

# Sélection naturelle de variants génétiques conférant une résistance au paludisme dans les populations humaines

Pascal RIHET

## Résumé

Le paludisme exerce une forte pression de sélection sur le génome humain depuis le développement de l'agriculture il y a 10 000 ans. Ce concept était initialement basé sur l'observation d'une fréquence élevée des déficiences érythrocytaires dans les populations vivant en zone d'endémie. Plusieurs variants génétiques affectant la physiologie du globule rouge, comme l'hémoglobine S, l'hémoglobine C, et les allèles G6PD A- et DARC/FY\*O, sont associés à la résistance contre le paludisme et ont été sélectionnés sous la pression du paludisme. Plusieurs gènes impliqués dans la réponse immunitaire sont aussi associés à la résistance contre le paludisme, avec un effet protecteur moindre. De plus, des études conduites à l'échelle du génome ont révélé des signatures moléculaires de la sélection positive dans plusieurs gènes du système immunitaire impliqués dans la protection contre le paludisme. Les progrès, à la fois des technologies à haut débit et de l'analyse des réseaux moléculaires, devraient permettre l'identification de nombreux gènes et réseaux moléculaires impliqués dans la résistance au paludisme et ouvrir ainsi de nouvelles perspectives dans le domaine de l'évolution du génome humain.

**Mots clés :** sélection naturelle, populations humaines, paludisme, *Plasmodium*, variant génétique, résistance.

## Abstract

Malaria has exerted a strong selective pressure in the recent history of the human genome, coinciding with the beginning of agriculture about 10,000 years ago.

This concept was initially based on the observation that erythrocyte defects are highly prevalent in populations living in malaria endemic areas. Several genetic variants that affect erythrocyte physiology, such as haemoglobin S, haemoglobin C, G6PD A-, and DARC/FY\*O, have been shown to confer resistance from malaria in several populations, and to be selected under the pressure of malaria. Several immune genes have also been associated with malaria resistance or susceptibility, yet with a more subtle effect.

In addition, genome-wide studies have revealed a molecular signature of a recent positive selection in several immune genes involved in malaria resistance.

Nevertheless, this may be the tip of the iceberg, and advances in both high-throughput technologies and the analysis of molecular networks should provide new insights into the evolution of human genome.

**Keywords:** natural selection, human populations, malaria, *Plasmodium*, genetic variant, resistance.

Le paludisme est un problème majeur de santé publique dans les zones tropicales. Le nombre de cas cliniques est estimé entre 300 et 500 millions par an, et la mortalité annuelle varie entre 1,5 et 2 millions. Plusieurs espèces plasmodiales sont responsables du paludisme humain,

*Plasmodium falciparum* étant responsable de la grande majorité des décès. Après transmission par un anophèle femelle (le vecteur), les parasites migrent vers le foie, envahissent les hépatocytes et prolifèrent. À la fin de la maturation, la membrane des hépatocytes se rompt, libérant ainsi les parasites dans la circulation sanguine. Les parasites envahissent alors les globules rouges, dans lesquels ils prolifèrent. La rupture des globules rouges infectés provoque la fièvre, conduisant à une forme clinique dite simple. Des complications peuvent survenir, souvent associée à la cytoadhérence des globules rouges infectés au niveau microvasculaires dans les organes profonds. Les stades sanguins sont responsables des formes cliniques simples (fièvre) et graves (anémie sévère, neuropaludisme et détresse respiratoire sévère).

La mortalité est essentiellement infantile, et elle était plus élevée avant l'apparition des mesures de contrôles (drogues anti-plasmodiales, moustiquaires imprégnées). Elle a probablement très fortement augmenté il y a 5 000 à 10 000 ans avec le développement de l'agriculture et de la sédentarisation des populations humaines ; ces changements auraient provoqué l'augmentation de la densité du vecteur transmettant le parasite et celle de la densité des populations humaines (Carter & Mendis 2002) et ainsi facilité la transmission du paludisme. Le nombre de cas observés annuellement et la mortalité infantile (avant l'âge de reproduction) expliquent pourquoi le paludisme est considéré comme la plus forte pression de sélection<sup>1</sup> naturelle dans l'histoire récente du génome humain. Cette hypothèse repose sur 1) l'identification de gènes et de leurs variants associés à la résistance au paludisme, 2) l'évaluation de la date d'apparition de ces variants génétiques, 3) la détection d'une signature de sélection au niveau de ces gènes.

Haldane a été le premier à proposer, d'une part, l'influence du paludisme sur le génome humain, et, d'autre part, l'effet protecteur contre le paludisme de variants génétiques provoquant la réduction ou l'absence d'une ou plusieurs sous-unité(s) de l'hémoglobine, causant ainsi les thalassémies (Weatherall 2008). De la même manière, des variants génétiques causant d'autres hémoglobinopathies ont été ensuite associés à la résistance au paludisme et des analyses récentes indiquent une sélection balancée de ces variants, sous la pression positive du paludisme et celle négative de l'hémoglobinopathie. Toutefois, si l'effet protecteur d'un variant génétique donné causant une hémoglobinopathie est très net, il n'explique pas la résistance observée chez les très nombreux sujets vivant

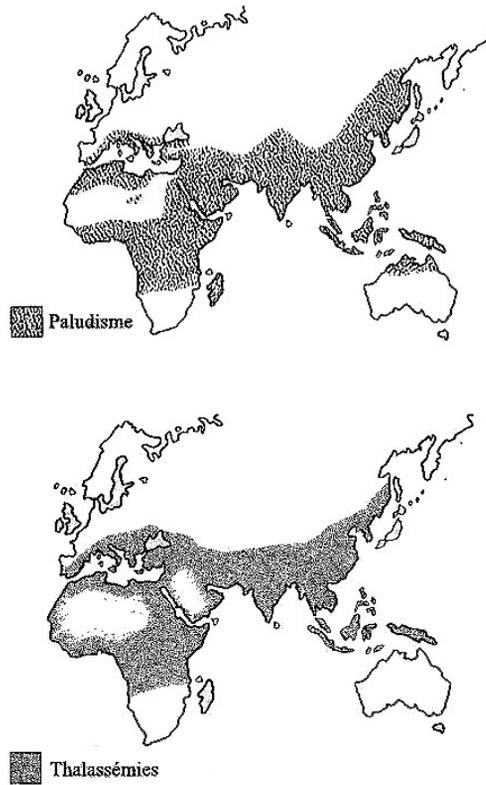
1. Pour une présentation générale de la sélection naturelle, voir l'article d'Alex Clamens, p. 17.

en zone d'endémie palustre qui ne portent pas ces variants génétiques. Cela indique clairement l'existence d'autres gènes de résistance au paludisme. De nombreux variants génétiques impliqués dans le système immunitaire ont été également associés à la résistance au paludisme, avec généralement un effet protecteur plus faible que celui observé pour les hémoglobinopathies (Rihet 2002 ; Kwiatkowski 2005).

L'objectif de l'article est de montrer non seulement quelques cas d'école permettant d'illustrer le concept de la sélection de variants génétiques conférant la résistance au paludisme, mais aussi les résultats et les perspectives de recherches plus récentes concernant notamment les gènes du système immunitaire.

### **Hémoglobinopathies, protection contre le paludisme à *Plasmodium falciparum* et sélection balancée ou positive**

Les thalassémies sont causées par de nombreux variants génétiques au niveau des gènes *HBA1*, *HBA2* (chromosome 16p13.3) et *HBB* (chromosome 11p15.5) codant respectivement pour les chaînes  $\alpha$ -globine et  $\beta$ -globine constituant les sous-unités de l'hémoglobine adulte. Les sujets atteints par une alpha-thalassémie sévère ont perdu deux ( $\alpha$ -/ $\alpha$ - ou  $\alpha\alpha$ /--) ou trois (- $\alpha$ /-- ou  $\alpha$ -/--) gènes codant pour la chaîne  $\alpha$ -globine, tandis que les sujets non atteints ont quatre gènes codant pour la chaîne  $\alpha$ -globine ( $\alpha\alpha/\alpha\alpha$ ) ; la perte des deux copies du gène *HBA1* combinée à celle des deux copies du gène *HBA2* (--/--) est létale. Les variants au niveau du gène *HBB* causant une bêta-thalassémie diminuent (type  $\beta^+$ ) voire suppriment (type  $\beta^0$ ) totalement l'expression du gène ; les symptômes les plus sévères, incluant une anémie sévère, sont observés chez les sujets  $\beta^+\beta^0$  et  $\beta^0\beta^0$ . Les thalassémies sont fréquentes et leur répartition géographique se superpose à la répartition passée et présente du paludisme en Europe, en Afrique et en Asie (figure 1 page suivante ; Weatherall 2008). De plus, une étude conduite dans les îles du Pacifique a montré une corrélation entre la fréquence des thalassémies et l'intensité de la transmission du paludisme. Ces observations suggèrent que les variants génétiques causant les thalassémies ont été maintenus sous la pression de sélection du paludisme. Cette hypothèse exprimée pour la première fois par Haldane en 1949 implique l'effet protecteur de ces variants contre le paludisme. Elle a été confirmée par des études d'association génétique dans plusieurs populations, notamment en Nouvelle-Guinée et en Afrique (Weatherall 2008). Les mécanismes de la résistance liée aux thalassémies ne sont pas complètement élucidés.



**Figure 1 :**

Comparaison des distributions géographiques passées et présentes du paludisme et des thalassémies en Afrique, en Eurasie et dans le sud-ouest du Pacifique (modifié d'après Weatherall 2008)

Néanmoins, des études *in vitro* montrent un mauvais développement du parasite dans les globules rouges des sujets thalassémiques et l'augmentation de la fixation d'anticorps à la surface des globules rouges infectés.

La drépanocytose ou anémie à hématies falciformes est causée par une mutation ponctuelle située au niveau du sixième codon<sup>2</sup> du gène *HBB* ( $GAG \Rightarrow GTG$ ) et provoquant le remplacement de l'acide glutamique par la valine ; la dénomination historique du variant impliqué est

2. Le codon est l'unité d'information du code génétique. Il est constitué de trois nucléotides successifs formant un triplet. Les quatre nucléotides présents dans les acides nucléiques (ADN ou ARN) permettent de constituer 64 codons (4<sup>3</sup>). Lors de la synthèse des protéines, l'appareil de traduction établit une correspondance entre un codon et un acide aminé. L'existence de codons « synonymes » limite à 20 le nombre d'acides aminés qui constituent les protéines. Le démarrage et la fin de la synthèse sont dus à l'existence de codons particuliers respectivement appelés codon initiateur et codon stop.

l'hémoglobine S et l'identifiant du polymorphisme dans les bases de données génomiques est rs334 (<http://www.ncbi.nlm.nih.gov/snp/>). Les propriétés physico-chimiques de ces acides aminés sont différentes, l'acide glutamique étant chargé négativement et la valine étant hydrophobe. Ce changement favorise, en condition d'hypoxie (faible taux d'oxygène), la polymérisation des chaînes de l'hémoglobine, la formation de fibrilles, la déformation et la rigidification des globules rouges, et leur adhésion aux vaisseaux sanguins, provoquant ainsi une forte anémie et une perturbation du flux sanguin. La drépanocytose est une maladie récessive, la maladie s'exprimant chez les homozygotes pour l'hémoglobine S (Hb SS); les homozygotes Hb SS ont une espérance de vie diminuée par rapport aux hétérozygotes (Hb AS) et aux individus sains (homozygotes Hb AA). Ainsi, la drépanocytose exerce une sélection négative sur l'hémoglobine S. Pourtant, Allison observait en 1954 que la fréquence de l'hémoglobine S était particulièrement élevée dans des zones hyper-endémiques pour le paludisme, principalement en Afrique sub-saharienne; il proposait que ce variant génétique ait un effet protecteur contre le paludisme et ait été maintenu sous la pression du paludisme. Par la suite, de nombreuses études ont montré l'association de l'hémoglobine S avec la protection contre le paludisme; en particulier, l'hétérozygotie pour l'hémoglobine S diminue d'un facteur 10 le risque de développer une forme grave de paludisme. Les études *in vitro* ont permis de proposer les mécanismes suivants: le parasite ne se multiplie pas en condition d'hypoxie dans des globules rouges Hb SS ou Hb AS, et les phagocytes éliminent préférentiellement les globules rouges Hb AS ou Hb SS infectés en coopérant avec les anticorps dirigés contre les antigènes présents à la surface des globules rouges infectés, les sujets porteurs de l'hémoglobine S ayant des taux plus élevés d'anticorps. L'absence d'effet physiopathologique majeur et l'effet protecteur contre le paludisme de l'hémoglobine S à l'état hétérozygote confèrent très probablement un avantage sélectif aux individus Hb AS. La combinaison des pressions négative (drépanocytose) et positive (paludisme) correspond à une sélection dite balancée, avec une fréquence de l'hémoglobine S atteignant au fil des générations une valeur intermédiaire (figure 2). L'analyse des polymorphismes voisins (figure 2) a montré que l'hémoglobine S était apparue récemment de manière indépendante au niveau d'haplotypes<sup>3</sup> différents dans différentes populations (table 1).

3. Un haplotype est une combinaison d'allèles (variants génétiques) liés physiquement car portés par le même chromosome. Un individu donné possède, pour une région chromosomique donnée, deux haplotypes, l'un hérité de son père et l'autre hérité de sa mère.

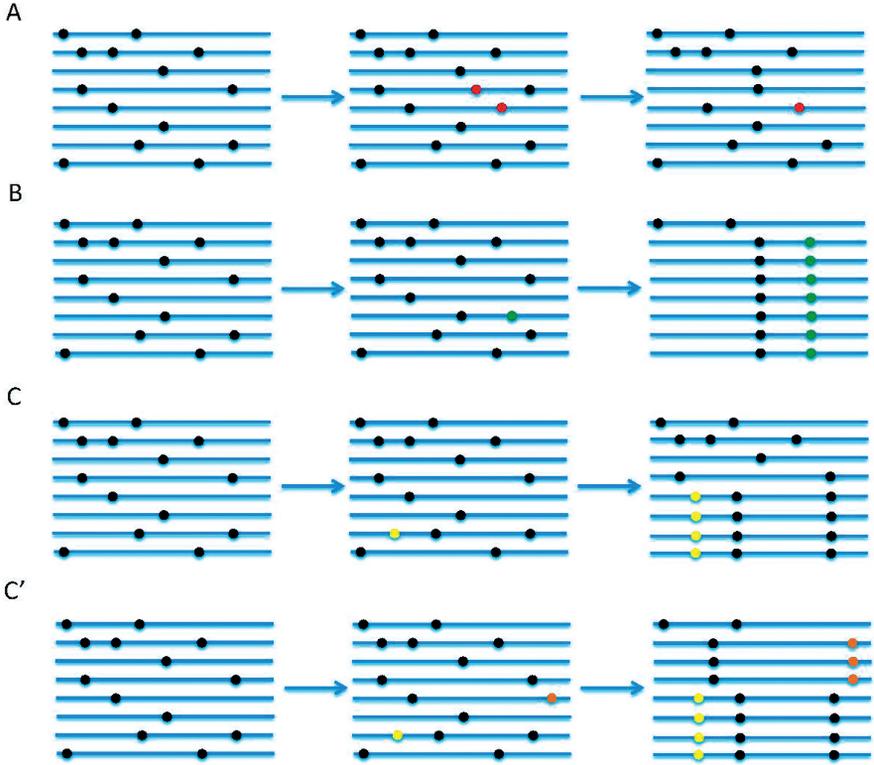
L'hémoglobine C (rs33930165) est un autre variant du gène *HBB* également situé au niveau du sixième codon, mais pas sur le même nucléotide que pour l'hémoglobine S (GAG=>AAG); l'acide glutamique est remplacé par une lysine. Les homozygotes ont des signes cliniques bénins, une anémie hémolytique modérée, et les hétérozygotes ne présentent pas de signe clinique. Les homozygotes et les hétérozygotes sont protégés contre le paludisme, l'effet protecteur étant toutefois plus marqué chez les homozygotes (Kwiatkowski 2005; Weatherall 2008). *In vitro*, le parasite se développe mal dans des globules rouges Hb CC, mais tout à fait normalement dans les globules rouges Hb AC. L'exposition des antigènes du parasite est modifiée à la surface des globules rouges AC et CC, augmentant probablement leur élimination par les phagocytes. L'hémoglobine C est concentrée en Afrique de l'Ouest, et dans certaines populations, sa fréquence dépasse celle de l'hémoglobine S. Les données actuelles suggèrent l'existence d'une sélection positive de l'hémoglobine C qui pourrait mener à la fixation de ce variant en Afrique de l'Ouest, correspondant à une fréquence de l'hémoglobine C approchant 100 %. L'analyse haplotypique (figure 2) de la région contenant l'hémoglobine C indique une origine unique en Afrique, avec l'apparition récente de la mutation au niveau d'un haplotype unique. Néanmoins, l'hémoglobine C a été également retrouvée, en Thaïlande, sur un haplotype non-Africain, indiquant que la mutation est apparue de manière indépendante dans les populations africaines et asiatiques.

### **Autres variations génétiques affectant la physiologie du globule rouge, protection contre le paludisme et sélection positive**

La déficience en G6PD<sup>4</sup> est l'enzymopathie la plus fréquente chez l'homme. Elle est causée par de nombreux variants affectant la séquence protéique et la fonction enzymatique. La distribution géographique de cette déficience et l'analyse haplotypique du locus G6PD indiquent une sélection positive récente sous la pression du paludisme (figures 2 et 3, et table 1). La forme commune G6PD A- (rs1050828) correspond à une mutation ponctuelle et à un changement d'acide aminé diminuant à la fois l'activité enzymatique et la multiplication du parasite dans les globules rouges *in vitro*. L'enzyme déficiente ne protégerait plus le parasite du stress

---

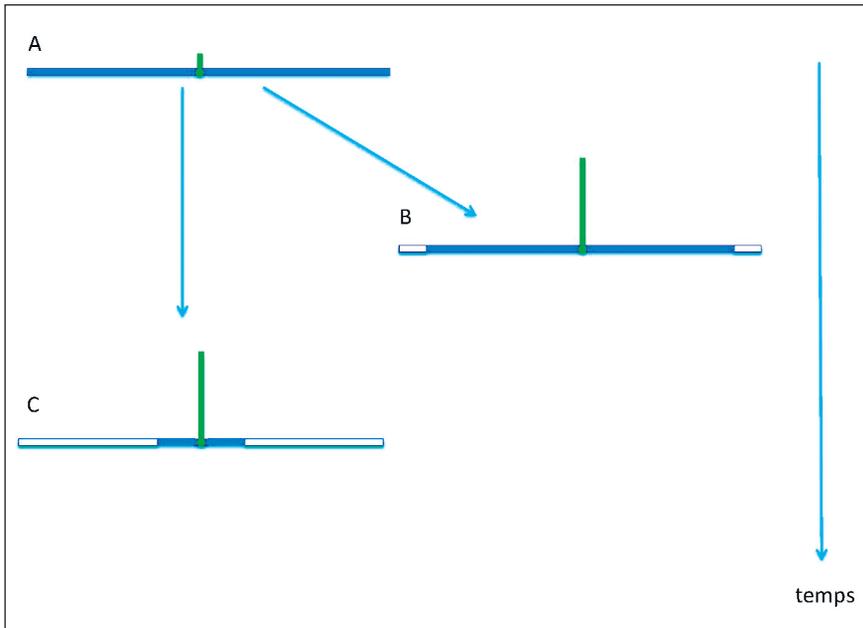
4. L'enzyme **G**lucose **6** **P**hosphate **D**éshydrogénase intervient dans la protection des globules rouges contre le stress oxydatif lié à la forte présence de dioxygène dans ces cellules. Elle permet en quelque sorte de détoxifier le dioxygène.



**Figure 2 :**  
**Sélection d'allèles et structure haplotypique dans une population**

Chaque trait représente une séquence d'ADN génomique, c'est-à-dire une région d'un chromosome donné. Les mutations neutres sont indiquées par un rond noir. Une combinaison particulière d'allèles (variants génétiques) liés physiquement est appelée haplotype ; 8 haplotypes différents sont représentés à l'étape initiale.

- A) **Sélection négative** d'allèles défavorables représentés en rouge. La mutation apparaît dans un haplotype particulier, créant ainsi un nouvel haplotype. Le mutant défavorable et les mutants neutres portés par le même chromosome sont maintenus à une très faible fréquence et progressivement éliminés.
- B) **Sélection positive** de l'allèle favorable représenté en vert. La mutation apparaît dans un haplotype particulier, créant ainsi un nouvel haplotype. La fréquence du mutant favorable et celle du mutant neutre porté par le même haplotype augmentent jusqu'à atteindre 1.
- C) **Sélection balancée** ou équilibrée. La mutation (jaune) apparaît dans un haplotype particulier, créant ainsi un nouvel haplotype. Le mutant subit une pression de sélection négative et une pression de sélection positive. L'hémoglobine S est le cas d'école, avec la sélection négative des homozygotes SS sous la pression de la drépanocytose, et la sélection positive des porteurs de l'hémoglobine S sous la pression du paludisme ; les hétérozygotes AS sont ainsi avantagés. La fréquence du mutant soumis aux pressions de sélection et celle du mutant neutre porté par le même haplotype augmentent jusqu'à un niveau intermédiaire.
- C') **Sélection balancée** ou équilibrée. Deux mutations favorables apparaissent dans deux haplotypes différents, conférant éventuellement un avantage aux hétérozygotes. La sélection positive des deux allèles conduit à l'évolution des fréquences vers un niveau intermédiaire.



**Figure 3 :**  
**Longueur des haplotypes et sélection positive récente**

La fréquence d'un allèle/variant génétique neutre peut augmenter sous l'effet du hasard, sans l'effet d'une pression de sélection particulière (théorie neutraliste, dérive génétique aléatoire). La fréquence de l'haplotype portant cet allèle augmente également. Cependant, les événements de recombinaison réduisent la longueur de l'haplotype au fil des générations. Une pression de sélection positive augmente relativement rapidement la fréquence d'un allèle favorable et celle de l'haplotype. Si la mutation est récente, les événements de recombinaison réduisent peu la longueur de l'haplotype. La présence d'haplotypes longs à une fréquence élevée dans la population d'étude est la signature d'une sélection positive récente.

- A) Haplotype portant le nouvel allèle (vert) avec une faible fréquence représentée par la barre verticale verte.
- B) Haplotype portant l'allèle favorable sélectionné avec une fréquence élevée représentée par la barre verticale verte.
- C) Haplotype portant l'allèle sélectivement neutre avec une fréquence élevée.

**Table 1 : Estimation de l'âge de variants génétiques protégeant contre le paludisme**

Gène-allèle	HbS	HbC	G6PD A-	DARC/FY	HLA-B53
Âge (années)	1125-1750	950-3000	3840-11760	3625-13125	2500

oxydatif que le parasite provoque lui-même en dégradant l'hémoglobine, entraînant ainsi l'accumulation du fer. La forme commune G6PD A- est responsable de 90 % des cas de déficience en G6PD en Afrique tropicale est associée à la protection contre le paludisme dans plusieurs populations vivant en zone d'endémie (Cappellini & Fiorelli 2008).

Les populations de l'Afrique sub-saharienne sont complètement résistantes à l'infection par *P. vivax*. Cette résistance est due à la mutation FY\*O (rs2814778) située dans le promoteur du gène *DARC/FY*<sup>5</sup> qui empêche l'expression de DARC à la surface du globule rouge (Kwiatkowski 2005). La protéine DARC, aussi connue sous le nom d'antigène Duffy, est un co-récepteur obligatoire de *P. vivax*, et son absence rend les globules rouges réfractaires à l'invasion par *P. vivax*. L'influence de la mutation FY\*O sur l'expression du gène est bien décrite au niveau moléculaire. Alors que le variant FY\*A permet la liaison du facteur de transcription GATA1 au niveau de son site de liaison AGATAA, le variant FY\*O réduit considérablement la fixation de GATA1 au niveau du site modifié AGGTAA, et ainsi l'expression de DARC/FY. Le variant FY\*O (rs2814778) est pratiquement fixé dans les populations de l'Afrique centrale et de l'Afrique de l'Ouest, avec une fréquence proche de 1 (Kwiatkowski 2005), tandis qu'il est généralement très rare dans les autres populations humaines. Le même variant est apparu indépendamment sur un haplotype non africain dans une population vivant en Nouvelle Guinée et est associé à une prévalence diminuée du paludisme à *P. vivax*.

### **Variations affectant des gènes de la réponse immunitaire, protection contre le paludisme à *Plasmodium falciparum* et sélection naturelle**

De nombreux gènes de la réponse immunitaire<sup>6</sup> ont été étudiés dans des populations humaines vivant en zone d'endémie palustre, et certains sont associés à la résistance ou à la susceptibilité au paludisme (table 2 ;

5. Le gène DARC/FY contrôle la synthèse d'une protéine membranaire des globules rouges qui est un récepteur de l'interleukine 8. Cette protéine permet également l'arrimage du parasite au globule rouge et participe ainsi à sa pénétration dans la cellule hôte.

6. Pour une présentation du système immunitaire, voir l'article d'Éric Espinosa, p. 121.

**Table 2 : Gènes impliqués dans la réponse immunitaire et associés à la résistance ou à la susceptibilité au paludisme**

Gène	Localisation	Protéine et fonction	Sélection positive
<i>CD36<sup>a</sup></i>	7q11.2	CD36, phagocytose et cytoadhérence	OUI
<i>CD40LG</i>	Xq26	CD40 ligand, prolifération et différenciation des lymphocytes B	OUI
<i>CR1<sup>a</sup></i>	1q32	CR1, récepteur aux molécules du complément, cytoadhérence et opsonisation	OUI
<i>FCGR2A</i>	1q21-q23	FCγRIIA, récepteur aux IgG, phagocytose et cytotoxicité	-
<i>HAVCR1</i>	5q33	TIM3, à la surface des lymphocytes T auxiliaires Th1	OUI
<i>HLA-B</i>	6p21.3	HLA-B, présentation des antigènes aux lymphocytes T cytotoxiques	-
<i>HLA-DR</i>	6p21.3	HLA-DR, présentation des antigènes aux lymphocytes T auxiliaires	-
<i>ICAM1<sup>a</sup></i>	19p13	CD54, cytoadhérence des leucocytes	-
<i>IFNAR1</i>	21q22.1	Récepteur à l'interferon alpha	-
<i>IFNG</i>	12q14	Interferon gamma, cytokine pro-inflammatoire	-
<i>IFNGR1</i>	6q23-q24	Récepteur à l'interferon gamma	-
<i>IL1A</i>	2q14	Interleukine 1 alpha, cytokine pro-inflammatoire	OUI
<i>IL1B</i>	2q14	Interleukine 1 beta, cytokine pro-inflammatoire	OUI
<i>IL4</i>	5q31	Interleukine 2, production d'anticorps	OUI
<i>IL10</i>	1q31-q32	Interleukine 10, cytokine anti-inflammatoire	-
<i>IL12B</i>	5q33	Interleukine 12, développement des lymphocytes T auxiliaires Th1 et des lymphocytes tueurs NK	-
<i>IL13</i>	5q31	Interleukine 13, production d'anticorps	OUI
<i>IRF1</i>	5q31	IRF1, facteur de transcription, régulation de l'expression de gènes (IFNA, IL12B, NOS2...)	OUI
<i>LTA</i>	6p21.3	Lymphotoxine exprimée par les lymphocytes T, développement d'organe lymphoïde, apoptose	-
<i>MBL2</i>	10q11	Lectine se liant au mannose, activation de la voie classique du complément	-
<i>NCR3<sup>b</sup></i>	6p21.3	NKp30, récepteur activateur des lymphocytes tueurs NK	-
<i>NOS2A</i>	17cen-q11.2	Enzyme impliquée dans la production d'une molécule cytotoxique, le NO	-
<i>PECAM1<sup>a</sup></i>	17q23	Molécule d'adhérence à la surface des leucocytes et des cellules endothéliales	OUI
<i>TIRAP</i>	11q24	Protéine impliquée dans la voie de signalisation de TLR4	-
<i>TNF</i>	6p21.3	Cytokine pro-inflammatoire, apoptose	-
<i>TLR4</i>	9q32-q33	Récepteur de l'immunité innée reconnaissant des motifs moléculaires de pathogènes	-

a. Plusieurs molécules du système immunitaire sont impliquées dans la cytoadhérence et la séquestration des globules rouges infectés au niveau des organes profonds, comme le cerveau. Cette séquestration joue un rôle majeur dans la physiopathologie du paludisme. L'antigène parasitaire PfEMP1 présent à la surface du globule rouge interagit directement avec des protéines membranaires de l'hôte.

b. Le récepteur NKp30 est présent à la surface des lymphocytes cytotoxiques NK. Il reconnaît l'antigène parasitaire PfEMP1, et est impliqué dans l'élimination des globules rouges infectés.

Rihet 2002 ; Kwiatkowski 2005). Plusieurs observations suggèrent qu'il y a eu une sélection négative, balancée, ou positive au niveau de ces gènes.

Ainsi, aucun allèle changeant la séquence en acide aminé n'est détecté dans les populations humaines vivant en zone d'endémie au niveau de gènes ayant un rôle central dans la réponse immunitaire, comme *IL12B*<sup>7</sup>. Ce type d'allèles est détecté chez quelques individus présentant une profonde immuno-déficiência nécessitant une hospitalisation (susceptibilité face à de multiples infections bactériennes). Ces observations suggèrent que ce type d'allèle existe à une très faible fréquence, et est sélectionné négativement sous la pression du paludisme et d'autres maladies infectieuses.

Le polymorphisme observé au niveau des gènes HLA<sup>8</sup> est un exemple classique de sélection balancée et de l'avantage aux hétérozygotes, *a priori* capables de produire une réponse immunitaire dirigée contre un plus grand nombre de souches d'un même pathogène ou de pathogènes d'espèces différentes. En effet, chez les hétérozygotes pour un gène HLA donné, les cellules sont capables de présenter aux lymphocytes T un répertoire antigénique plus grand par les deux variants de la molécule HLA. Néanmoins, certains allèles du HLA sont plus favorables, comme l'allèle HLA-B53, associé à la protection contre le paludisme et impliqué dans la présentation par les hépatocytes infectés d'un peptide antigénique du parasite aux lymphocytes T cytotoxiques (Rihet 2002 ; Kwiatkowski 2005). L'âge estimé de l'allèle est indiqué dans la table 1.

La table 2 montre les gènes du système immunitaire associés à la résistance ou à la susceptibilité au paludisme. Plusieurs de ces gènes sont localisés dans les régions chromosomiques 5q31-q33 et 6p21.3. Pour certains d'entre eux, une signature moléculaire de sélection positive a été détectée sur la base de l'étude des haplotypes (figure 3). Quelques travaux réalisés dans des populations africaines sur des gènes candidats ont révélé l'existence d'une sélection positive récente pour les gènes *CD40LG* et *CD36*. Plusieurs études récentes conduites à l'échelle du génome dans différentes populations humaines ont permis la détection de signatures moléculaires de la sélection positive récente pour de nombreuses régions génomiques, avec un enrichissement significatif en régions correspondant

7. L'interleukine 12 est impliquée dans la différenciation des lymphocytes T auxiliaires, et oriente ainsi la nature de la réponse immunitaire. Elle est aussi un facteur de croissance essentiel des cellules tueuses NK.

8. Les molécules HLA de classe I permettent aux cellules de l'organisme de présenter des antigènes aux récepteurs des lymphocytes T cytotoxiques. Les molécules HLA de classe II sont exprimées à la surface de cellules spécialisées dites présentatrices (monocytes/macrophages, cellules dendritiques, lymphocytes B), et permettent la présentation des antigènes aux récepteurs des lymphocytes T auxiliaires.

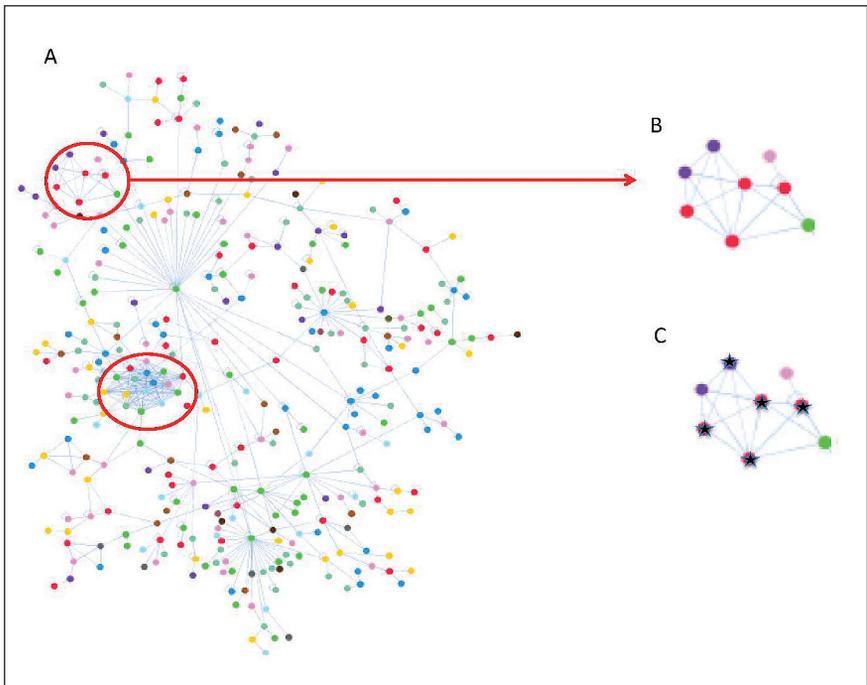
à des gènes de la réponse immunitaire (Barreiro & Quintana-Murci 2010). Ces données suggèrent un impact important des maladies infectieuses – y compris le paludisme – sur la sélection de variants génétiques humains.

Malgré ces observations, nous n'avons probablement qu'une idée partielle de la sélection naturelle des variants génétiques impliqués dans le système immunitaire. Celui-ci est en effet un système biologique complexe, impliquant de nombreuses interactions moléculaires et cellulaires, et la réponse de l'hôte à une infection dépend, sur les plans quantitatif et qualitatif, d'une combinaison d'allèles au niveau de plusieurs gènes et non pas d'un seul gène. Les interactions moléculaires directes impliquent l'existence de gènes dont les variations de séquence peuvent influencer l'interaction au sens physico-chimique entre les deux partenaires moléculaires ; il est tout à fait admis que les gènes impliqués puissent co-évoluer sous une pression de sélection particulière, aboutissant à l'augmentation de la fréquence de combinaisons favorables d'allèles des deux partenaires. Cette idée a été proposée notamment pour les gènes codant pour les cytokines et leurs récepteurs. De la même façon, il est vraisemblable que plusieurs gènes impliqués dans le même réseau moléculaire subissent une pression de sélection aboutissant à l'augmentation de la fréquence de certaines combinaisons alléliques favorables.

Si la biologie des systèmes permet d'étudier l'influence des variants génétiques sur le comportement des réseaux moléculaires, il y a peu de réflexion sur l'évolution de tels systèmes complexes dans les populations humaines soumises à une pression de sélection donnée. De nouvelles approches dans ce domaine devraient offrir un éclairage très intéressant sur l'évolution des gènes de la réponse immunitaire sous la pression des maladies infectieuses, comme le paludisme. À titre d'exemple, la figure 4 montre une partie de l'interactome humain, c'est-à-dire la carte des interactions directes entre les protéines humaines ; on y distingue des sous-réseaux constitués de protéines dont les gènes sont localisés sur différents chromosomes. Il est possible également de dessiner des réseaux moléculaires basés sur les interactions ADN-protéine ou ARN-ARN<sup>9</sup>. Quelle que soit la nature du réseau moléculaire, on peut proposer que les combinaisons d'allèles favorables qui modifient les interactions physiques entre les molécules soient co-sélectionnées sous l'effet d'une pression de sélection donnée, comme le paludisme.

---

9. Il s'agit de réseaux de régulation d'expression génique. La protéine peut être un facteur de transcription se liant à un site de liaison spécifique au niveau du promoteur, régulant ainsi la transcription du gène cible. Un micro ARN est un petit ARN se liant à un site spécifique situé dans une partie non traduite de l'ARN messager, ce qui provoque sa dégradation.



**Figure 4 :**  
**Le réseau d'interaction entre protéines humaines et la sélection naturelle :**  
**une hypothèse**

- A) La figure représente une partie du réseau d'interaction entre protéines. Chaque nœud/rond représente une protéine, et chaque trait reliant les protéines une interaction directe ; les couleurs représentent la localisation des gènes correspondant sur différents chromosomes. À titre d'exemples, deux sous-réseaux particulièrement denses sont entourés ; plusieurs protéines ont la particularité d'interagir avec plusieurs autres protéines (entre 2 et 16 interacteurs). Ce type de sous-réseaux correspond à des ensembles fonctionnels, comme des complexes multi-protéiques ou à des voies de signalisation.
- B) L'un des sous-réseaux est agrandi. Les gènes correspondants ne présentent pas de mutations modifiant la séquence protéique.
- C) Le même sous-réseau est représenté avec des variants génétiques affectant la séquence protéique et les interactions entre les protéines. Les mutations sont représentées sous forme d'étoile. Elles modifient la fonctionnalité du réseau. On propose ici que la combinaison de variants favorables au processus physiologique, dans lequel le sous-réseau est impliqué, puisse être sélectionnée.



## Conclusions et perspectives

Les travaux sur le paludisme ont été les premiers à illustrer le concept de la sélection récente de variants génétiques favorables dans des populations humaines. Des cas de sélection positive et balancée pour des variants génétiques ayant un fort effet protecteur ont été ainsi mis en évidence au niveau des gènes impliqués dans la physiologie du globule rouge, l'un des principaux habitats du parasite. L'hémoglobine C est un cas documenté très récemment. Correspondant à une mutation voisine de celle produisant l'hémoglobine S, les deux se produisant dans le même codon et affectant le même acide aminé.

Plusieurs gènes impliqués dans la réponse immunitaire sont également associés à la résistance ou à la susceptibilité au paludisme, avec toutefois un effet protecteur moindre comparé à celui des gènes impliqués dans la physiologie du globule rouge. Plusieurs études conduites à l'échelle du génome humain<sup>10</sup> visaient l'identification de signatures moléculaires de la sélection naturelle ; elles ont pointé de nombreux gènes du système immunitaire, parmi lesquels plusieurs sont associés à la résistance au paludisme. Néanmoins, il est difficile d'affirmer que les variants génétiques ont été sélectionnés sous la pression exclusive du paludisme. Par ailleurs, le système immunitaire est un système biologique complexe, et de nouvelles découvertes sont attendues dans la compréhension des mécanismes protecteurs contre le paludisme et également de la sélection de variants génétiques favorables. La biologie des systèmes est une nouvelle approche permettant de prédire le comportement d'un réseau moléculaire et son influence sur un phénotype à l'échelle cellulaire, du tissu ou de l'organisme ; les découvertes dans ce domaine devraient orienter les recherches dans le domaine de l'évolution.

Enfin, d'autres gènes et réseaux moléculaires sont très probablement impliqués dans la résistance ou la susceptibilité au paludisme. Les recherches continuent activement dans ce domaine. Les technologies à haut débit<sup>11</sup> génèrent aujourd'hui une masse considérable de données et nous donnent l'espoir d'identifier, sans hypothèse *a priori*, de nouveaux gènes

10. Ces études sont basées sur l'analyse, dans des populations humaines, de plusieurs centaines de milliers de polymorphismes répartis sur le génome. Il ne s'agit pas de séquençage du génome complet de chaque sujet inclus dans l'étude. Un projet en cours, « le projet 1 000 génomes », vise à la séquence complète du génome de plus de 1 000 individus. L'analyse de ces données devrait affiner les signatures moléculaires de sélection.

11. Ces technologies, incluant la technologie des biopuces, permettent l'analyse dans de nombreux échantillons de plusieurs centaines de milliers de polymorphismes ou de l'ensemble des ARN. Les technologies les plus récentes sont celles du séquençage à très haut débit, qui permettent l'analyse des variations génétiques, épigénétiques et transcriptomiques.

impliqués dans la résistance ou la susceptibilité aux maladies infectieuses, et au paludisme en particulier. Ainsi, plusieurs études génétiques conduites à l'échelle du génome dans plusieurs populations humaines ont permis la localisation de nouveaux loci de résistance au paludisme. De plus, les études du transcriptome (l'ensemble des ARN messagers présents dans une cellule ou un tissu) de sujets vivant en zone d'endémie et de souris infectées par *Plasmodium* ont permis l'identification de groupes de gènes sur- ou sous-exprimés chez les individus résistants. Les études chez la souris ont été conduites avec des lignées consanguines génétiquement résistantes ou génétiquement susceptibles, indiquant que les différences de séquence de l'ADN sont responsables à la fois des variations d'expression génique et du phénotype de résistance au paludisme. Plusieurs voies physiologiques ont été identifiées, et certaines permettent de prédire, dans le modèle souris, la mortalité causée par *Plasmodium*. Les gènes impliqués dans ces voies physiologiques devront être séquencés chez l'homme dans le but d'identifier les variants protecteurs. Nous espérons ainsi identifier de nouveaux variants génétiques conférant une certaine résistance au paludisme et potentiellement sélectionnés dans les populations humaines vivant en zone d'endémie.

## Bibliographie

- BARREIRO L.B. & QUINTANA-MURCI L. (2010). From evolutionary genetics to human immunology: how selection shapes host defence genes. *Nat. Rev. Genet.* **11**, 17-30.
- CAPPELLINI M.D. & FIORELLI G. (2008). Glucose-6-phosphate dehydrogenase deficiency. *Lancet* **371**, 64-74.
- CARTER R. & MENDIS K (2002). Evolutionary and historical aspects of the burden of malaria. *Clinical Microbiology Review* **15**, 564-594.
- KWIATKOWSKI D.P. (2005). How malaria has affected the human genome and what human genetics can teach us about malaria. *Am. J. Hum. Genet.* **77**, 171-192.
- RIHET P. (2002). Genetic predisposition to malaria: search of genes controlling the human resistance to malaria induced by *Plasmodium falciparum*. In *Susceptibilité aux maladies infectieuses* (éd. G. Milon). Annales de l'Institut Pasteur, Elsevier Masson SAS, Paris, 65-80.
- WEATHERALL D.J. (2008). Genetic variation and susceptibility to infection: The red cell and malaria. *Br. J. Haematol.* **141**, 276-286.