 0,97 | PFD1160w | nd | - |
| PF ATBP | 0 | -0,85 | -0,85 | -3,74 | -1,31 | PFE0305w | nd | + |
| CHR5 21208 | 0 | -0,4 | 0,61 | 0,78 | -0,14 | PFE0420c | + | + |
| PF TOPO 1 | 0 | -0,5 | -0,82 | 0,79 | 1,04 | PFE0520c | - | - |
| PFCYC-3 (TRYP-2) | 0 | -0,01 | -1,22 | -3,53 | -1,33 | PFE0920c | nd | - |
| EF-1B | 0 | 0,27 | 0,68 | -0,11 | -0,5 | PFI0645w | + | + |
| PF B7 | 0 | 0,69 | 0,44 | 0,49 | 0,15 | PFI0930c | - | none |

| | | | | | | | | |
|------------------|---|-------|-------|-------|-------|----------|----|------|
| 3-PHOSPHO | 0 | 0,84 | 1,85 | 1 | -1,31 | PFI1105w | + | + |
| PF PHOS | 0 | 0,2 | 1,46 | 0,03 | -0,91 | PFI1245c | + | + |
| CHR9 IX78E09,P1C | 0 | 0,83 | -0,96 | -0,42 | -1,47 | PFI1280c | + | + |
| PF PKA | 0 | 0,82 | 0,41 | 2,08 | 2,84 | PFI1685w | + | + |
| CHR11 N3190 | 0 | 0,47 | 0,45 | 1,03 | 0,64 | PFL0080c | nd | + |
| PF HMG1 | 0 | -0,24 | 0,15 | 0,5 | 0,4 | PFL0145c | + | + |
| PF HMG2 | 0 | -0,3 | -0,17 | 0,11 | 0,15 | PFL0290w | + | + |
| KROX | 0 | 0,09 | 0,36 | 0,85 | 0,98 | PFL0465c | + | none |
| NF-LIK1 | 0 | -0,58 | -0,18 | -2,09 | -1 | PFL1030w | - | - |
| CHR12 06 | 0 | -0,43 | -1,47 | -2,12 | -1,17 | PFL1285c | - | - |
| PFCYC-2 | 0 | -0,56 | -0,88 | 0,23 | 0,1 | PFL1330c | + | + |
| PF NEK-1 | 0 | 0,03 | 0,9 | 0,9 | 0,47 | PFL1370w | + | + |
| phdB | 0 | 0,79 | 0,51 | -0,85 | -0,67 | PFL1905w | + | none |
| FSTH | 0 | 0,18 | 1,13 | 1,69 | 1,48 | PFL1925w | + | + |
| PF TBP | 0 | 0,14 | 0,97 | 0,98 | 0,84 | PFL2345c | + | + |

Supplementary data 2. Comparison of transcriptome data obtained from our experiments for 3D7 and F12 clone.

3D7 clone transcriptome data are shown in black. F12 clone transcriptome data are shown in green.

Average of three experiments for each clone is shown

| gene name | R | ET | LT | ES | LS | accession number |
|---------------|---|-------|-------|-------|-------|------------------|
| 3-PHOSPHO | 0 | 0,84 | 1,85 | 1 | -1,31 | PF11105w |
| 3-PHOSPHO | 0 | -1,04 | 2,79 | 0,06 | 1,36 | |
| ADP-RIBO FAC1 | 0 | -0,14 | 0,8 | -0,16 | 0,1 | PF10_0203 |
| ADP-RIBO FAC1 | 0 | 0,1 | 0,67 | -0,27 | -0,04 | |
| C2143 | 0 | -0,12 | 0,67 | 1,86 | -2,81 | PF07_0073 |
| C2143 | 0 | 1,03 | 0,39 | -0,82 | -1,18 | |
| CALCINEURIN | 0 | -0,11 | 0,34 | -0,03 | -0,51 | PF14_0492 |
| CALCINEURIN | 0 | -0,27 | 0,32 | 0,01 | -0,01 | |
| CALMODULIN | 0 | 0,78 | 1,07 | 1,42 | 1,58 | PF14_0323 |
| CALMODULIN | 0 | 0,53 | 1,13 | 1,83 | 1,27 | |
| cdc48 | 0 | 0,79 | 0,41 | 1,56 | 1,31 | MAL6P1.232 |
| cdc48 | 0 | 0,74 | 1,3 | 1,6 | 0,38 | |
| CHR11 N3190 | 0 | 0,47 | 0,45 | 1,03 | 0,64 | PFL0080c |
| CHR11 N3190 | 0 | 0,49 | -0,49 | 1,62 | -0,81 | |
| CHR11 S127 | 0 | -0,96 | -1,5 | -0,13 | 1,45 | PF11_0464 |
| CHR11 S127 | 0 | -1,26 | -0,83 | -0,62 | 0,15 | |
| CHR11 S189 | 0 | 0,33 | 0,38 | 0,37 | 0,23 | PF11_0096 |
| CHR11 S189 | 0 | -0,27 | 0,4 | 0,35 | -0,29 | |
| CHR11 S51 | 0 | -0,37 | 0,7 | 1,08 | 1,2 | PF11_0220 |
| CHR11 S51 | 0 | -0,48 | 0,84 | 1,59 | 2,7 | |
| CHR12 06 | 0 | -0,43 | -1,47 | -2,12 | -1,17 | PFL1285c |
| CHR12 06 | 0 | -1,23 | -1,53 | -1,69 | -0,1 | |
| CHR13 21780 | 0 | -0,16 | -0,26 | -0,03 | -0,1 | PF13_0258 |
| CHR13 21780 | 0 | -1,35 | 0,72 | 1,42 | 0,93 | |
| CHR13 32969 | 0 | 0,59 | -0,99 | 2,11 | 1,11 | PF13_0211 |
| CHR13 32969 | 0 | 0,51 | 0,51 | 2,31 | 1,65 | |
| CHR13 40872 | 0 | -0,46 | -0,7 | -0,97 | 0,97 | MAL13P1.278 |
| CHR13 40872 | 0 | -1,02 | -0,87 | -0,92 | 0,23 | |
| CHR13 40983 | 0 | 0,27 | 0,36 | -0,06 | 0,24 | MAL13P1.78 |
| CHR13 40983 | 0 | 0,57 | 0,55 | -1,14 | -0,84 | |
| CHR13 41763 | 0 | -0,47 | 0,93 | 0,09 | -0,69 | MAL13P1.196 |

| | | | | | | |
|------------------|---|-------|-------|-------|-------|------------|
| CHR13 41763 | 0 | -0,01 | 0,44 | 0,46 | -0,57 | |
| CHR14 128 B | 0 | 0,32 | -0,25 | -1 | -0,55 | PF14_0423 |
| CHR14 128 B | 0 | 0,95 | -0,72 | -1,19 | -0,27 | |
| CHR14 132 | 0 | 0,12 | -0,38 | -1,25 | -0,89 | PF14_0346 |
| CHR14 132 | 0 | 0,35 | 0,5 | -0,88 | -1 | |
| CHR5 21208 | 0 | -0,4 | 0,61 | 0,78 | -0,14 | PFE0420c |
| CHR5 21208 | 0 | 0,32 | 0,67 | -0,2 | -0,37 | |
| CHR9 9716 | 0 | 0,3 | 0,47 | 0,52 | 0,35 | PF14_0392 |
| CHR9 9716 | 0 | 0,22 | 0,6 | 0,77 | 0,65 | |
| CHR9 IX78E09,P1C | 0 | 0,83 | -0,96 | -0,42 | -1,47 | PF11280c |
| CHR9 IX78E09,P1C | 0 | 0,92 | -0,27 | -0,45 | -0,08 | |
| CHRBLOB 36958 | 0 | 0,05 | -1,9 | -2,81 | 1,11 | MAL6P1.191 |
| CHRBLOB 36958 | 0 | -0,12 | -0,93 | -0,75 | 0,89 | |
| CHRBLOB 68156 | 0 | 0,32 | 1,16 | 0,59 | -0,39 | MAL6P1.271 |
| CHRBLOB 68156 | 0 | -0,4 | 0,62 | 0,5 | -0,46 | |
| CHRBLOB BU60D08 | 0 | -0,12 | -0,84 | -1,22 | -0,19 | PF08_0044 |
| CHRBLOB BU60D08 | 0 | -0,12 | -0,84 | -1,22 | -1,19 | |
| CYCLIN ANIA | 0 | -0,65 | -0,24 | -1,86 | -1,01 | PF13_0022 |
| CYCLIN ANIA | 0 | -0,25 | -0,59 | -0,88 | -0,78 | |
| D11CY3 | 0 | -0,52 | -0,96 | -1 | 0,28 | PF14_0316 |
| D11CY3 | 0 | 0,11 | -1,47 | -1,41 | -1,33 | |
| DNA HELICASE | 0 | -1,17 | -0,97 | -0,24 | -2,12 | PFB0730w |
| DNA HELICASE | 0 | -0,88 | -0,16 | 0,79 | -0,86 | |
| DNA PRIM | 0 | -0,3 | 0,42 | 0,37 | -0,27 | PF14_0366 |
| DNA PRIM | 0 | 0,17 | 0,62 | 0,36 | 0,36 | |
| EF-1B | 0 | 0,27 | 0,68 | -0,11 | -0,5 | PF10645w |
| EF-1B | 0 | 0,53 | 0,69 | -0,17 | -0,72 | |
| EIF-4A | 0 | -0,04 | 0,69 | -0,51 | -0,61 | PFC0445w |
| EIF-4A | 0 | -0,42 | 0,34 | 0,07 | -1,14 | |
| FSTH | 0 | 0,18 | 1,13 | 1,69 | 1,48 | PFL1925w |
| FSTH | 0 | -0,92 | 1,35 | 1,79 | 1,8 | |
| GBP130 | 0 | 0,08 | 1,33 | 0,45 | -0,12 | PF10_0159 |
| GBP130 | 0 | 0,37 | 1,26 | 0,45 | -0,7 | |
| HISTONE H2A | 0 | -0,38 | 0,28 | 1,07 | 1,86 | MAL6P1.249 |
| HISTONE H2A | 0 | 0,97 | 1,08 | 1,13 | 0,7 | |
| HISTONE H2B | 0 | 1,03 | 1,09 | 2,08 | 2,44 | PF11_0062 |

| | | | | | | |
|-------------|---|-------|-------|-------|-------|------------|
| HISTONE H2B | 0 | -1,31 | 1,82 | 1,98 | 2,02 | |
| HISTONE H3 | 0 | 0,48 | 0,34 | 0,83 | 1,44 | MAL6P1.248 |
| HISTONE H3 | 0 | -0,95 | 0,71 | 2,02 | 0,74 | |
| HISTONE H4 | 0 | 0,04 | 0,63 | 1,43 | 2,21 | PF11_0061 |
| HISTONE H4 | 0 | 0,4 | 0,23 | 1,44 | 1,78 | |
| HOX 1 | 0 | -0,07 | -0,32 | 0,4 | 0,46 | PFD0340c |
| HOX 1 | 0 | -0,19 | 0,33 | 0,24 | 0,18 | |
| HOX 2 | 0 | 0,36 | -0,46 | 0,24 | 0,97 | PFD1160w |
| HOX 2 | 0 | -0,66 | -0,05 | 0,03 | 0,32 | |
| KHARP | 0 | 0,37 | -1,34 | -1,35 | -2 | PFB0100c |
| KHARP | 0 | 0,2 | -0,2 | -1,73 | -1,7 | |
| KROX | 0 | 0,09 | 0,36 | 0,85 | 0,98 | PFL0465c |
| KROX | 0 | -0,21 | 0,22 | 0,63 | 0,7 | |
| KRU | 0 | -0,2 | 0,02 | 1,2 | -0,12 | PF14_0383 |
| KRU | 0 | 1,27 | -0,03 | -0,33 | -0,23 | |
| NF-LIK1 | 0 | -0,58 | -0,18 | -2,09 | -1 | PFL1030w |
| NF-LIK1 | 0 | 0,23 | 0,04 | -2,04 | -0,8 | |
| PF ACTII | 0 | -0,31 | -0,39 | -0,67 | -0,91 | PF14_0124 |
| PF ACTII | 0 | 0,25 | -0,08 | -0,34 | -0,48 | |
| PF ATBP | 0 | -0,85 | -0,85 | -3,74 | -1,31 | PFE0305w |
| PF ATBP | 0 | -0,75 | -0,54 | -1,3 | 0,55 | |
| PF B7 | 0 | 0,69 | 0,44 | 0,49 | 0,15 | PFI0930c |
| PF B7 | 0 | -0,53 | 0,95 | 2,09 | 0,93 | |
| PF BO520W | 0 | 1,25 | 0,35 | 0,21 | 0,14 | PF BO520W |
| PF BO520W | 0 | 1,17 | 0,69 | -0,27 | 0,22 | |
| PF BO665W | 0 | 0,24 | 0,21 | 0,56 | 1,29 | PF BO665W |
| PF BO665W | 0 | 0,01 | -0,23 | -0,28 | 0,29 | |
| PF BO815W | 0 | -0,24 | -0,66 | 0,8 | 1,75 | PF BO815W |
| PF BO815W | 0 | -0,88 | 0,9 | 3,01 | 1,38 | |
| PF CDC ATP | 0 | -0,65 | 0,46 | -0,35 | -1,02 | PF07_0047 |
| PF CDC ATP | 0 | 0,14 | 0,23 | 0,17 | -0,44 | |
| PF CDPK2 | 0 | -0,13 | 0,48 | 0,14 | -0,1 | MAL6P1.108 |
| PF CDPK2 | 0 | -0,07 | 0,27 | -0,49 | -0,65 | |
| PF CK1 | 0 | -0,37 | -1,59 | 0,4 | 0,42 | PF11_0377 |
| PF CK1 | 0 | -0,58 | -1,34 | 0,02 | 0,34 | |
| PF CPK | 0 | 0,05 | -0,5 | -0,21 | 2,61 | PFB0815w |

| | | | | | | |
|----------|---|-------|-------|-------|-------|------------|
| PF CPK | 0 | 0,41 | -1,29 | 0,59 | 1,34 | |
| PF CRK1 | 0 | 0,48 | 0,85 | 0,74 | 0,56 | PFD0865c |
| PF CRK1 | 0 | 0,85 | 0,88 | 1,35 | 0,92 | |
| PF CRK2 | 0 | -0,31 | -0,31 | -0,6 | -1,1 | PFD0740w |
| PF CRK2 | 0 | -0,4 | -0,16 | -1,08 | -0,87 | |
| PF CYC-1 | 0 | 0,06 | -0,22 | -1,04 | 0,28 | PF14_0605 |
| PF CYC-1 | 0 | -0,51 | -0,79 | -0,45 | -0,3 | |
| PF ERB | 0 | -0,77 | 0,35 | 0,93 | 0,85 | PF11_0098 |
| PF ERB | 0 | -1,2 | 0,74 | 3,19 | 0,86 | |
| PF ERC | 0 | 0,67 | 0,63 | 1,09 | 0,78 | PF11_0098 |
| PF ERC | 0 | -0,64 | 1,33 | 2,75 | 1,65 | |
| PF FEST | 0 | 0,14 | 0,83 | 0,24 | 0,29 | MAL7P1.91 |
| PF FEST | 0 | -0,53 | 0,98 | 0,28 | 0,45 | |
| PF HMG1 | 0 | -0,24 | 0,15 | 0,5 | 0,4 | PFL0145c |
| PF HMG1 | 0 | -0,42 | 0,34 | 0,64 | 0,56 | |
| PF HMG2 | 0 | 0,22 | -0,26 | 0,22 | 0,37 | PFL0290w |
| PF HMG2 | 0 | -0,3 | -0,17 | 0,11 | 0,15 | |
| PF MAP2 | 0 | -1,27 | -1,87 | -1,94 | -1,77 | PF11_0147 |
| PF MAP2 | 0 | -1,56 | -1,94 | -1,88 | -0,79 | |
| PF MBP | 0 | -0,59 | -0,1 | 1,29 | 0,02 | PFB0725c |
| PF MBP | 0 | -0,36 | -0,15 | 1,03 | 0,47 | |
| PF MT C | 0 | -0,68 | -2,12 | -1,5 | 0,38 | PFD0370w |
| PF MT C | 0 | -3,05 | -3,37 | -3,63 | -1,86 | |
| PF MYB3 | 0 | 0,01 | 0,46 | 0,83 | -0,16 | PF10_0143 |
| PF MYB3 | 0 | -0,97 | 1,05 | 1,56 | 1,37 | |
| PF NEK-1 | 0 | 0,03 | 0,9 | 0,9 | 0,47 | PFL1370w |
| PF NEK-1 | 0 | 0,02 | 0,66 | 1,16 | 0,39 | |
| PF PCNA | 0 | -0,6 | 1,06 | 2,22 | 2,04 | PF13_0328 |
| PF PCNA | 0 | 0,23 | 1,25 | 4,91 | 2,54 | |
| PF PHOS | 0 | 0,2 | 1,46 | 0,03 | -0,91 | PF11245c |
| PF PHOS | 0 | 0,39 | 1,16 | 0,65 | 0,23 | |
| PF PK2 | 0 | 0,95 | 0,16 | -0,44 | 0,04 | MAL6P1.108 |
| PF PK2 | 0 | 0,68 | 0,05 | -0,07 | 0,32 | |
| PF PK4 | 0 | -0,34 | -0,07 | -0,02 | -0,3 | MAL6P1.146 |
| PF PK4 | 0 | -0,34 | 0,89 | -0,33 | 0,15 | |
| PF PK6 | 0 | -0,34 | 1,43 | 2,08 | -1,02 | PF13_0206 |

| | | | | | | |
|-------------|---|-------|-------|-------|-------|-----------|
| PF PK6 | 0 | -0,22 | 1,22 | 2,66 | -0,02 | |
| PF PKA | 0 | 0,82 | 0,41 | 2,08 | 2,84 | PF11685w |
| PF PKA | 0 | -1,49 | 0,58 | 2,03 | 2,05 | |
| PF POLA | 0 | -0,83 | 0,71 | 1,7 | 2,56 | PFD0590c |
| PF POLA | 0 | -1,38 | 0,21 | 1,66 | 0,3 | |
| PF PUGBP130 | 0 | -0,61 | 0,98 | -0,13 | -0,92 | PF10_0159 |
| PF PUGBP130 | 0 | -0,52 | 0,8 | 0,14 | 0,11 | |
| PF RAD2 | 0 | -0,57 | 0,28 | -1,12 | -2,01 | PFB0265c |
| PF RAD2 | 0 | -0,18 | 0,34 | -0,54 | -0,71 | |
| PF RAN-1 | 0 | -0,15 | 0,81 | -0,15 | -0,09 | PF11_0183 |
| PF RAN-1 | 0 | -0,36 | 0,64 | 0,2 | 0,13 | |
| PF RNR1 | 0 | 0,58 | 0,14 | 1,08 | 1,31 | PF14_0352 |
| PF RNR1 | 0 | -0,11 | 1,13 | 2,32 | 2,41 | |
| PF RNR2 | 0 | 1,03 | 1,22 | 2,03 | 1,48 | PF14_0053 |
| PF RNR2 | 0 | -0,42 | 0,39 | 1,92 | 1,7 | |
| PF S40 | 0 | 0,36 | 0,35 | 1,1 | 0,82 | PF11_0098 |
| PF S40 | 0 | -0,55 | -0,48 | 1,55 | 0,97 | |
| PF SNRNP | 0 | 0,14 | 0,67 | 0,05 | 0,18 | PFB0865w |
| PF SNRNP | 0 | -0,16 | -0,33 | -0,39 | -0,7 | |
| PF STARP | 0 | -0,16 | -0,29 | -1,2 | -2,65 | PF07_0006 |
| PF STARP | 0 | -0,48 | -0,02 | -1,15 | -2,19 | |
| PF TBP | 0 | 0,14 | 0,97 | 0,98 | 0,84 | PFL2345c |
| PF TBP | 0 | -0,07 | 0,85 | 0,98 | 0,28 | |
| PF TOPO 1 | 0 | -0,5 | -0,82 | 0,79 | 1,04 | PFE0520c |
| PF TOPO 1 | 0 | 0,5 | 1,44 | 1,94 | 1,1 | |
| PF TOPO 2 | 0 | 0,13 | -0,22 | -0,72 | -0,01 | PF14_0316 |
| PF TOPO 2 | 0 | 0,35 | -0,63 | -0,71 | -1,04 | |
| PFC0060C | 0 | -0,55 | -1,68 | -0,95 | 0,02 | PFC0060C |
| PFC0060C | 0 | -1,8 | -2,11 | -3,08 | -1,84 | |
| PFC0195W | 0 | 0,01 | 0,67 | -0,13 | -0,62 | PFC0195w |
| PFC0195W | 0 | -0,2 | 0,38 | 0,35 | -0,84 | |
| PFC0485W | 0 | 0,07 | -0,57 | -2,31 | -0,82 | PFC0485w |
| PFC0485W | 0 | -0,77 | 0,09 | 1,31 | 0,16 | |
| PFC0510W | 0 | 1,67 | 1,86 | 3,41 | 2,05 | PFC0510w |
| PFC0510W | 0 | 0,64 | 1,36 | 1,7 | 0,91 | |
| PFC0525C | 0 | -0,17 | -0,62 | 0,68 | -0,51 | PFC0525c |

| | | | | | | |
|------------------|---|-------|-------|-------|-------|------------|
| PFC0525C | 0 | -0,87 | -0,6 | 0,91 | -0,35 | |
| PFC0635C | 0 | 0,1 | -0,68 | -0,53 | 0,39 | PFC0635c |
| PFC0635C | 0 | 0,09 | -0,58 | -0,93 | 0,15 | |
| PFC0740C | 0 | -0,78 | -1,12 | -1,24 | -0,59 | PFC0740c |
| PFC0740C | 0 | -0,48 | -0,54 | -0,83 | -0,04 | |
| PFC0755C | 0 | -0,6 | 0,01 | 1,34 | 1,22 | PFC0755C |
| PFC0755C | 0 | 0,35 | 0,97 | 1,57 | 0,99 | |
| PFC0805W | 0 | 1,03 | 0,46 | 0,21 | 0,29 | PFC0805w |
| PFC0805W | 0 | -0,47 | 0,89 | 0,49 | 0,79 | |
| PFC0825C | 0 | 0,62 | 0,91 | 0,29 | 0,43 | PFC0825C |
| PFC0825C | 0 | 0,16 | 0,61 | 0,07 | -0,25 | |
| PFC0915W | 0 | -0,2 | 0,88 | 0,75 | -0,84 | PFC0915w |
| PFC0915W | 0 | -1,11 | 1,21 | 1,74 | 1,6 | |
| PFC0920W | 0 | -0,41 | 0,92 | 0,17 | -0,51 | PFC0920w |
| PFC0920W | 0 | 0,76 | 0,99 | -0,36 | -0,11 | |
| PFCYC-2 | 0 | -0,56 | -0,88 | 0,23 | 0,1 | PFL1330c |
| PFCYC-2 | 0 | -2,57 | -1,21 | -0,54 | 0,32 | |
| PFCYC-3 (TRYP-2) | 0 | -0,01 | -1,22 | -3,53 | -1,33 | PFE0920c |
| PFCYC-3 (TRYP-2) | 0 | 0,3 | -0,21 | -2 | -1,48 | |
| PFMYB2 | 0 | -0,52 | -1,21 | -0,73 | -0,53 | PF10_0327 |
| PFMYB2 | 0 | -0,14 | -0,17 | -0,82 | -0,63 | |
| PHd2 | 0 | -1,01 | -1,77 | -1,87 | -0,27 | MAL13P1.76 |
| PHd2 | 0 | -1,33 | -2,3 | -2,32 | -0,07 | |
| phdB | 0 | 0,79 | 0,51 | -0,85 | -0,67 | PFL1905w |
| phdB | 0 | 0,52 | 0,38 | -0,45 | -0,31 | |
| PP1-LIKE | 0 | -0,53 | 0,37 | -0,19 | -0,32 | PF14_0224 |
| PP1-LIKE | 0 | 0,01 | 0,19 | -0,5 | -0,46 | |
| RNA HELICASE | 0 | 1,14 | -1,38 | -1,58 | -1,41 | PFB0860c |
| RNA HELICASE | 0 | 1,94 | -0,86 | -2,64 | -2,14 | |
| RNA POL1 | 0 | 0,37 | 0,34 | -1,88 | -1,35 | PFC0155c |
| RNA POL1 | 0 | 0,3 | -0,94 | -2,08 | -2,23 | |
| RNA POLBP | 0 | -0,17 | -2,13 | -3,3 | -0,01 | PF13_0123 |
| RNA POLBP | 0 | -0,41 | -1,75 | -1,77 | -1,21 | |
| RNA POLIII | 0 | -0,08 | -0,46 | -0,45 | -0,35 | PF13_0150 |
| RNA POLIII | 0 | 0,11 | -0,24 | 0,21 | 0,27 | |
| RPB4-LIKE | 0 | 0,47 | -0,26 | -1 | 0,29 | PFB0245c |

| | | | | | | |
|-----------|---|-------|-------|-------|-------|-------------|
| RPB4-LIKE | 0 | 0,67 | -0,32 | -0,91 | 0,24 | |
| RSUA | 0 | 0,13 | 1,14 | 0,51 | -0,79 | PFB0890c |
| RSUA | 0 | -0,4 | 0,5 | 0,07 | -0,32 | |
| SAR1 | 0 | -0,71 | 0,42 | 0,52 | 0,27 | PFD0810w |
| SAR1 | 0 | -0,06 | 0,13 | 1,04 | 0,35 | |
| SPOU | 0 | 0,06 | 0,96 | -0,29 | -0,23 | PFB0855c |
| SPOU | 0 | 0,09 | 0,81 | -0,06 | -0,24 | |
| TFIIS | 0 | 0,23 | -0,23 | -1,53 | -1,55 | PFB0290c |
| TFIIS | 0 | 0,17 | -0,58 | -0,87 | -0,35 | |
| YOU2 | 0 | 1,32 | -0,08 | 0,04 | 0,22 | PFB0620w |
| YOU2 | 0 | 1,23 | 0,19 | -0,33 | -0,22 | |
| ZnSnf2 | 0 | -0,9 | 0,68 | -2,12 | -0,5 | MAL13P1.216 |
| ZnSnf2 | 0 | -0,05 | 0,8 | -1,97 | -0,61 | |

Primer used for qPCR experiments

| Primer sequence | Gene Name |
|--------------------------------|------------------|
| TTTTCTCCGCCGAAAGG | 18s-1786F |
| TCATGCTTACTAGGCATTCTCG | 18s-1890R |
| CCTTAGGTGCCTTCCATTCCA | GBP 130-125F |
| TACAGCCTTTTCGTATTTGTCTCCA | GBP 130-228R |
| CCGTGCATCAAGAGCTGGA | H2A-126F |
| GCTGCGGCGTAAACTGCT | H2A-239R |
| CAAGATACCGAAATTGTTAGACCTCAA | HMG2-5659F |
| TGACAAGATCCCAAATGGTGC | HMG2-5821R |
| TGCTACACCAGGCGCAAA | kahrp-1512F |
| CTTCAGCAGCACATTGTCCTTT | kahrp-1612R |
| CGAATGTACTTTCTTTGATTAATGGTAAAA | krox-3425F |
| CATCACACCGTTATGTAGTTTATCTCTTT | krox-3579R |
| ATAGACCATCGGCTGCTGATGT | MAL6P1.146-8936F |
| TTCCCTGAGGTGGTTTGTTC | MAL6P1.146-9058R |
| TTACTCCTTGATGACCCTTCAGG | MYB3-6186F |
| CATATCGTCTTTTGGTGGACTTTTTA | MYB3-6339R |
| TGTGCCCTGCTTCTTTCTAA | PF07_0073-1058R |
| ACCTCAGAACAACCATTATGTGCTT | PF07_0073-946F |
| CATGAAAGTGGAGCAATCAATGTAA | PF11_0098-730F |
| GCAATTTTAAGTGCTTTTTCATTCC | PF11_0098-830R |
| CATATCCGATTTTGGCTTATCTACAA | PF13_0258-5013F |
| CTTACAGCAAATGACCATACATCTGA | PF13_0258-5162R |
| GATGCTGAAGTAGCTCTCAAACCA | PF13_0328-535F |
| TTGATGGCAAAGGATTGTTTAATTT | PF13_0328-635R |
| TTGGACAGACGTTACGAGAAATATCT | PF14_0383-3431F |
| GACGTCGAGGAAGAACATATTTTCA | PF14_0383-3585R |
| CCATATTTACGTACCCCGGT | PFB0865w-192R |
| GTGTTAAAGACAACGCACAGGTTCT | PFB0865w-74F |
| GGTAAACAACAGTGGACAGCCA | PFC0485w -1759F |
| TCGGCTCTTCAAACCTGGGA | PFC0485w -1810R |
| TGTGCTAGAGAAATGGCTGGAA | PFC0805w-2638F |
| TTTGACCCACTAGCCACCATACT | PFC0805w-2723R |
| TTGAGTTTCATGCCTGTAGCGA | pfg27-122F |
| TGATTCTTGCGAACATATCTCCTG | pfg27-242R |
| AATGTGCCCATGAACAAATGAA | PFI0930c-137F |
| TGTATTAGCCAAAAACCTGGAATT | PFI0930c-249R |
| TACTCAAATAATGTGTTTCATGTCATGG | PHd2-858F |
| CGTGCATCGATATTAATGTATACCA | PHd2-988R |
| CCAACATTTTTTATGAACACAAGCAG | PhDB-192R |
| GAATAATTCAGATGAGCAGGCAGA | PhDB-72F |

4. Données et matériels supplémentaires de l'article 3.

Supplementary data 1: microarray data used for SAM comparison between untreated samples and *pfmyb1* dsRNA treated samples.

| Gene Name | Accession Number | control | control | myb1 ds RNA | myb1 ds RNA | control | control | myb1 ds RNA | myb1 ds RNA | |
|---------------|------------------|---------|---------|-------------|-------------|---------|---------|-------------|-------------|------|
| PF CRK1 | PFD0865c | | 211 | 272 | 226 | 235 | 301 | 394 | 633 | 658 |
| PF PHOS | PFI1245c | | 132 | 197 | 155 | 163 | 222 | 368 | 165 | 194 |
| DNA HELICASE | PFB0730w | | 130 | 152 | 173 | 158 | 230 | 224 | 182 | 267 |
| PFBO715 | PFB0715w | | 168 | 158 | 179 | 218 | 419 | 424 | 428 | 505 |
| PF LSA-3 | PFB0915w | | 341 | 328 | 515 | 485 | 1298 | 1249 | 1631 | 1735 |
| PF CRK2 | PFD0740w | | 96 | 106 | 143 | 143 | 205 | 187 | 130 | 116 |
| PFEMP2 | PFE0040c | | 133 | 159 | 192 | 171 | 253 | 330 | 229 | 198 |
| PF MAP2 | PF11_0147 | | 183 | 210 | 235 | 256 | 191 | 215 | 252 | 227 |
| ADP-RIBO FAC1 | PF10_0203 | 6478 | 5641 | 7771 | 7293 | 11031 | 11551 | 9699 | 10010 | |
| CHR12 14 | PFL2280w | | 296 | 420 | 300 | 260 | 540 | 602 | 488 | 404 |
| SPOU | PFB0855c | | 203 | 232 | 310 | 288 | 529 | 506 | 440 | 410 |
| CHR9 9716 | PF14_0392 | | 243 | 309 | 223 | 456 | 399 | 460 | 256 | 270 |
| PFC0525C | PFC0525C | | 352 | 463 | 436 | 417 | 957 | 855 | 1167 | 1137 |
| HISTONE H4 | PF11_0061 | | 296 | 289 | 303 | 337 | 440 | 494 | 669 | 309 |
| SAR1 | PFD0810w | | 620 | 540 | 812 | 944 | 1403 | 1616 | 1490 | 1391 |
| PF MYB3 | PF10_0143 | | 731 | 857 | 599 | 735 | 1636 | 1739 | 1143 | 1318 |
| CHR10 3810 | PF11_0227 | | 248 | 194 | 233 | 284 | 394 | 415 | 422 | 427 |
| RNA POL1 | PFC0155c | | 370 | 268 | 292 | 192 | 209 | 214 | 228 | 222 |
| CHR13 40983 | MAL13P1.78 | 18621 | 19850 | 26760 | 26441 | 9896 | 10093 | 6573 | 6275 | |
| K9:1 | MAL13P1.76 | | 723 | 826 | 828 | 956 | 2690 | 2913 | 2526 | 2297 |
| PFCYC-3 | PFE0920c | | 94 | 129 | 128 | 111 | 204 | 215 | 146 | 192 |
| PF ERB | PF11_0098 | | 1619 | 1744 | 1303 | 1469 | 2403 | 2538 | 3672 | 3742 |
| PF POLA | PFD0590c | | 518 | 565 | 530 | 583 | 1278 | 1101 | 1288 | 1535 |
| PF ADENYL | PF11_0395 | | 281 | 310 | 513 | 343 | 738 | 325 | 605 | 617 |
| PF CSP | PFC0210c | | 207 | 144 | 151 | 163 | 186 | 204 | 173 | 159 |
| PF MBP | PFB0725c | | 145 | 145 | 138 | 115 | 192 | 151 | 147 | 164 |
| CHR12 06 | PFL1285c | | 373 | 344 | 214 | 125 | 281 | 386 | 186 | 195 |
| CHRBLOB 36958 | MAL6P1.191 | | 191 | 199 | 153 | 186 | 366 | 361 | 286 | 291 |
| PF PK1 | PF08_0044 | | 234 | 148 | 222 | 464 | 114 | 186 | 166 | 123 |
| PF CDPK3 | PFC0420w | | 285 | 285 | 366 | 384 | 386 | 350 | 394 | 459 |
| PF PHOS | PFI1245c | | 152 | 173 | 240 | 188 | 302 | 278 | 231 | 256 |

| | | | | | | | | | |
|-------------|-------------|-------|-------|-------|-------|-------|-------|-------|-------|
| PF RAD2 | PFB0265c | 221 | 182 | 176 | 179 | 348 | 397 | 370 | 327 |
| PFC0635C | PFC0635C | 129 | 143 | 114 | 112 | 212 | 212 | 172 | 184 |
| PF HMG2 | PFL0290w | 1549 | 1360 | 1077 | 1193 | 2408 | 2534 | 2309 | 2214 |
| CHR11 N3190 | PFL0080c | 213 | 193 | 210 | 255 | 468 | 488 | 417 | 500 |
| CHR14 128 A | PF14_0431 | 129 | 144 | 134 | 96 | 220 | 174 | 159 | 169 |
| P34CDC2 | MAL13P1.279 | 578 | 519 | 2008 | 1669 | 447 | 447 | 1561 | 1179 |
| PF PK4 | MAL6P1.146 | 61618 | 54741 | 60605 | 66758 | 52202 | 49314 | 31241 | 29459 |
| DNA PRIM | PF14_0366 | 819 | 815 | 554 | 579 | 2237 | 2854 | 729 | 858 |
| RPB4-LIKE | PFB0245c | 139 | 181 | 171 | 195 | 191 | 227 | 283 | 315 |
| PFC0740C | PFC0740C | 383 | 479 | 621 | 544 | 1913 | 1870 | 1864 | 1951 |
| PF STARP | PF07_0006 | 132 | 149 | 137 | 128 | 287 | 315 | 287 | 280 |
| PFCYC-2 | PFL1330c | 384 | 451 | 385 | 435 | 1280 | 1473 | 1377 | 1377 |
| TRAP | PF13_0201 | 188 | 199 | 186 | 192 | 230 | 230 | 171 | 227 |
| PF PCNA | PF13_0328 | 959 | 964 | 1185 | 1040 | 2331 | 2077 | 3265 | 3535 |
| 3-PHOSPHO | PF11105w | 1743 | 1540 | 992 | 943 | 1211 | 972 | 693 | 585 |
| PF CPK | PFB0815w | 2411 | 2425 | 2868 | 2343 | 6890 | 7079 | 7476 | 7993 |
| ALIEN1 | | 943 | 919 | 942 | 921 | 3036 | 3891 | 2370 | 2652 |
| CHR11 S56 | PF11_0488 | 245 | 194 | 178 | 189 | 453 | 474 | 287 | 289 |
| RSUA | PF14_0630 | 234 | 240 | 229 | 249 | 693 | 620 | 530 | 592 |
| IX78E09.P1C | PF08_0044 | 345 | 294 | 300 | 310 | 707 | 673 | 544 | 535 |
| M10:1 | PF13_0289 | 506 | 492 | 672 | 607 | 527 | 494 | 446 | 387 |
| CALMODULIN | PF14_0323 | 299 | 255 | 285 | 291 | 1211 | 1399 | 1423 | 1295 |
| PFC0485W | PFC0485W | 234 | 223 | 222 | 285 | 425 | 511 | 430 | 552 |
| PF S40 | PF11_0098 | 2003 | 1942 | 2569 | 2418 | 4994 | 4650 | 5994 | 5927 |
| PF HMG1 | PFL0145c | 1201 | 1154 | 1179 | 1340 | 2954 | 2736 | 1795 | 2233 |
| CHR11 N1106 | PFL2250c | 307 | 247 | 234 | 349 | 511 | 433 | 400 | 349 |
| PFC0825C | PFC0825C | 304 | 296 | 221 | 282 | 360 | 347 | 449 | 364 |
| CHR13 41763 | MAL13P1.196 | 512 | 544 | 312 | 458 | 999 | 922 | 988 | 922 |
| K10:1 | PFL1905w | 202 | 240 | 134 | 162 | 256 | 243 | 199 | 254 |
| PFAGBPH | PF14_0010 | 202 | 205 | 139 | 157 | 227 | 153 | 172 | 191 |
| PF CK1 | PF11_0377 | 1780 | 1554 | 1865 | 1855 | 4834 | 3946 | 5985 | 5253 |
| PFC0340W | PFC0340W | 339 | 474 | 397 | 396 | 816 | 822 | 760 | 731 |
| PF RAN-1 | PF11_0183 | 1077 | 1094 | 1402 | 1366 | 2713 | 2762 | 3349 | 3883 |
| PF SNRNP | PFB0865w | 1133 | 915 | 1344 | 1082 | 1475 | 1581 | 1314 | 1560 |
| PF LSA-1 | PF10_0356 | 506 | 427 | 480 | 462 | 291 | 291 | 193 | 291 |
| PFC0805W | PFC0805W | 600 | 512 | 609 | 506 | 601 | 593 | 669 | 653 |

| | | | | | | | | | |
|-------------|------------------|------|------|------|------|------|------|------|------|
| CHR13 21780 | PF13_0258 | 453 | 456 | 352 | 353 | 251 | 261 | 248 | 337 |
| PFEMP1 ATS | PF08_0107 | 382 | 326 | 407 | 435 | 1337 | 1383 | 1211 | 1352 |
| | 68156 MAL6P1.271 | 898 | 1106 | 776 | 792 | 1424 | 1257 | 1537 | 1285 |
| PF MAP1 | PF14_0294 | 259 | 210 | 276 | 229 | 258 | 204 | 273 | 226 |
| PF SNF1 | PF14_0516 | 272 | 213 | 214 | 163 | 196 | 156 | 291 | 440 |
| PF BO150C | PFBO150C | 338 | 248 | 312 | 243 | 492 | 386 | 493 | 457 |
| PF B7 | PFI0930c | 720 | 593 | 676 | 602 | 1107 | 1042 | 927 | 766 |
| EIF-4A | PFC0445w | 720 | 719 | 822 | 769 | 1998 | 1786 | 1974 | 1472 |
| HOX 1 | PFD0340c | 396 | 288 | 401 | 324 | 921 | 1337 | 1250 | 1053 |
| CHR11 S127 | PF11_0464 | 830 | 748 | 1181 | 973 | 1513 | 1536 | 1635 | 948 |
| CHR14 128 B | PF14_0423 | 451 | 373 | 446 | 402 | 440 | 363 | 367 | 333 |
| PF MT C | PFD0370w | 80 | 95 | 146 | 81 | 130 | 153 | 115 | 129 |
| PF PK6 | PF13_0206 | 101 | 141 | 141 | 143 | 150 | 237 | 177 | 221 |
| HISTONE H2B | PF11_0062 | 418 | 428 | 344 | 326 | 222 | 388 | 499 | 421 |
| PFC0955C | PFC0955C | 304 | 355 | 327 | 346 | 389 | 691 | 868 | 916 |
| GBP130 | PF10_0159 | 435 | 387 | 284 | 303 | 309 | 492 | 542 | 601 |
| PFC0060C | PFC0060C | 151 | 171 | 241 | 223 | 250 | 235 | 261 | 203 |
| PF ERC | PF11_0098 | 1911 | 2037 | 1566 | 1592 | 5408 | 5640 | 6280 | 6221 |
| PF TOPO 2 | PF14_0316 | 524 | 509 | 374 | 363 | 970 | 960 | 1206 | 1195 |
| PF C0155C | PF C0155C | 255 | 271 | 195 | 185 | 369 | 414 | 370 | 308 |
| CHR13 32969 | PF13_0211 | 298 | 270 | 163 | 203 | 491 | 494 | 552 | 645 |
| C2143 | PF07_0073 | 349 | 376 | 201 | 259 | 489 | 430 | 672 | 726 |
| BL57G11.Q1T | PF07_0072 | 365 | 487 | 249 | 260 | 656 | 602 | 726 | 734 |
| CBPR | PFL0635c | 76 | 82 | 61 | 44 | 74 | 79 | 116 | 127 |
| PF BO520W | PFBO520W | 133 | 154 | 130 | 130 | 243 | 366 | 220 | 247 |
| PF RNR2 | PF14_0053 | 700 | 675 | 622 | 705 | 1714 | 1765 | 2086 | 2106 |
| PF NEK-1 | PFL1370w | 204 | 223 | 150 | 137 | 388 | 342 | 274 | 273 |
| HOX 2 | PFD1160w | 197 | 198 | 215 | 217 | 289 | 286 | 198 | 222 |
| CHR11 S189 | PF11_0096 | 260 | 286 | 207 | 199 | 814 | 816 | 905 | 884 |
| PF PKA | PFI1685w | 1114 | 1294 | 1210 | 1184 | 5260 | 5188 | 5073 | 4254 |
| PFC0920W | PFC0920W | 3065 | 2764 | 3917 | 3885 | 2738 | 2647 | 2537 | 2265 |
| PFC0420W | PFC0420W | 9062 | 9305 | 8885 | 8938 | 3036 | 3080 | 3153 | 2831 |
| PFC0510W | PFC0510W | 292 | 315 | 204 | 222 | 522 | 556 | 583 | 610 |
| RNA POLB | PFL2185w | 321 | 303 | 269 | 256 | 205 | 178 | 102 | 101 |
| PF DNA GEN | | 1408 | 1247 | 1571 | 1327 | 891 | 878 | 909 | 859 |
| KHARP | PFB0100c | 1100 | 1154 | 1140 | 961 | 1819 | 1721 | 1986 | 1828 |

| | | | | | | | | | |
|----------------|-------------|-------|-------|-------|-------|------|------|-------|------|
| CYCLIN ANIA | PF13_0022 | 97 | 166 | 92 | 79 | 101 | 117 | 98 | 175 |
| PF FEST | MAL7P1.91 | 190 | 159 | 156 | 120 | 348 | 327 | 274 | 303 |
| PF PK2 | PFL1885c | 814 | 825 | 651 | 670 | 814 | 752 | 560 | 565 |
| PF BO665W | PFBO665W | 382 | 314 | 337 | 282 | 474 | 507 | 424 | 421 |
| PF TBP | PFL2345c | 873 | 808 | 480 | 454 | 1499 | 1406 | 873 | 844 |
| EF-1B | PFI0645w | 301 | 319 | 282 | 256 | 647 | 588 | 693 | 661 |
| NF-LIK1 | PFL1030w | 132 | 118 | 79 | 86 | 140 | 142 | 136 | 149 |
| CHR11 S436 | PF11_0242 | 149 | 143 | 98 | 66 | 251 | 243 | 208 | 210 |
| CHR5 21208 | PFE0420c | 5938 | 5239 | 4889 | 4927 | 6706 | 6939 | 5967 | 5761 |
| PP1-LIKE | PF14_0224 | 6010 | 6186 | 3217 | 3240 | 6825 | 6621 | 4845 | 3934 |
| PFC0755C | PFC0755C | 823 | 862 | 719 | 648 | 2064 | 2328 | 2702 | 2450 |
| HISTONE H3 | MAL6P1.248 | 3307 | 2797 | 1102 | 1095 | 2895 | 3000 | 1849 | 1733 |
| PFC0915W | PFC0915W | 373 | 381 | 333 | 458 | 1085 | 993 | 1013 | 950 |
| PFMYB2 | PF10_0327 | 199 | 179 | 151 | 209 | 410 | 506 | 354 | 411 |
| PFC0385C | PFC0385C | 307 | 335 | 138 | 131 | 556 | 584 | 367 | 383 |
| PF CDPK2 | MAL6P1.108 | 21620 | 20658 | 23581 | 25034 | 9281 | 8732 | 6233 | 4985 |
| PF DNA POLD | PF10_0165 | 236 | 236 | 271 | 306 | 245 | 220 | 206 | 160 |
| RNA POLBP | PFL2185w | 142 | 140 | 158 | 91 | 210 | 191 | 151 | 134 |
| CHR13 40872 | MAL13P1.278 | 395 | 439 | 472 | 326 | 1281 | 1058 | 1103 | 989 |
| PF HSP70 | PF08_0054 | 4112 | 4166 | 3189 | 2748 | 8307 | 8588 | 10191 | 8057 |
| BU60D08.P1C | PF08_0044 | 161 | 164 | 177 | 171 | 328 | 373 | 240 | 270 |
| FSTH | PFL1925w | 177 | 178 | 304 | 305 | 365 | 776 | 298 | 616 |
| PF BO605W | PFBO605W | 660 | 561 | 704 | 763 | 2108 | 2057 | 1995 | 2164 |
| PF RNR1 | PF14_0352 | 2000 | 1857 | 2509 | 1913 | 4471 | 4529 | 6065 | 6140 |
| PF CYC-1 | PF14_0605 | 228 | 237 | 269 | 247 | 322 | 404 | 356 | 393 |
| KROX | PFL0465c | 346 | 342 | 324 | 440 | 593 | 612 | 515 | 607 |
| CHR11 S245 | PF11_0239 | 139 | 140 | 72 | 121 | 184 | 181 | 183 | 210 |
| RNA HELICASE 1 | PFE1390w | 163 | 149 | 105 | 67 | 224 | 191 | 210 | 168 |
| CHR14 185 | PF14_0476 | 140 | 132 | 70 | 79 | 189 | 120 | 139 | 132 |
| PFC0195W | PFC0195W | 347 | 304 | 316 | 308 | 535 | 542 | 441 | 530 |
| HISTONE H2A | MAL6P1.249 | 675 | 604 | 325 | 336 | 611 | 667 | 301 | 329 |
| PF CDC ATP | PF07_0047 | 290 | 315 | 387 | 382 | 1111 | 1263 | 1271 | 1329 |
| YOU2 | PFB0620w | 157 | 177 | 186 | 215 | 205 | 217 | 270 | 264 |
| RNA POLIII | PF13_0150 | 173 | 185 | 122 | 136 | 407 | 425 | 311 | 354 |
| PF DNA GEN | | 543 | 609 | 595 | 509 | 224 | 353 | 266 | 254 |
| PF CDK7 | PF10_0141 | 146 | 172 | 140 | 172 | 156 | 143 | 183 | 179 |

| | | | | | | | | | |
|-------------|-----------|------|------|------|------|------|------|------|------|
| PF TOPO 1 | PFE0520c | 233 | 238 | 311 | 369 | 468 | 358 | 444 | 434 |
| PF BO815W | PFBO815w | 1870 | 1823 | 2045 | 2086 | 6693 | 6034 | 7861 | 6918 |
| PF ATBP | PFE0305w | 198 | 206 | 166 | 243 | 214 | 209 | 380 | 442 |
| PF ACTII | PF14_0124 | 154 | 157 | 159 | 152 | 402 | 461 | 502 | 555 |
| KRU | PF14_0383 | 532 | 462 | 678 | 780 | 1921 | 1754 | 1676 | 1731 |
| CHR11 S51 | PF11_0220 | 236 | 235 | 163 | 188 | 594 | 532 | 497 | 479 |
| CALCINEURIN | PF14_0492 | 520 | 481 | 644 | 661 | 1040 | 935 | 869 | 893 |

Supplementary data 2: microarray data used for SAM comparison between untreated and unrelated dsRNA treated samples.

| Gene Name | Accession Number | control | control | unrelated dsRNA | unrelated dsRNA |
|---------------|------------------|---------|---------|-----------------|-----------------|
| PF CRK1 | PFD0865c | 1031 | 975 | 1024 | 988 |
| PF PHOS | PF11245c | 225 | 242 | 202 | 219 |
| DNA HELICASE | PFB0730w | 360 | 421 | 439 | 312 |
| PFBO715 | PFB0715w | 1271 | 1284 | 1256 | 1131 |
| PF LSA-3 | PFB0915w | 3046 | 2935 | 2928 | 3021 |
| PF CRK2 | PFD0740w | 472 | 466 | 483 | 342 |
| PFEMP2 | PFE0040c | 1104 | 1132 | 1159 | 939 |
| PF MAP2 | PF11_0147 | 243 | 239 | 235 | 255 |
| ADP-RIBO FAC1 | PF10_0203 | 6710 | 6670 | 6779 | 6795 |
| CHR12 14 | PFL2280w | 210 | 225 | 233 | 211 |
| SPOU | PFB0855c | 956 | 971 | 739 | 954 |
| CHR9 9716 | PF14_0392 | 526 | 520 | 719 | 562 |
| PFC0525C | PFC0525C | 1019 | 1000 | 1048 | 998 |
| HISTONE H4 | PF11_0061 | 295 | 295 | 271 | 225 |
| SAR1 | PFD0810w | 861 | 901 | 828 | 914 |
| PF MYB3 | PF10_0143 | 3116 | 3464 | 3238 | 3533 |
| CHR10 3810 | PF11_0227 | 362 | 332 | 345 | 437 |
| RNA POL1 | PFC0155c | 339 | 351 | 335 | 337 |
| CHR13 40983 | MAL13P1.78 | 8303 | 8364 | 8571 | 8015 |
| K9:1 | MAL13P1.76 | 1471 | 1437 | 1442 | 1333 |
| PFCYC-3 | PFE0920c | 216 | 214 | 266 | 165 |
| PF ERB | PF11_0098 | 3749 | 4363 | 4033 | 3536 |
| PF POLA | PFD0590c | 1054 | 1033 | 997 | 942 |
| PF ADENYL | PF11_0395 | 295 | 269 | 302 | 297 |
| PF CSP | PFC0210c | 278 | 193 | 280 | 257 |
| PF MBP | PFB0725c | 228 | 227 | 239 | 242 |
| CHR12 06 | PFL1285c | 413 | 297 | 437 | 297 |
| CHRBLOB 36958 | MAL6P1.191 | 308 | 270 | 276 | 289 |
| PF PK1 | PF08_0044 | 226 | 190 | 208 | 215 |
| PF CDPK3 | PFC0420w | 315 | 327 | 458 | 320 |
| PF PHOS | PF11245c | 309 | 322 | 305 | 284 |
| PF RAD2 | PFB0265c | 399 | 380 | 387 | 397 |

| | | | | | |
|-----------------|-------------|-------|-------|-------|-------|
| PFC0635C | PFC0635C | 193 | 233 | 239 | 252 |
| PF HMG2 | PFL0290w | 2334 | 2672 | 2544 | 2617 |
| CHR11 N3190 | PFL0080c | 273 | 308 | 267 | 241 |
| CHR14 128 A | PF14_0431 | 243 | 207 | 211 | 214 |
| P34CDC2 | MAL13P1.279 | 518 | 428 | 481 | 524 |
| PF PK4 | MAL6P1.146 | 16904 | 17205 | 16972 | 17030 |
| DNA PRIM | PF14_0366 | 1448 | 1459 | 1182 | 1405 |
| RPB4-LIKE | PFB0245c | 257 | 261 | 268 | 269 |
| PFC0740C | PFC0740C | 1648 | 1689 | 1687 | 1643 |
| PF STARP | PF07_0006 | 238 | 245 | 216 | 225 |
| PFCYC-2 | PFL1330c | 264 | 318 | 337 | 339 |
| TRAP | PF13_0201 | 655 | 659 | 689 | 641 |
| PF PCNA | PF13_0328 | 5770 | 5615 | 6143 | 5885 |
| 3-PHOSPHO | PF1105w | 2118 | 1923 | 2253 | 1773 |
| PF CPK | PFB0815w | 339 | 318 | 308 | 286 |
| ALIEN1 | | 1089 | 1221 | 1159 | 1217 |
| CHR11 S56 | PF11_0488 | 374 | 292 | 380 | 246 |
| RSUA | PF14_0630 | 590 | 498 | 486 | 570 |
| CHR9 IX78E09.P1 | PF08_0044 | 874 | 939 | 839 | 940 |
| M10:1 | PF13_0289 | 960 | 827 | 985 | 945 |
| CALMODULIN | PF14_0323 | 361 | 293 | 311 | 258 |
| PFC0485W | PFC0485W | 683 | 702 | 638 | 766 |
| PF S40 | PF11_0098 | 5521 | 4964 | 5474 | 5152 |
| PF HMG1 | PFL0145c | 508 | 525 | 529 | 565 |
| CHR11 N1106 | PFL2250c | 250 | 244 | 253 | 245 |
| PFC0825C | PFC0825C | 483 | 505 | 491 | 558 |
| CHR13 41763 | MAL13P1.196 | 876 | 861 | 843 | 861 |
| K10:1 | PFL1905w | 576 | 640 | 647 | 521 |
| PFAGBPH | PF14_0010 | 212 | 206 | 231 | 225 |
| PF CK1 | PF11_0377 | 2751 | 2424 | 2634 | 2178 |
| PFC0340W | PFC0340W | 635 | 709 | 726 | 677 |
| PF RAN-1 | PF11_0183 | 2566 | 2511 | 2476 | 2501 |
| PF SNRNP | PFB0865w | 1498 | 1536 | 1436 | 1570 |
| PF LSA-1 | PF10_0356 | 236 | 302 | 292 | 375 |
| PFC0805W | PFC0805W | 1931 | 1839 | 1684 | 1808 |
| CHR13 21780 | PF13_0258 | 255 | 229 | 249 | 258 |

| | | | | | |
|---------------|------------|-------|-------|-------|-------|
| PFEMP1 ATS | PF08_0107 | 1478 | 1447 | 1427 | 1520 |
| CHRBLOB 68156 | MAL6P1.271 | 1081 | 1046 | 1097 | 1083 |
| PF MAP1 | PF14_0294 | 364 | 388 | 429 | 340 |
| PF SNF1 | PF14_0516 | 191 | 221 | 184 | 141 |
| PF BO150C | PFBO150C | 239 | 232 | 245 | 281 |
| PF B7 | PFI0930c | 2342 | 2142 | 2491 | 2269 |
| EIF-4A | PFC0445w | 6990 | 6091 | 6050 | 6069 |
| HOX 1 | PF0340c | 541 | 541 | 586 | 563 |
| CHR11 S127 | PF11_0464 | 256 | 241 | 216 | 242 |
| CHR14 128 B | PF14_0423 | 853 | 819 | 799 | 835 |
| PF MT C | PF0370w | 2932 | 2605 | 2813 | 2619 |
| PF PK6 | PF13_0206 | 803 | 736 | 886 | 846 |
| HISTONE H2B | PF11_0062 | 1694 | 1651 | 2196 | 1914 |
| PFC0955C | PFC0955C | 1552 | 1626 | 1633 | 1636 |
| GBP130 | PF10_0159 | 208 | 165 | 169 | 232 |
| PFC0060C | PFC0060C | 367 | 490 | 284 | 328 |
| PF ERC | PF11_0098 | 8406 | 8205 | 8443 | 8707 |
| PF TOPO 2 | PF14_0316 | 264 | 255 | 201 | 232 |
| PF C0155C | PF C0155C | 247 | 244 | 280 | 271 |
| CHR13 32969 | PF13_0211 | 711 | 703 | 737 | 795 |
| C2143 | PF07_0073 | 10391 | 11078 | 10555 | 11805 |
| BL57G11.Q1T | PF07_0072 | 474 | 534 | 474 | 473 |
| CBPR | PFL0635c | 1407 | 1568 | 1432 | 1718 |
| PF BO520W | PFBO520W | 3363 | 3467 | 3246 | 3415 |
| PF RNR2 | PF14_0053 | 5038 | 4888 | 5025 | 5163 |
| PF NEK-1 | PFL1370w | 461 | 647 | 578 | 618 |
| HOX 2 | PF01160w | 844 | 810 | 864 | 832 |
| CHR11 S189 | PF11_0096 | 255 | 244 | 284 | 238 |
| PF PKA | PFI1685w | 1571 | 1467 | 1432 | 1426 |
| PFC0920W | PFC0920W | 717 | 684 | 732 | 689 |
| PFC0420W | PFC0420W | 2274 | 2364 | 2409 | 2310 |
| PFC0510W | PFC0510W | 881 | 770 | 870 | 878 |
| RNA POLB | PFL2185w | 425 | 488 | 590 | 546 |
| PF DNA GEN | | 785 | 781 | 770 | 830 |
| KHARP | PFB0100c | 149 | 183 | 159 | 160 |
| CYCLIN ANIA | PF13_0022 | 153 | 187 | 128 | 132 |

| | | | | | |
|----------------|-------------|------|------|------|------|
| PF FEST | MAL7P1.91 | 265 | 272 | 243 | 236 |
| PF PK2 | PFL1885c | 467 | 466 | 422 | 482 |
| PF BO665W | PFBO665W | 225 | 293 | 213 | 207 |
| PF TBP | PFL2345c | 258 | 218 | 193 | 256 |
| EF-1B | PFI0645w | 842 | 924 | 1064 | 648 |
| NF-LIK1 | PFL1030w | 2702 | 2925 | 2519 | 2981 |
| CHR11 S436 | PF11_0242 | 305 | 329 | 321 | 235 |
| CHR5 21208 | PFE0420c | 127 | 177 | 214 | 95 |
| PP1-LIKE | PF14_0224 | 284 | 308 | 316 | 283 |
| PFC0755C | PFC0755C | 988 | 970 | 741 | 963 |
| HISTONE H3 | MAL6P1.248 | 589 | 636 | 543 | 643 |
| PFC0915W | PFC0915W | 776 | 775 | 839 | 807 |
| PFMYB2 | PF10_0327 | 384 | 381 | 362 | 269 |
| PFC0385C | PFC0385C | 360 | 267 | 352 | 396 |
| PF CDPK2 | MAL6P1.108 | 3702 | 3822 | 3734 | 3898 |
| PF DNA POLD | PF10_0165 | 841 | 909 | 845 | 832 |
| RNA POLBP | PFL2185w | 951 | 933 | 967 | 1037 |
| CHR13 40872 | MAL13P1.278 | 208 | 180 | 165 | 185 |
| PF HSP70 | PF08_0054 | 1210 | 1144 | 1257 | 1204 |
| BU60D08.P1C | PF08_0044 | 225 | 205 | 189 | 179 |
| FSTH | PFL1925w | 231 | 194 | 216 | 189 |
| PF BO605W | PFBO605W | 287 | 245 | 299 | 314 |
| PF RNR1 | PF14_0352 | 2399 | 2634 | 2718 | 2666 |
| PF CYC-1 | PF14_0605 | 287 | 321 | 253 | 314 |
| KROX | PFL0465c | 995 | 988 | 993 | 993 |
| CHR11 S245 | PF11_0239 | 1313 | 1229 | 1307 | 1316 |
| RNA HELICASE 1 | PFE1390w | 239 | 244 | 254 | 257 |
| CHR14 185 | PF14_0476 | 164 | 169 | 152 | 121 |
| PFC0195W | PFC0195W | 320 | 284 | 299 | 224 |
| HISTONE H2A | MAL6P1.249 | 2339 | 2238 | 2436 | 2412 |
| PF CDC ATP | PF07_0047 | 4979 | 4837 | 4934 | 4854 |
| YOU2 | PFB0620w | 288 | 322 | 299 | 315 |
| RNA POLIII | PF13_0150 | 167 | 236 | 266 | 243 |
| PFDNA GEN | | 1909 | 2184 | 2056 | 2139 |
| PF CDK7 | PF10_0141 | 550 | 450 | 487 | 551 |
| PF TOPO 1 | PFE0520c | 335 | 212 | 205 | 224 |

| | | | | | |
|-------------|-----------|------|------|------|------|
| PF BO815W | PFBO815w | 513 | 513 | 466 | 578 |
| PF ATBP | PFE0305w | 786 | 697 | 796 | 797 |
| PF ACTII | PF14_0124 | 4939 | 5033 | 5206 | 5064 |
| KRU | PF14_0383 | 307 | 235 | 310 | 300 |
| CHR11 S51 | PF11_0220 | 220 | 193 | 189 | 206 |
| CALCINEURIN | PF14_0492 | 160 | 164 | 156 | 148 |

Supplementary material 1: primers used for qPCR experiments.

| Primer sequence | Primer name |
|------------------------------|--------------------------|
| GCTCTTGACGTTGATAATAGTGGAAC | PFCDPK2-1141F |
| GACCTGAGGCGTTTGAGTCAAT | PFCDPK2-1273R |
| ACAGGAGGAAAGGCCCA | histone h3-34F |
| GTTCTGGACGATATCTATGTGGC | histone h3-137R |
| AGCAGGAAATGCAGCAAGAGAT | histone h2a -210F |
| CCTGAAGCGAAGGTGACACC | histone h2a -329R |
| TCAGGTGCTGAGGTAAAAGCTGT | Pf tbp homolog (m)-1272F |
| TCTGCGTCTTGCTTCATAACCTT | Pf tbp homolog (m)-1396R |
| TTGAAAGAAATCGACGAACTCGA | PP1-like PF14_0224-439F |
| TCTGAGGATCGGACCATAACAGA | PP1-like PF14_0224-541R |
| TCGGTGATAAAAGAAAAGTCGACC | PCNA-23F |
| TCAGCACTTGATTGTTCTTCTTCATC | PCNA-129R |
| TGAAGCGGAAAATAGGACATCC | PF14_0366 -1218F |
| TTCATTCAATCGTTTCTTTTTTCAAC | PF14_0366 -1326R |
| TTTCGGTACTGCTCATCGTGC | PGK PFI1105w-492F |
| TGGCTAGTAGTGGTCTTTGTGGGT | PGK PFI1105w-628R |
| ATGAACCTTACCATGGGAATCA | Pfpk5-722F |
| TGTTATCTTTGGTTTGGATCAAGC | Pfpk5-822R |
| TGTGCTAGAGAAATGGCTGGAA | PFC0805w-2638F |
| TTTGACCCACTAGCCACCATACT | PFC0805w-2723R |
| AATCGATTATATGCCCCCTGAAA | PFC0385c-4665F |
| CGGTAAATGGAGGAAATCCAATA | PFC0385c-4780R |
| ATCCTTGTTTTAACCACTATGGAGG | MAL8P1.154-5831F |
| TGAGACGTTTCCCCTAGAAAAATAA | MAL8P1.154-5979R |
| TATTCCACCTCCACCACCCA | PFD0465w -1644F |
| TGTTATCCAATTTGGCACTATTATTG | PFD0465w -1753R |
| TTTCAAAGATAAAATTTCCGATTCTGA | hsp 70-1758F |
| TCTTGCGATATGGCCCATG | hsp 70-1904R |
| TACTCAAATAATGTGTTTCATGTCATGG | MAL13P1.76 F |
| CGTGCATCGATATTAATGTATACCA | MAL13P1.76 R |
| CGAAGATATGTATGGCGAAGCA | MYB2-2118F |
| ACATTGGGACGTATCATCACCTTT | MYB2-2262R |
| TTACTCCTTGATGACCCTTCAGG | MYB3-6186F |
| CATATCGTCTTTTGGTGGACTTTTTA | MYB3-6339R |

Supplementary material 2: primers used for ChIP experiments.

| Primer sequence | Primer name |
|------------------------------|-------------------|
| TATGTTATTTTCTTACGTCCAG | Pfpk5 f |
| AAAAATATTCTTATATTCTT | pfpk5 r |
| AAATATTATATGTAGCCTATG | pgk f |
| GATTATAAAATAGAGGGTT | pgk r |
| TTATTTATATTTATGGCATT | PF07_0073 f |
| CTCCTTCACCTTTAATCCAA | PF07_0073 r |
| ATCCTTGTTTTAACCCTATGGAGG | MAL8P1.154 F |
| TGAGACGTTTCCCCTAGAAAAATAA | MAL8P1.154 R |
| CAAGTAAGACCTGGTAAAATCGAAAA | PF07_0037 F |
| TAACCAGATTAGATGAATAGTCCGATTC | PF07_0037-R |
| TATCCACCTCCACCACCCA | PFD0465w -F |
| TGTTATCCAATTTGGCACTATTATTTG | PFD0465w -R |
| AAACATAATACGGAACAACCT | PFL1885c f |
| GCGATGCACACAGAATAGC | PFL1885c r |
| GGGGATATATAAATGAACCTA | histones h3/h2a f |
| TACGGTATATTAACGAATGC | histones h3/h2a r |
| ATTTATTGTTTGGCATATCTC | PFL1285c f |
| ATTTTTATAAAGCGTCTTGT | PFL1285c r |
| AATTTTTATATTTTCGTCTC | PFL2345c f |
| TGATTTAGATTGGTCGTA | PFL2345c r |
| ATTCATGACGGTCGATAG | PF14_0224 f |
| ACAATTCTGTGGTGTACCAA | PF14_0224 r |
| TTTTCCAACAGCCATCTAA | PF08_0054 f |
| TTTGCTCTTATTTATAACAG | PF08_0054 r |
| AGACATATGTATATTTGAGA | PF10_0159 f |
| AGGGGTGATGTGACTA | PF10_0159 r |
| TTGGGTAACATGATAGTATGA | PF11_0096 f |
| TATATGGCATATAACCTTC | PF11_0096 r |