International Journal for Parasitology: Drugs and Drug Resistance 7 (2017) 174–180

Contents lists available at ScienceDirect

International Journal for Parasitology: Drugs and Drug Resistance

journal homepage: www.elsevier.com/locate/ijpddr

Genetic diversity of next generation antimalarial targets: A baseline for drug resistance surveillance programmes

Ana Rita Gomes [a], Matt Ravenhall [a], Ernest Diez Benavente [a], Arthur Talman [b], Colin Sutherland [a], Cally Roper [a], Taane G. Clark [a, c], Susana Campino [a, *]

[a] Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, UK
[b] Wellcome Trust Sanger Institute, Hinxton Cambridge, UK
[c] Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, UK

ARTICLE INFO

Article history:
Received 27 January 2017
Received in revised form
2 March 2017
Accepted 6 March 2017
Available online 9 March 2017

Keywords:
Antimalarial drugs
Drug-resistant mutations
Gene targets
Genetic diversity

ABSTRACT

Drug resistance is a recurrent problem in the fight against malaria. Genetic and epidemiological surveillance of antimalarial resistant parasite alleles is crucial to guide drug therapies and clinical management. New antimalarial compounds are currently at various stages of clinical trials and regulatory evaluation. Using ~2000 *Plasmodium falciparum* genome sequences, we investigated the genetic diversity of eleven gene-targets of promising antimalarial compounds and assessed their potential efficiency across malaria endemic regions. We determined if the *loci* are under selection prior to the introduction of new drugs and established a baseline of genetic variance, including potential resistant alleles, for future surveillance programmes.

1. Introduction

The continuous emergence and spread of resistance to first line antimalarial treatments, including artemisinin and its derivatives, threatens global efforts to reduce the burden of malaria. The development of a fully effective vaccine has been hampered by the complex life cycle of the malaria parasite and the high genetic diversity of key parasite antigens. Thus, antimalarial drugs, particularly those targeting basic cellular machinery common to all stages of the parasite life cycle, are the most promising approaches to control malaria.

The pipeline of antimalarial drugs has greatly expanded over the past decade, particularly because of the strong public-private partnerships and significant investment in innovative technologies (Flannery et al., 2013; Wells et al., 2015). A set of next generation antimalarial compounds, for which the molecular targets are known or being investigated, are currently at various phases of preclinical and clinical assessment (Wells et al., 2015).

Knowledge of parasite molecular drug targets can be exploited to monitor the potential emergence and spread of resistant alleles, particularly from the introduction of a drug, and rapidly inform local policies to tailor interventions. Without knowledge of antimalarial gene targets, the identification and surveillance of resistant alleles needs to be based on accurate clinical drug efficacy trials and genome-wide population genetic studies of field collected samples (Anderson et al., 2011). This approach can be both costly and labour intensive. Alternatively, a powerful strategy to identify mutations linked to resistance, prior to the licensing of a drug, is the use of laboratory-adapted strains to induce selection *in vitro* with sub-lethal and increasing concentrations of drugs. This strategy has led to the identification of polymorphisms in the *Plasmodium (P) falciparum kelch13* gene underlying resistance to artemisinin (Ariey et al., 2014). This gene was confirmed subsequently in association studies in field collected samples (Miotto et al., 2015) and using a reverse genetics approach (Ghorbal et al., 2014).

Here we consider eleven gene-targets of key investigated compounds that due to their efficiency might become the next antimalarial drugs, and for which mutations conferring resistance have been identified in *in vitro* studies (Baragaña et al., 2015; Dong et al., 2011; Flannery et al., 2015; Herman et al., 2015; Kato et al., 2016; LaMonte et al., 2016; Lim et al., 2016; McNamara et al., 2013; Ross et al., 2014). These 11 genes were also selected because they are gene-targets for a range of new antimalarial compounds already under evaluation in clinical trials. We present a survey of the natural genetic variation (SNPs, insertions and

* Corresponding author.
E-mail address: Susana.campino@lshtm.ac.uk (S. Campino).

http://dx.doi.org/10.1016/j.ijpddr.2017.03.001

deletions (indels), copy number variants (CNVs)) and diversity in these gene-targets using a publicly available global collection of ~2000 *P. falciparum* "field" parasite genomes from 18 countries. We use the variation to establish whether these regions are already under selective pressure, and report a baseline reference to assist future surveillance programmes with observing emergence of resistance mutations.

2. Materials and methods

Eleven gene-targets were analysed: *PF3D7_1113300 (Pfugt), PF3D7_1036800 (Pfact), PF3D7_0109800 (PfcPheRs), PF3D7_0603300 (Pfdhodh), PF3D7_0509800 (Pfpi4k), PF3D7_0321900 (Pfcarl), PF3D7_1211900 (Pfatp4), PF3D7_1451100 (PfeEF2), PF3D7_1320600 (Pfrab11A), PF3D7_1213800 (Pfprs)* and *Mal_Mito_3 (PfCYTB)* (see Table 1). Genome variation data was analysed for isolates from East Africa (Kenya, Tanzania, n = 35), West Africa (Burkina Faso, The Gambia, Ghana, Guinea, Mali, Nigeria, n = 521), Central Africa (Democratic Republic of Congo (DRC), n = 56), South America (Colombia, Peru, n = 24), South Asia (Bangladesh, n = 53) and Southeast Asia (Cambodia, Laos, Myanmar, Papua New Guinea, Thailand, Vietnam, n = 1187).

Sequencing data was generated by the Pf3k project (www.malariagen.net/pf3k), is open access and is described in (Miotto et al., 2015). Whole genome analysis of these data has also been recently described (Ravenhall et al., 2016) and we used a set of characterised high quality SNPs and indels identified in the 11 candidate target genes. In addition, larger structural variants (e.g. CNVs) in these regions were identified using *Delly* software (Rausch et al., 2012). Using the SNP variants, population genetic analyses were performed to establish if targeted coding regions are under selection. In particular, the Tajima's D method was applied to detect regions under balancing selection (R package Pegas); extended haplotype homozygosity approaches (|iHS|, XP-EHH) were applied to identify long-range positive directional selection, and F_{ST} statistics were used to assess population differentiation (see (Ravenhall et al., 2016) for a detailed description of these methods).

3. Results

Across the eleven gene-targets, a total of 778 SNPs were identified, with half (n = 424, 54.5%) leading to non-synonymous changes (Table 1, Supplementary Table 1). The overall genetic diversity was low, with the majority of SNPs (75.1%) having minor allele frequencies of less than 5%. The SNP density (number of SNPs per kbp) across genes was similar (~1 SNP per 33bp), except for those coding for the ras-related protein (*Rab11A*, 1 SNP per 258.6bp), elongation factor 2 (*eEF2*, 1 SNP per 73.4bp) and the acetyl-CoA transporter (*ACT*, 1 SNP per 64.2bp), all with lower density, suggesting greater gene conservation. The *pfact* gene was recently identified to be the target, together with the UDP-galactose transporter gene-target (*Pfugt*), of a variety of imidazolopiperazine compounds. One of these compounds (KAF156) has potent activity against gametocytes and parasite liver stages, and is currently in Phase II clinical trial (Lim et al., 2016). Rab11A is a molecular target for aminopyridine class compounds (McNamara et al., 2013), and eEF2 is the target for quinoline-4-carboxamide (DDD107498) compounds, both with activity against multiple lifecycle parasite stages (Baragaña et al., 2015). The eEF2 protein mediates GTP-dependent translocation of the ribosome along the mRNA and is required during protein synthesis. The Rab11A protein is likely involved in cytokinesis and interacts with another antimalarial gene-target, the *Pfpi4k* (McNamara et al., 2013). Only one non-synonymous SNP was detected for each of these two genes, supporting their likely essential function. A low number of non-synonymous SNPs (19.6%) was also detected for the mitochondrial cytochrome *b* (*MtcytB*) gene. This gene is the target for several antimalarial compounds under evaluation (Dong et al., 2011) and atovoquone, a longstanding antimalarial drug used in combination with proguanil in Malarone™ for the curative and prophylactic treatment of malaria.

The *Pfpi4k* gene has the highest percentage of non-synonymous

Table 1
Drug targets and genetic polymorphisms.

Gene-target ID	Gene	Active Compounds	Drug development stage	SNP (non-synonymous)	Non-reference allele frequency >5% (non-synonymous)	Known antimalarial resistant mutations
PfcPheRs (PF3D7_0109800)	Phenylalanine-tRNA ligase alpha subunit	Bicyclic azetidine	Nonclinical development	44 (28)	5 (3)	L550V
Pfcarl (PF3D7_0321900)	Cyclic amine resistance locus protein	Imidazolopiperazines, benzimidazolyl piperidines	Clinical trials	152 (90)	34 (21)	0
Pfpi4k (PF3D7_0509800)	Phosphatidylinositol-4 kinase	Imidazopyrazines, aminopyridine class, quinoxaline, 2-aminopyradines	Clinical trials	182 (129)	62 (49)	0
Pfdhodh (PF3D7_0603300)	Ddihydroorotate dehydrogenase	triazolopyrimidine-based inhibitor, N-alkyl-5-thiophene-2-carboxamides	Clinical trials	50 (26)	8 (3)	0
Pfact (PF3D7_1036800)	Acetyl-CoA transporter	Imidazolopiperazines	Clinical trials	46 (22)	13 (6)	0
Pfugt (PF3D7_1113300)	UDP-galactose transporter	Imidazolopiperazines	Clinical trials	30 (12)	10 (6)	0
Pfatp4 (PF3D7_1211900)	P-type cation transporting ATPase	Spiroindolones, sulfonamide, carboxamide, pyrazoles,dihydroisoquinolones	Clinical trials	123 (75)	34 (23)	0
Pfprs (PF3D7_1213800)	Proline-tRNA synthetase	Febrifugine and derivates	Nonclinical development	66 (31)	16 (9)	0
Pfrab11A (PF3D7_1320600)	Ras-related protein Rab-11A	Aminopyridine class	Clinical trials	5 (1)	0 (0)	0
PfeEF2 (PF3D7_1451100)	Elongation factor 2	Quinoline-4-carboxamide (DDD107498)	Preclinical development	34 (1)	4 (0)	0
PfCYTB (mal_mito_3)	Cytocrome b	Atovaquone, tetracyclic benzothiazepine, benzylsulfonamide, decoquinate	Clinical drug, clinical trial	46 (9)	7 (2)	0