Gadalla, Nahla B; Malmberg, Maja; Adam, Ishag; Oguike, Mary C; Beshir, Khalid; Elzaki, Salah-Eldin; Mukhtar, Izdihar; Gadalla, Amal A; Warhurst, David C; Ngasala, Billy; +4 more... Mårtensson, Andreas; El-Sayed, Badria B; Gil, J Pedro; Sutherland, Colin J; (2014) Alternatively spliced transcripts and novel pseudogenes of the Plasmodium falciparum resistance-associated locus pfcrt detected in East African malaria patients. The Journal of antimicrobial chemotherapy, 70 (1). pp. 116-123. ISSN 0305-7453 DOI: https://doi.org/10.1093/jac/dku358

Downloaded from: http://researchonline.lshtm.ac.uk/id/eprint/1987648/

DOI: https://doi.org/10.1093/jac/dku358

Usage Guidelines:

Please refer to usage guidelines at https://researchonline.lshtm.ac.uk/policies.html or alternatively contact researchonline@lshtm.ac.uk.

Available under license: http://creativecommons.org/licenses/by-nc-nd/2.5/
Alternatively spliced transcripts and novel pseudogenes of the *Plasmodium falciparum* resistance-associated locus *pfcrt* detected in East African malaria patients

Nahla B. Gadalla1,2†, Maja Malmberg3‡, Ishag Adam4, Mary C. Oguike1, Khalid Beshir1, Salah-Eldin Elzaki2, Izdihar Mukhtar3, Amal A. Gadalla2, David C. Warhurst6, Billy Ngasala7, Andreas Mårtensson3,8, Badria B. El-Sayed2, J. Pedro Gil9–11 and Colin J. Sutherland1*

1Department of Immunology and Infection, London School of Hygiene and Tropical Medicine, London, UK; 2Department of Epidemiology, Tropical Medicine Research Institute, Khartoum, Sudan; 3Malaria Research, Infectious Diseases Unit, Department of Medicine Solna, Karolinska Institutet, Stockholm, Sweden; 4Faculty of Medicine, University of Khartoum, Khartoum, Sudan; 5National Health Laboratory, Federal Ministry of Health, Khartoum, Sudan; 6Department of Pathogen Molecular Biology, London School of Hygiene and Tropical Medicine, London, UK; 7Department of Parasitology, Muhimbili University of Health and Allied Sciences, Dar es Salaam, Tanzania; 8Global Health (IHCAR), Department of Public Health Sciences, Karolinska Institutet, Stockholm, Sweden; 9Drug Resistance Unit, Division of Pharmacogenetics, Department of Physiology and Pharmacology, Karolinska Institutet, Stockholm, Sweden; 10Drug Resistance and Pharmacogenetics, Center for Biodiversity, Functional and Integrative Genomics, Faculdade de Ciências, Universidade de Lisboa, Lisbon, Portugal; 11The Harpur College of Arts and Sciences, Binghamton University, The State University of New York, Binghamton, NY, USA

*Corresponding author. Tel: +44-(0)20-7927-2338; Fax: +44-(0)20-7636-8739; E-mail: colin.sutherland@lshtm.ac.uk
†Present address: National Institute of Allergy and Infectious Diseases, National Institutes of Health, 12735 Twinbrook Parkway, Rockville, MD 20854, USA.
‡Present address: Department of Biomedical Sciences and Veterinary Public Health, Section of Virology, Swedish University of Agricultural Sciences, Uppsala, Sweden.

Received 29 March 2014; returned 23 May 2014; revised 21 July 2014; accepted 14 August 2014

**Objectives:** Polymorphisms in the lysosomal transporter encoded by the *pfcrt* gene directly impact on *Plasmodium falciparum* susceptibility to aminoquinolines. The Lys76Thr mutation is the critical change conferring chloroquine resistance *in vitro* and *in vivo*, but always occurs with additional non-synonymous changes in the *pfcrt* coding sequence. We sought to better describe *pfcrt* polymorphisms distal to codon 76.

**Methods:** Small-volume samples (≤500 µL) of parasite-infected blood collected directly from malaria patients presenting for treatment in Sudan and Tanzania were immediately preserved for RNA extraction. The *pfcrt* locus was amplified from cDNA preparations by nested PCR and sequenced directly to derive full-length mRNA sequences.

**Results:** In one of two sites in Sudan, two patients were found with an unorthodox spliced form of *pfcrt* mRNA in which two exons were skipped, but it was not possible to test for the presence of the putative protein products of these aberrant transcripts. Genomic DNA sequencing from dried blood spots collected in parallel confirmed the presence of spliced *pfcrt* pseudogenes in a minority of parasite isolates. Full-length cDNA from conventionally spliced mRNA molecules in all study sites demonstrated the existence of a variety of *pfcrt* haplotypes in East Africa, and thus provides evidence of intragenic recombination.

**Conclusions:** The presence of pseudogenes, although unlikely to have any direct public health impact, may confound results obtained from simple genotyping methods that consider only codon 76 and the adjacent residues of *pfcrt*.

**Keywords:** malaria, chloroquine resistance transporters, alternative splicing

© The Author 2014. Published by Oxford University Press on behalf of the British Society for Antimicrobial Chemotherapy. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com
Introduction

Resistance of the malaria parasite *Plasmodium falciparum* to antimalarial drugs has posed a serious threat to public health in Africa in the past, and may do so again in the future. Molecular markers of resistance to chloroquine, pyrimethamine and sulfadoxine were identified and validated some years after the efficacy of these drugs had waned; the utility of markers was thus mainly in gaining understanding of events that had already happened. In the context of artemisinin-based combination therapy (ACT) being the mainstay of uncomplicated malaria management worldwide, malaria researchers have recently focused on the identification of candidate markers for diminished efficacy of artemisinin and of the partner drugs used in ACT. Such markers are needed as molecular tools for surveillance for the purpose of preventing the global spread of resistant parasites from foci of origin, and this objective has become more urgent since the recent description in Cambodia and neighbouring countries of parasites with reduced susceptibility to artemisinin compounds.

It is now well established that ACT selects for particular alleles of two genes, *pfmdr1* and *pfcrt*, both firmly associated with resistance to the 4-aminoquinolines chloroquine and amodiaquine. In a number of African studies of ACT efficacy in symptomatic children and adults, the aminoquinoline-susceptible haplotypes of these genes were significantly more common in recurrent parasitaemia detected during follow-up after treatment with the most commonly used ACT, artemether/lumefantrine.

Encoding a transporter located in the parasite’s digestive vacuole membrane, *pfcrt* is an important determinant of susceptibility to antimalarial chemotherapy. Variant genes encoding a threonine instead of a lysine at codon 76 (the K76T mutation) are highly resistant to antimalarial chemotherapy. Variant genes encoding a threho-
product and an amplicon of \(~\leq 500\) bp. In sample 24 this band was more intense than the 827 bp product, while in sample 18 both bands were of moderate intensity (Figure 2a). Direct sequencing of the \(~\leq 500\) bp PCR product for sample 24 revealed that, for the majority of cDNA molecules amplified, exons 3 and 4 were completely absent, resulting in a truncated cDNA sequence (521 bp; Figure 1b and Figure 2b). This transcript, if translated, encodes a putative 322 amino acid hypothetical protein, which would be...
102 amino acids shorter than the canonical PfCRT encoded by the 3D7 clone reference (PF3D7_0709000, www.PlasmidB.org). It is interesting to note that exons 2 and 5 are both type 0 exons, which means that splicing occurs exactly on codon triplet boundaries (i.e. at the third nucleotide), and so the removal of both exons 3 and 4 does not change the reading phase of the codons. Figure 1(b) illustrates the predicted structure of the full-length spliced cDNA. The sequence of pfcrt mRNA in all remaining patient samples from Gedarif and Fukayosi displayed the expected spliced form. This suggests that the splice variant is not always present.

Intron deletion in genomic copies of pfcrt

In order to investigate whether the observed splicing reflects rearrangements at the genomic level, and to address whether pfcrt introns are structured differently in these samples, DNA was extracted from glass fibre membranes by the Chelex method, which does not preserve RNA, from samples 18 and 24, and from three additional randomly selected samples for comparison, 22, 5 and 75 from New Halfa. These were subjected to amplification and sequencing, which was successful only for the 3′ region of the pfcrt gene, coding for the protein carboxy-terminus. Interestingly, amplification produced two PCR products of unexpected sizes: a 600 bp product and an ≏450 bp product in samples 18 and 24, and also in one of the comparators, 22.

For each of these isolates, the two PCR products were excised from agarose gels following electrophoresis, purified and sequenced. Four primers (NG2_FOR, P11, P13 and P17) located in this region were employed for sequencing these products. Direct sequencing of the 600 bp PCR product revealed complete removal of introns, with the analysed partial genomic sequence found to be identical to the predicted mRNA sequence. Figure 3 shows the partial PfCRT amino acid sequence encoded by an intronless gDNA segment amplified from samples 18 and 24. Interpretable sequence data were not obtained from the 450 bp amplicon in any of these isolates. Thus some P. falciparum isolates from New Halfa harbour additional genome copies of pfcrt that lack introns. We were not able to demonstrate that these unusual genomic pfcrt loci were the origin of the unusual transcripts identified in the previous cDNA analysis, and these may be inactive, processed pseudogenes that are not expressed. Alternatively, these may be transcribed by blood stage parasites, giving rise to the unconventionally spliced transcripts observed.

Haplotypes deduced from pfcrt transcript sequences

cDNA sequence data were used to construct extended haplotypes for the pfcrt locus in each P. falciparum isolate that exhibited the wild-type splicing pattern. Two new polymorphisms were observed in a single isolate from Sudan: the non-synonymous nucleotide change C28A generating a Gln to Lys mutation at codon 10 (Q10K), and the nucleotide change G70T resulting in an Asp to Tyr mutation at codon 24 (D24Y). This codon 24 polymorphism was also observed in three isolates from Tanzania. Using the single-letter amino acid nomenclature, the following non-synonymous pfcrt mutations were observed in addition to codons 10 and 24: M74I, N75E, K76T, A220S, Q271E, N326S and R371I. (In the following paragraphs, only these variant positions are considered.)

Among the 23 Sudanese isolates where only the 5′-region of the cDNA was fully sequenced, the wild-type haplotype MNKA (at codons 74–76 and 220) was present in 4 (17.4%), and the remaining 19 isolates carried the mutant IETS haplotype, associated with chloroquine resistance in Africa. Of the 26 isolates for which full-length transcript sequences were obtained, chloroquine-resistant haplotypes dominated, with five distinct
full-length haplotypes (Figure 3). A single isolate displayed the reference (wild-type) haplotype MNKQAQR at codons 74–76, 220, 271, 326 and 371. Importantly, two individuals in Sudan harboured parasites with hybrid haplotypes, IETSQNR and IETSQNI, indicative of intragenic recombination within the pfcrt locus in this population.

Among the 13 isolates from Fukayosi, Tanzania, for which only the 5′-region was sequenced, the wild-type MNKA at codons 74–76 and 220 was the predominant allele, being present in 7 isolates (53.8%), with 3 (23.1%) carrying the IETS haplotype. Three patients (23.1%) had a mixture of both the IET and MNK, all of whom carried the mutant 220S. A total of 35 isolates were successfully sequenced for the full-length transcript (Figure 3). Two hybrid haplotypes, IETSQNR and MNKAQNI at codons 74–76, 220, 271, 326 and 371, were detected in a single Tanzanian isolate each. Thus in both Sudan and Tanzania, recombinant pfcrt alleles comprising sequence elements of both chloroquine-resistant and wild-type pfcrt were observed. The prevalence of each haplotype observed in both Sudan and Tanzania is presented in Figure 4.

Discussion

The CRT membrane transporter is an important modulator of parasite response to antimalarial chemotherapy, and yet few studies have addressed the structure or abundance of pfcrt transcripts in natural P. falciparum infections. Here, we sequenced cDNA amplified from pfcrt mRNA present in P. falciparum-infected peripheral blood samples taken prior to treatment from malaria patients in eastern Sudan and coastal Tanzania. This approach provided evidence that, in vivo, peripheral blood parasites express alternatively spliced pfcrt mRNA, the first such demonstration either in vivo or in vitro, and that naturally occurring parasite genomes can harbour additional pfcrt copies in the form of spliced pseudogenes. Although these extra gene copies may be the source of the unorthodox pfcrt copies we detected, additional studies are needed to confirm a direct relationship. Further, we determined extended pfcrt haplotypes, identifying new polymorphisms at codons 10 and 24, and providing the first evidence that chloroquine-resistant and chloroquine-susceptible forms of pfcrt recombine in natural parasite populations.

The unusual variant transcript of pfcrt described here was detected in clinical isolates from only one of two study sites in eastern Sudan (n=49) and Tanzania (n=40). Number of observations of each haplotype defined at codons 74–76 and 220 (amino-terminus and full-length data) and 271, 326 and 371 (full-length data only). Data from isolates with multiple genotypes at the pfcrt locus were not included.
As exons 2 and 5 are spliced precisely in-frame with codon boundaries (known as type 0 splicing), the downstream sequence frame is maintained, such that the alternative splicing event can generate mRNA molecules that encode in-frame putative polypeptides. This also raised the question as to whether the alternative spliced format of \textit{pfcrt} and the full-length mRNA were both being expressed from the established \textit{pfcrt} locus on chromosome 7 (exemplified by PF3D7_0709000).

Alternative splicing is an important mechanism by which precursor mRNAs are transcribed into several forms of mature mRNAs, leading to multiple polypeptides encoded by a single gene, and the generation of plurifunctional variant protein isoforms. The first report of a spliced variant mRNA in \textit{P. falciparum} was the 41-3 gene encoding a variable antigen involved in immune regulation and sought as a vaccine candidate. Further examples are the guanylyl cyclase gene involved in the signalling pathway of male gametogenesis, the erythrocyte-binding protein MAEBL and a CDK-related protein kinase. Recent transcriptomic studies have identified further variable splicing events for various \textit{P. falciparum} genes in laboratory-cultured clones such as 3D7 and NF54, but there have been no reports of alternative splicing in the \textit{pfcrt} gene in laboratory lines or in vivo. Interestingly, Otto et al. described four new splicing events on the \textit{pfcrt} locus, but none of these encompassed the \textit{pfcrt} locus.

In attempting to understand the genomic origin of the variant cDNA sequences we observed, which corresponded to the 5' half of the normal \textit{pfcrt} mRNA, the 3' half of an additional genomic \textit{pfcrt} locus in which all introns were missing was identified in parasite gDNA amplified from dried filter-paper blood spots. To our knowledge, this is the first description of pseudogenes derived from drug resistance-associated functional genes in \textit{Plasmodium}. A plausible explanation for our observations, which we were unable to confirm with the material available, is that the intron-less \textit{pfcrt} locus we have described is a classic pseudogene that has arisen in this parasite population, and is transcribed, as \textit{P. falciparum} pseudogenes frequently are. As such it may have no significant effect on parasite responses to antimalarial drugs, despite being expressed in peripheral blood parasites, although an effect on regulation of transcription from the native \textit{pfcrt} locus cannot be ruled out. The unusual cDNA we sequenced encompassed the region of \textit{pfcrt} around the codon 76 polymorphism, suggesting the same may be true for the additional \textit{pfcrt} pseudogene locus; this is therefore very likely to contribute to genotyping analysis of parasite DNA from peripheral blood. It remains possible that such loci are widespread, and have been part of the 'signal' obtained whenever field samples are genotyped at this locus, whether by PCR-RFLP, qPCR or direct sequencing. It is very important that this possible source of confounding of genotyping studies is investigated thoroughly. Interestingly, the pseudogene encodes the mutant 76Thr allele, which may nevertheless provide a source of the mutant sequence that could be reactivated by recombination should chloroquine pressure be resumed in the future. In this context, it should be noted that evidence suggests that considerable variability was present in the \textit{pfcrt} locus before the era of chloroquine exposure.

\textbf{Figure 5.} A possible linear arrangement of transmembrane regions in the full-length CRT protein (a) and putative truncated CRT protein encoded by alternative-spliced \textit{pfcrt} mRNA from Sudanese malaria patients (b). Transmembrane regions are depicted as rectangles; blue denotes presence in putative truncated form of CRT and white denotes presence only in full-length CRT. Freeform blue line denotes extramembrane polypeptide tracts in the cytosol or inside the digestive vacuole lumen. Transverse yellow bars denote site of splicing event and loss of contiguous nucleotide/amino acid sequence. Polymorphic residues are depicted by yellow or orange (codons 76 and 220) dots. Amino terminal polymorphisms at codons 10 and 24 are not shown.
and Tanzania. Such events may indicate that pfcrt genes are still evolving in African parasite populations, generating further heterogeneity in the structure of the CRT protein, and this diversity is likely to be under active selection by ongoing ACT pressure across the continent.

Conclusions
In this study we describe aberrant spliced pfcrt transcripts and confirm the presence of pfcrt pseudogenes in a small number of patient parasite isolates. We also provide evidence from both Sudan and Tanzania that hybrid forms of pfcrt with sequence characteristics of both chloroquine-resistant and wild-type alleles are present in natural populations, probably due to meiotic recombination. A weakness of our study is that the unusual transcripts and alleles of pfcrt were detected in non-replenishable parasite RNA and DNA, and so confirmatory analyses were not possible. Similarly, no phenotypic analysis of any of the parasite isolates was carried out. However, as the variant alleles of pfcrt were identified in more than one individual, and from both mRNA and gDNA sequences, we are confident that these parasites were present in New Halfa. Further studies of pfcrt cDNA sequences from patient peripheral blood are needed to explore this phenomenon further.

Funding
This work was funded by the EU FP7 MALACTRES Consortium, SIDA-SAREC Grant ref. SWE-2009-165 and EU FP7 NanoMal Consortium Grant ref. 304948. N. B. G. was supported by TDR and C. J. S. is supported by Public Health England.

Transparency declarations
None to declare.

References


