

Origine et adaptation de *Plasmodium falciparum*, agent de la malaria, en Amérique du Sud

Franck Prugnolle¹, Erhan Yalcindag¹, Eric Elguero¹, Celine Arnathau¹, Eric Legrand², Bernard Carme³, Christine Chevillon¹, Patrick Durand¹, François Renaud¹

1. Laboratoire MIVEGEC, UMR 224-5290 CNRS-IRD-UMI, IRD Montpellier, 911 Avenue Agropolis, BP 64501, 34394 Montpellier Cedex 5, France.

2. CNR Chimioresistance du Paludisme, Institut Pasteur de Cayenne, Cayenne, France

3. Laboratoire de Parasitologie-Mycologie, Université Antilles Guyane, Cayenne, France.

Introduction

Une manière d'appréhender l'adaptation des pathogènes à de nouveaux environnements est d'analyser d'anciens événements d'émergence (i.e. d'anciens événements de colonisation). L'agent de la malaria humaine, *Plasmodium falciparum*, est, en ce sens, un bon modèle biologique. En effet, au cours de son histoire évolutive, *P. falciparum* a colonisé de nouvelles populations et de nouveaux territoires de multiples fois. Un des événements connus d'émergence de *P. falciparum* dans une nouvelle aire géographique a été son introduction en Amérique du Sud, à partir de l'Afrique, au cours de la traite des Noirs (1-3). Au cours de cette colonisation du nouveau monde (intervenue dans les 500 dernières années), le parasite a rencontré des conditions environnementales radicalement différentes de celles qu'il connaissait sur le continent africain: un nouvel environnement hôte humain (des populations amérindiennes, des populations européennes, des populations africaines mixtes) ainsi qu'un nouvel environnement vecteur. *Anopheles darlingi* (sous-genre *Nyssorhynchus*), est le vecteur principal de *P. falciparum* en Amérique du sud (4), tandis qu'en Afrique les vecteurs majeurs sont *A. gambiae* et *A. funestus* (sous-genre *Cellia*). Les populations amérindiennes sont génétiquement très différentes des populations africaines (5). De la même manière, *A. darlingi* est très différent génétiquement des deux anophèles vecteurs majeurs d'Afrique, puisqu'il a divergé d'avec elles il y a environ 100 millions d'années (6).

Les changements environnementaux qui sont intervenus pendant et depuis la colonisation du nouveau monde par *P. falciparum* ont très probablement exercé de fortes pressions de sélection sur le parasite. En retour, le parasite a dû s'adapter à ces nouvelles conditions environnementales en évoluant vers de nouveaux phénotypes et donc de nouveaux génotypes. Comment le parasite s'est-il génétiquement adapté à ces nouveaux environnements? Quels gènes ont été impliqués? Comment ces gènes ont évolué dans ces nouveaux environnements?

L'objectif majeur de ce projet de recherche fut donc de déterminer mais aussi de comprendre l'évolution dans des environnements différents, des régions du génome de *P. falciparum* qui ont été la cible d'une sélection positive récente en Amérique du Sud et qui ont ainsi joué un rôle dans l'adaptation à son nouvel environnement, aussi bien vecteur qu'humain.

Afin de répondre à cet objectif, un préalable nécessaire était tout d'abord de bien caractériser l'histoire démographique et l'origine de *P. falciparum* en Amérique du Sud et ses conséquences sur la distribution du polymorphisme dans les populations sud-américaines du parasite.

Ce projet s'articulait donc autour de deux grandes questions :

- 1) Quelle est l'histoire démographique et l'origine des populations de *P. falciparum* en Amérique du Sud ?
- 2) Quelles régions du génome ont participé à l'adaptation du parasite aux nouvelles populations humaines colonisées ainsi qu'à la nouvelle espèce vectrice ?

Afin de répondre à ces différentes questions, nous avons utilisé les outils et les approches de génétique et de la génomique des populations. Nous avons d'une part utilisés différentes méthodes de phylogéographie pour reconstruire l'origine et l'histoire démographique de *P. falciparum* en Amérique du Sud [7] et d'autre part des méthodes de scan génomique pour déterminer les régions soumises à sélection positive récente dans les populations sud-américaines (7).

RESULTATS OBTENUS (ET PUBLIES)

Structuration et origine des populations Sud-Américaines de *P. falciparum*

Pour cerner l'histoire démographique de *P. falciparum* en Amérique du Sud, nous avons échantillonné 26 populations du parasite de part le monde: en Afrique, tout d'abord, car c'est la source supposée du pathogène. Sur ce continent, nous avons tout particulièrement échantillonné les zones d'origine des esclaves amenés en Amérique du Sud c'est-à-dire toute la bande côtière d'Afrique de l'ouest. Pour comparaison nous avons également échantillonné deux autres populations plus éloignées dont l'une est au Kenya et l'autre à Djibouti. En Amérique du Sud, nous avons essayé de couvrir au maximum l'aire de répartition actuelle du parasite (une grosse partie de l'échantillonnage a été réalisé par nos partenaires de l'Institut Pasteur et de l'Université Antilles-Guyane à Cayenne). Enfin, à titre comparatif, nous avons également collecté des échantillons en Asie (Iran, Thaïlande, Myanmar et Laos). Ces différentes collectes se sont faites au travers de multiples collaborations avec des chercheurs travaillant dans ces différents pays.

Pour l'ensemble de ces échantillons, nous avons génotypé 12 marqueurs microsatellites ainsi que 384 marqueurs SNPs (l'objectif initial était de ne génotyper que 12 marqueurs microsatellites).

Ce travail nous a permis de confirmer l'origine Africaine de *P. falciparum* mais également et surtout de mettre en évidence l'existence de deux clusters génétiques bien différenciés en Amérique du Sud, fait jusqu'alors inconnu. Un cluster « nord », incluant les populations Colombiennes et un cluster « sud » incluant les populations de Guyane Française, du Brésil et de Bolivie. Entre ces deux clusters, nous observons l'existence d'une zone hybride où les deux clusters se rejoignent (entre le Venezuela et le Pérou). Ces deux clusters génétiques ont tous les deux une origine Africaine. L'utilisation de méthodes statistiques ABC (pour « Approximate Bayesian Computation ») (8) nous a par ailleurs permis de démontrer que ces deux clusters génétiques avaient été introduits de manière indépendante en Amérique du Sud à partir de l'Afrique ainsi que d'estimer les paramètres démographiques des populations sud américaines comme la taille efficace (d'autres méthodes ont également été utilisées pour estimer ces tailles efficaces (9, 10)), la force des goulots d'étranglements au moment de leur introduction ainsi les temps de divergence avec les populations africaines (11). Nous avons ainsi pu estimer que les populations de *P. falciparum* avaient été introduites entre 300 et 700 ans Amérique du Sud (11).

Cette double introduction de *P. falciparum* en Amérique du Sud peut s'expliquer par l'histoire de la colonisation européenne du continent sud-américain et plus particulièrement la manière dont les différents empires (espagnols et Portugais) ont géré le commerce d'esclaves Africains (traite transatlantique). Entre le 16^{ème} et le 19^{ème} siècle, l'empire espagnol a importé ses esclaves en Amérique du sud en les débarquant au niveau du port de Cartagena (actuelle Colombie) alors que l'empire portugais les débarquait dans des ports proches des actuelles villes de Bahia et Rio de Janeiro, au Brésil.

Adaptation génétique aux nouvelles conditions environnementales

Concernant le second volet du projet, la partie relative à l'adaptation génétique du parasite à de nouveaux environnements, nous n'avons pas pu réaliser le génotypage des 3000 SNPs initialement prévu dans les populations Sud-Américaines afin de déterminer les régions du génome ayant connu un évènement de sélection positive récent. La raison à cela fut un problème d'ordre technique. L'ADN du parasite présent dans les extraits de sang de

patients infectés en Amérique du Sud n'était malheureusement pas en quantité suffisante pour réaliser ce typage, malgré une amplification globale de génome réalisée au préalable (comme prévu dans l'ANR).

Plusieurs alternatives à ce problème ont été envisagées et développées :

1) tout d'abord, nous avons génotypé 384 marqueurs SNPs répartis sur l'ensemble du génome dans 15 populations de part le monde, incluant toutes les populations sud-américaines. Bien qu'une partie de ces SNPs était située dans des régions non codantes, une bonne partie d'entre eux était présente dans des régions codantes. Le choix des SNPs codants s'est fait de deux manières : une partie a été sélectionnée de manière aléatoire dans le génome et une autre partie a été sélectionnée dans des gènes d'intérêts soit connus pour jouer un rôle dans la résistance aux anti-paludéens soit dans l'interaction hôte-parasite. Les résultats de cette analyse n'ayant pas encore été publiés, nous n'irons pas plus loin dans leur description.

2) Nous avons également décidé d'analyser de manière plus approfondie la sélection opérant sur les gènes EBAs (EBA 175, EBA 140 et EBA 181) dans les populations Sud-Américaines. Ces gènes jouent un rôle dans l'invasion du globule rouge par le parasite et interagissent avec la protéine Glycophorine A présente à la surface des hématies. Nous avons séquencé chacun de ces gènes (~ 1500bp pour chaque gène) dans 6 populations en Amérique du Sud, 2 en Afrique et une en Asie. Pour comparaison, nous avons également séquencé dans les mêmes populations deux gènes de ménage (SERCA (~4000bp et ADSL ~2000bp). A nouveau, les résultats de ces analyses n'ayant pas encore été publiés, nous n'irons pas plus dans leur description.

Résultats annexes 1.

Afin de pouvoir mener à bien ce projet et de palier aux problèmes rencontrés (problèmes aussi bien techniques, empiriques que théoriques), plusieurs développements ont été nécessaires et réalisés au cours de cette ANR. Parmi ces développements : 1) nous avons réalisé des simulations afin de comprendre comment la reproduction clonale et les caractéristiques d'échantillonnage chez *P. falciparum* pouvaient affecter les estimateurs des paramètres utilisés classiquement en génétique des populations (9). 2) Nous avons également développé une méthode permettant d'estimer l'effectif efficace de *P. falciparum* à partir de paramètres classiquement utilisés en génétique des populations (10).

Résultats annexes 2

Au cours de cette ANR et lors du génotypage des populations africaines de *P. falciparum* humains, nous avons décidé de génotyper, pour comparaison, des échantillons de *Plasmodium* issus de sangs de grands singes que nous avons à disposition au laboratoire et qui semblaient infectés par des parasites similaires morphologiquement à *P. falciparum*. Ce travail nous a permis de révéler la circulation de *P. falciparum* chez les gorilles et les petits singes en Afrique, fait jusqu'alors inconnu. Ces découvertes ont fortement bouleversé l'idée que l'on avait de l'origine de *P. falciparum* chez l'homme en Afrique. Cette ANR fut donc également à l'origine d'une découverte majeure concernant l'évolution de *P. falciparum* chez l'homme en Afrique (12-15).

REFERENCES

1. Anderson, T. J., et al. (2000) Microsatellite markers reveal a spectrum of population structures in the malaria parasite *Plasmodium falciparum*. *Mol Biol Evol* 17, 1467-82.
2. Conway, D. J. (2003) Tracing the dawn of *Plasmodium falciparum* with mitochondrial genome sequences. *Trends Genet* 19, 671-4.
3. Conway, D. J., et al. (2000) Origin of *Plasmodium falciparum* malaria is traced by mitochondrial DNA. *Mol Biochem Parasitol* 111, 163-71.
4. Mirabello, L. & Conn, J. E. (2006) Molecular population genetics of the malaria vector *Anopheles darlingi* in Central and South America. *Heredity* 96, 311-21.
5. Manica, A., Prugnolle, F. & Balloux, F. (2005) Geography is a better determinant of human genetic differentiation than ethnicity. *Hum Genet* 118, 366-71.
6. Krzywinski, J., Wilkerson, R. C. & Besansky, N. J. (2001) Toward understanding Anophelinae (Diptera, Culicidae) phylogeny: insights from nuclear single-copy genes and the weight of evidence. *Syst Biol* 50, 540-56.
7. Nielsen, R., et al. (2005) Genomic scans for selective sweeps using SNP data. *Genome Res* 15, 1566-75.
8. Cornuet, J. M., et al. (2008) Inferring population history with DIY ABC: a user-friendly approach to approximate Bayesian computation. *Bioinformatics* 24, 2713-2719.
9. Prugnolle, F. & De Meeus, T. (2010) Apparent high recombination rates in clonal parasitic organisms due to inappropriate sampling design. *Heredity* 104, 135-40.
10. Prugnolle, F., Durand, P., Renaud, F. & Rousset, F. (2010) Effective size of the hierarchically structured populations of the agent of malaria: a coalescent-based model. *Heredity* 104, 371-7.
11. Yalcindag, E., Elguero, E., & Prugnolle, F. (2011) Multiple independent introductions of *Plasmodium falciparum* in South America. *Proc Natl Acad Sci U S A*.
12. Ollomo, B., et al. (2009) A new malaria agent in African hominids. *PLoS Pathog* 5, e1000446.
13. Prugnolle, F., et al. (2011) *Plasmodium falciparum* is not as lonely as previously considered. *Virulence* 2, 71-6.
14. Prugnolle, F., et al. (2011) A fresh look at the origin of *Plasmodium falciparum*, the most malignant malaria agent. *PLoS Pathog* 7, e1001283.
15. Prugnolle, F., et al. (2011) African monkeys are infected by *Plasmodium falciparum* nonhuman primate-specific strains. *Proc Natl Acad Sci U S A* 108, 11948-53.