Antimicrobial Agents and Chemotherapy, Apr. 2003, p. 1347–1354

0066-4804/03/$08.00+0 DOI: 10.1128/AAC.47.4.1347–1354.2003

Copyright © 2003, American Society for Microbiology. All Rights Reserved.

Molecular Determination of Point Mutation Haplotypes in the Dihydrofolate Reductase and Dihydropteroate Synthase of Plasmodium falciparum in Three Districts of Northern Tanzania

Richard J. Pearce,1* Chris Drakeley,1,2 Daniel Chandramohan,1 Frank Mosha,2,3 and Cally Roper1

Department of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, WC1E 7HT, United Kingdom,1 and Kilimanjaro Christian Medical Centre3 and The Joint Malaria Programme (JMP)2 Moshi, Tanzania

Received 24 September 2002/Returned for modification 12 November 2002/Accepted 28 January 2003

The antimalarial combination of sulfadoxine and pyrimethamine (SP) was introduced as first-line treatment for uncomplicated malaria in Tanzania during 2001 following 18 years of second-line use. The genetic determinants of in vitro resistance to the two drugs individually are shown to be point mutations at seven sites in the dihydrofolate reductase gene (dhfr) conferring resistance to pyrimethamine and five sites in the dihydropteroate synthase (dhps) gene conferring resistance to sulfadoxine. Different combinations of mutations within each gene confer differing degrees of insensitivity, but information about the frequency with which allelic haplotypes occur has been lacking because of the complicating effects of multiple infection. Here we used a novel high-throughput sequence-specific oligonucleotide probe-based approach to examine the present resistance status of three Plasmodium falciparum populations in northern Tanzania. By using surveys of asymptomatic infections and screening for the presence of all known point mutations in dhfr and dhps genes, we showed that just five dhfr and three dhps allelic haplotypes are present. High frequencies of both triple-mutant dhfr and double-mutant dhps mutant alleles were found in addition to significant interregional heterogeneity in allele frequency. In vivo studies have shown that the cooccurrence of three dhfr mutations and two dhps mutations in an infection prior to treatment is statistically predictive of treatment failure. We have combined data for both loci to determine the frequency of two-locus genotypes. The triple-dhfr/double-dhps genotype is present in all three regions with frequencies ranging between 30 and 63%, indicating that treatment failure rates are likely to be high.

Sulfadoxine-pyrimethamine (SP) has now replaced chloroquine as the first-line curative antimalarial in much of East Africa. The earliest reports of emerging SP resistance in Africa were from Muheza district in Tanzania during 1994 and 1995 (32, 37). SP remains in use throughout Tanzania and was recently officially adopted as the national first-line treatment for nonsevere Plasmodium falciparum malaria. It is now a priority to learn how widespread genetic determinants of SP resistance currently are in the larger regions of northern Tanzania.

To investigate this issue, we have carried out a population-based genetic analysis of P. falciparum in the North and South Pare Mountains and Hai district, which are distinct geographical areas with their own microclimates and tribes. There had been no malaria research in these districts since 1965 until recent work showed relatively low levels of transmission, with an estimated entomological inoculation rate of 24 infective bites/person/year in Hai district (C. Drakeley and D. Chandramohan, unpublished data) compared with an entomological inoculation rate in the range of 34 to 405 infective bites/person/year in Muheza district (14).

It is widely understood that people self-treat with antimalarial drugs, which can be freely purchased (24); as a consequence, a proportion of people attending health facilities with signs and symptoms of malaria may have had recent prior exposure to drugs. In this study we have analyzed material from community surveys of asymptomatic infections. These data, we believe, are less subject to bias due to prior drug selection and are therefore representative of the parasite population at large.

Resistance to SP is associated in vitro with a series of substitutions within the active site of target enzymes of the folate biosynthesis pathway, dihydropteroate synthase (DHPS) (4, 38) and dihydrofolate reductase (DHFR) (9, 29, 36), and this has been demonstrated through laboratory-based in vitro sensitivity tests and transfection experiments with DHFR with respect to pyrimethamine (43, 44) and DHPS with respect to sulfadoxine (39, 40, 42). The sequence changes coding for substitutions that are naturally occurring worldwide are summarized in Table 1. A Ser-to-Asn substitution at codon 108 of DHFR decreases sensitivity to pyrimethamine 100-fold (43). Additional substitutions at codons 51 (N51I), 59 (C59R), and 164 (I164L) progressively increase levels of resistance to pyrimethamine (15, 43). Isolates containing all four substitutions have yet to be reported in Africa (21, 25, 41). A total of 14 DHPS substitutions at five sites have been characterized worldwide, of which 6 have been recorded in Africa, with A437G and

* Corresponding author. Mailing address: Department of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, Keppel St., London WC1E 7HT, United Kingdom. Phone: (44) 207 927 2416. Fax: (44) 207 636 8739. E-mail: Richard.Pearce@lshtm.ac.uk.
The role of the point mutations at each locus in conferring resistance to SP in vivo has been inferred from studies showing predictive association of particular mutations with treatment failure (21, 28) and from overrepresentation of mutations in samples from the North and South Pare Mountains. A mixed infection containing these two double-mutant alleles is less resistant to pyrimethamine than an infection containing the triple-mutant allele, despite all three mutations being present in either case. When comparing populations, it is important to measure the frequency of haplotypes rather than the prevalence of each point mutation separately, because haplotypes are the determinants of resistance levels.

K540E mutations being the most frequently reported (12, 41, 42).

The study area and the samples. P. falciparum-positive samples were collected from two separate studies within the same area of northeast Tanzania; one was a study of infants and young children in Hai district, and the second was a study of people up to 45 years of age in the North and South Pare sites. Both studies were cross-sectional malarialometric surveys across an altitude band of 550 to 1,600 m, and most of the study subjects were asymptomatic.

For the villages in Hai district, samples were collected in May 2001. All children less than 3 years old from 16 randomly selected villages were invited to attend the survey clinic at a central clinic. A finger-prick blood sample for blood slides and a filter paper blood sample were collected from each participating child. The filter paper blood samples were air dried and stored at 4°C with desiccant. Bloodspots from blood film-positive children were selected retrospectively for genotyping.

Samples from the North and South Pare Mountains were collected in November 2001 during malarialometric cross-sectional surveys. A random sample of 1,250 individuals (250 per village) under 45 years of age was recruited, and a blood sample was taken into EDTA Microtainer tubes. Filter paper bloodspots were made with 10 μl of packed cells from samples of individuals found to be parasite positive.

Materials and Methods

The study area and the samples. P. falciparum-positive samples were collected from two separate studies within the same area of northeast Tanzania; one was a study of infants and young children in Hai district, and the second was a study of people up to 45 years of age in the North and South Pare sites. Both studies were cross-sectional malarialometric surveys across an altitude band of 550 to 1,600 m, and most of the study subjects were asymptomatic.

For the villages in Hai district, samples were collected in May 2001. All children less than 3 years old from 16 randomly selected villages were invited to attend the survey clinic at a central clinic. A finger-prick blood sample for blood slides and a filter paper blood sample were collected from each participating child. The filter paper blood samples were air dried and stored at 4°C with desiccant. Bloodspots from blood film-positive children were selected retrospectively for genotyping.

Samples from the North and South Pare Mountains were collected in November 2001 during malarialometric cross-sectional surveys. A random sample of 1,250 individuals (250 per village) under 45 years of age was recruited, and a filter paper blood sample was taken into EDTA Microtainer tubes. Filter paper bloodspots were made with 10 μl of packed cells from samples of individuals found to be parasite positive.

Ethics. Scientific and ethical clearance for both studies was granted from the National Institute of Medical Research in Tanzania and the London School of Hygiene and Tropical Medicine. Consent was obtained from all individuals or their guardians prior to enrollment in the surveys.
DNA extraction. DNA extraction from bloodspots on filter paper was carried out in a 96-well plate format. A segment of the bloodspot was first soaked in 0.5% saponin–1× phosphate-buffered saline overnight and was then washed twice in 1 ml of 1× phosphate-buffered saline. The segment was then boiled for 8 min in 100 μl of PCR-quality water–50 μl of 20% Chelex suspension in distilled water (pH 9.5).

PCR amplification of dhfr and dhps. A 711-bp fragment of dhfr containing the polymorphic codons were independently amplified by nested PCR in a 96-well plate format. PCR primer sequences and reaction conditions are indicated in Table 2. The 25-μl PCR mix contained primers at 0.25 μM final concentration, 2 mM MgCl2, 250 μM each dNTP, and 1× Bioline Taq polymerase. Template DNA (1 μl) was introduced to outer reaction mixtures. The dhfr outer PCR product (1 μl) was introduced into a 25-μl inner amplification mixture. Aliquots of 1 μl of threefold-diluted dhfr outer PCR product were introduced into a 25-μl inner amplification reaction mixture.

Use of SSOP for molecular genotyping of point mutations. Final-round PCR products were heat denatured (95°C for 2 min), cooled, and then spotted onto nylon membranes in 1-μl volumes in a 12 by 8 grid. A panel of four PCR samples of known sequences representing all common sequence variants was spotted on nylon membranes in 1-μl volumes in a 12 by 8 grid. A panel of four PCR samples were each designed to complement the sequence polymorphism at each site separately. Images of blots probed with Dynamics Storm 840 PhosphorImager (Amersham Pharmacia Biotech).

Scoring. We scored the presence, absence, or relative abundance of the variant sequence polymorphism at each site separately. Images of blots probed with variant sequences for a single locus were transferred as TIF files to ImageMaster Total Lab software (Amersham Pharmacia Biotech). In the array analysis subsection of the software, a standard area of each spot was defined and the intensity of chemiluminescence in that area was measured. Background data were adjusted for by subtraction of the volume of the negative controls from the volume data. Thus, the volume of chemiluminescence for each spot was calculated as Volume = (maximum intensity × spot area) − background. To determine the threshold of detection per se, the presence-flagging option was employed. By this method, the faintest spot considered present and not background was selected to set the threshold volume value. To compare SNP-specific probes at a single site, Microsoft Excel bar charts were drawn comparing the volume data for each probe on every sample to the presence-flagging result in each case. The following rules were used to determine whether a SNP was present or absent at each site.

(i) An SNP was considered present in a PCR product when the volume value of a particular probe was higher than that of the background. When volume values were low, presence flagging provided an internal control for avoiding possible biasing between probes or blots.

(ii) An SNP was considered absent when all volume values below the first gridline on the chart were rejected. No set value can be given for this criterion, as volume value comparison is relative and differs depending on the strength of the probe labeling and binding.

RESULTS

Of the 165 bloodspots that yielded PCR products, 10.3% were mixed at dhfr and 1.2% were mixed at dhps; no majority haplotypes were found. The low number of mixed infections was a reflection of the low level of transmission in the area. On stratification of the populations of the North Pare and South Pare Mountains into the age ranges of 0 to 4 and 5 to 45 years, no significant difference was found in the frequencies of dhfr and dhps allelic haplotypes, allowing comparisons to be made between the samples from those sites and those from Hai district, where samples were taken exclusively from subjects <5 years of age.

The dhfr and dhps allelic haplotypes present in each region are shown in Fig. 1. Three point mutations were found in dhfr, and of the eight possible haplotypic conformations, five were found; this matches well with findings of point mutations present in single-genotype infections elsewhere in East Africa that have been described previously (26, 41). Three point mutations were likewise found in dhps, and of the eight possible haplotypes, only three were found in dhps.

Regional haplotype frequencies are summarized in Fig. 2. Not all alleles at dhfr and dhps were present in each of the

### Table 2. Table of PCR primer sequences and reaction conditions for the nested amplification of dhfr and dhps

<table>
<thead>
<tr>
<th>Gene</th>
<th>Primer sequence</th>
<th>PCR conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>dhfr Outer, M1</td>
<td>5′ TTATGATGGAAACAGTCTGC 3′</td>
<td>94°C × 3 min</td>
</tr>
<tr>
<td>650 bp, M7</td>
<td>5′ CTATGATACATACGCTAAC 3′</td>
<td>94°C × 1 min, 52°C × 2 min, 72°C × 1 min, 40×; 72°C × 10 min</td>
</tr>
<tr>
<td>Inner, M3b</td>
<td>5′ TGATGGAACAGTCTGCCAGGT 3′</td>
<td>94°C × 3 min</td>
</tr>
<tr>
<td>594 bp, M9</td>
<td>5′ CTGAAAAATACATACATCATTAG 3′</td>
<td>94°C × 1 min, 44°C × 2 min, 72°C × 1 min, 4×; 94°C × 1 min, 44°C × 1 min, 72°C × 1 min, 34×; 72°C × 10 min</td>
</tr>
<tr>
<td>dhps Outer, N1</td>
<td>5′ GATCTTTTCTACAGATGAG 3′</td>
<td>94°C × 3 min</td>
</tr>
<tr>
<td>770 bp, N2</td>
<td>5′ TTCTCTATGATTACATCG 3′</td>
<td>94°C × 1 min, 51°C × 2 min, 72°C × 1 min, 40×; 72°C × 10 min</td>
</tr>
<tr>
<td>Inner, R2</td>
<td>5′ AACCTAAAACGCTGTGCT 3′</td>
<td>As described above for dhfr outer primer sequence</td>
</tr>
<tr>
<td>711 bp, R</td>
<td>5′ AATTGTGATTTGTCCACA 3′</td>
<td></td>
</tr>
</tbody>
</table>

Statistical analysis. Statistical analysis of population differences in haplotype frequencies (Wright's F<sub>ST</sub> [fixation index]) (6) and linkage disequilibrium (LD), given as D' coefficient, was carried out using Arlequin software (33). Statistical analysis of contingency tables of the association of haplotypes within two-locus combinations was performed using a χ² test.
There was no significant difference in the remaining pairwise comparisons, despite the differences in haplotype frequencies between the North and South Pare Mountains.

Two-locus combinations were derived from single infections for which only one haplotype was recorded in both dhfr and dhps. The map in Fig. 3 shows the spatial distribution of frequencies of two-locus combinations. The reduction in sample size (n) reflects the loss due to mixed infections. The frequency of the most highly resistant genotype triple-mutant dhfr/double mutant dhps (ACIRNV1-SGEAA; substitutions are underlined) was found to be extremely high. The frequency was 63.2% in Hai district, 50% in the South Pare Mountains, and 22% in the North Pare Mountains. In the North Pare Mountains, there was more diversity at both loci and consequently a greater diversity of 2 locus genotypes. The Wright’s Fst comparing the three populations at both loci further confirms the differences between the North and South Pare populations (Fst = 0.0583 [P = 0.0054]) and between the North Pare Mountains and Hai district (Fst = 0.149 [P < 0.0001]). Pairwise comparison of the South Pare and Hai district populations (Fst = 0.011) showed no significant difference and were merged for subsequent analysis. Statistical analysis of observed and expected two-locus combinations was performed on the population of the North Pare Mountains and the combined populations of Hai district and the South Pare Mountains. We found a significant departure from expected in the merged Hai-South Pare Mountains population (χ² test [P = 0.0018, 6 df]), whereas the distribution in North Pare was nonsignificant (χ² test [P = 0.835, 8 df]). LD analysis was performed on the combined Hai and South Pare Mountain data set, and we found three two-locus combinations to be in LD, namely, ACICNV1-SAKAA (D’ = 0.277, P = 0.01), ACNRNVI-AA KAA (D’ = 1.0, P < 0.0001), and ACIRNV1-SGEAA (D’ = 0.229, P = 0.031) (substitutions are underlined). No other pair of alleles was found to be in LD.

**DISCUSSION**

We have described a new approach by which blood survey material was used to determine the frequency of point mutation haplotypes in *P. falciparum* populations. This approach allows quantitation of resistance at the population level and enables direct comparison of population resistance levels even when they differ widely in the proportions of multiply infected individuals. The issue of multiple infections can be problematic when genotyping blood-stage parasites, because it causes haplotype conformations of point mutations to be obscured and rare mutations to be oversampled. By recording one genotype per infection and discounting minority genotypes, we were able to avoid the oversampling of rare genotypes and to estimate the frequency of mutation haplotypes in the population in a manner which was standardized over all populations of different transmission intensities. With this consideration in mind, the SSOP method employed is designed for high-throughput screening of blood-stage infections to derive haplotype frequencies from survey material.

We have shown that only a subset of point mutations reported globally were present in North Tanzania in these populations, confirming previous reports of analyses of natural *P. falciparum* populations of East Africa (26, 41). Using single-
and majority-genotype infections, we were able to determine the point mutation haplotypes that were present in the three districts surveyed. We found five alleles at *dhfr*, and in reviewing the single-infection data from other studies in Africa, it is possible to see these same haplotypes in Kenya, Malawi, Tanzania, and Mauritania. We found three haplotypes in *dhps*, the sensitive-allele haplotype, the single-mutant allele S436A haplotype, and the double-mutant allele A437G K540E haplotype which has been widely recorded in East Africa (18, 19, 21, 25–27, 41) and not in West Africa (12, 30, 41).

**Regional variation in population resistance.** Allelic haplotype frequencies in the three regions differed significantly. Those of the Hai district and South Pare populations were highly resistant at *dhfr*, with triple-mutant allele (N51I+C59R+S108N) frequencies of >70% and *dhps* double-mutant allele (A437G+K540E) frequencies of >50%.

![](image.png)
The same alleles were present in the North Pare population but at lower frequencies. Sensitive alleles at either locus were absent in the Hai district and South Pare populations but were present at low frequency in the North Pare population. The North Pare population was significantly distinct from the other two populations, with the most significant difference found between those of the North Pare and Hai district populations. These findings point to a slower development of SP resistance in the North Pare Mountains. Possible causes of weaker selective pressure are greater use of alternative antimalarials such as amodiaquine and reduced use of antimalarials per se due to poor access to healthcare facilities or raised levels of acquired immunity. Such differences are maintained in the face of gene flow, which operates to make the parasite populations more homogeneous. Ongoing investigations into treatment-seeking behavior and overprescription by providers aim to further illuminate differences in drug use among the three study populations (T. Swarthout, D. Chandramohan, F. Mosha, A. Bell, G. Masuki, C. Drakeley, and H. Reyburn, Abstr. 3rd MIM Pan-African Malaria Conf., 2002).

Recent work in real populations in Malawi has found statistical association of the presence of all three dhfr mutations and both dhps mutations with failure to clear parasitemia after SP treatment (21). It is probable that in many cases this is due to the presence of the two most highly resistant alleles, triple dhfr and the double dhps—and this idea is supported by analysis of recrudescence infections following SP treatment (22). Here we directly measured the frequency of the highly resistant two-locus genotype consisting of the ACIRNVI-SGEAA (substitutions are underlined). We found that frequencies were twice as high in Hai district as in the South Pare Mountains and three times as high as in North Pare Mountains; hence, there is a clear and testable indication that SP treatment failure rates in these three regions are likely to differ.
Selection on dhfr and dhps by SP use. It has been observed that dhfr mutations appear earlier than dhps mutations in the development of SP resistance within populations in Africa (23, 26, 34). Our own data suggest the same sequence of events, with resistant dhfr being fixed in Hai and the South Pare Mountains yet with sensitive dhps being relatively common. The observed LD between the ACICNVI-SAKAA and the absence of significant LD between the dhfr double mutants and the dhps double-mutant alleles supports the idea of an interaction between dhfr and dhps in the development of resistance.

Furthermore, there was a statistically significant association between the triple dhfr mutant and the double dhps mutant in the combined Hai district and the South Pare populations. dhfr is on chromosome 4 and dhps is on chromosome 8, so the fact that LD was found between two highly resistant haplotypes at these unlinked loci is indicative of the nonindependence of SP selection. This relationship has been reported previously in Kenya (28) and Malawi (21). The same relationship was not apparent in North Pare, where the frequency of resistance alleles at both loci was also significantly lower; both this and the absence of statistically significant LD imply that drug selection is weaker. This finding emphasizes the transient nature of linkage, particularly in areas of high transmission intensity, where recombination rapidly breaks down the linkage between dhfr and dhps, and argues for caution in the use of indicator mutations as a proxy for resistance genotyping. While the cooccurrence of all five mutations in an infection in the South Pare Mountains and Hai district was in fact a reliable indicator of the presence of a quintuple genotype, this was not the case in the North Pare Mountains. The widespread use of the antibiotic Septrin, which contains trimethoprim and sulfamethoxazole, to treat other infections may indirectly select for dhfr and dhps resistance mutations and further complicate the relationship.

Past, present and future of pyrimethamine and sulfadoxine use in northern Tanzania. The three districts described here are found in an area between 200 km and 400 km northwest of Muheza district, a region historically associated with exceptionally high levels of antimalarial drug resistance. Resistance to pyrimethamine alone was reported in 1954 in Mheza in Muheza district, following mass administration of prophylactic doses of pyrimethamine monotherapy over a five-month period during 1953 (5). Use of the SP combination began in Tanzania in 1982, when it was introduced as a second-line treatment for use in cases of chloroquine failure, and as early as 1984 it was the policy of the Muheza district hospital to use SP as a first-line antimalarial (25). Studies from the region report that SP was highly effective during the eighties, but resistance was recorded in Magoda village near Muheza in 1994 (32) and was subsequently reported in villages in the surrounding area (18, 19, 37). The emergence of resistance to SP in 1994 was attributed by the authors to be in part a result of the prophylactic intervention with weekly dapsone pyrimethamine treatment to all children less than 10 years old. Resistance to SP in Muheza district hospital is now reported to be as high as 45% (25).

Two explanations for the high frequency of resistance alleles in the three districts described here are the widespread use of SP or related drugs and the movement of resistance from Muheza. However, levels of resistance do not show a simple decline with distance from Muheza to Hai (Fig. 3), and drug use is clearly a very important factor. What is striking from the data presented here are the significant interpopulation differences, and it is probable that these differences have arisen from differing patterns of drug use in these communities.

In this paper we have described a new technique for detection of resistance point mutations in dhfr and dhps. We have argued for the use of haplotype frequencies as a measure of resistance at the population level and as an appropriate tool for describing the spread of resistance. In displaying the power of this technique, we have shown significant regional heterogeneity and identified a region of northern Tanzania where resistance to SP is already unprecedentedly high. Since policy recently mandated a switch to use of this drug as the first-line treatment, we recommend that in vivo studies be performed to confirm that the high frequency of resistance alleles is indicative of treatment failure, as is predicted on the basis of studies conducted elsewhere in Africa.

ACKNOWLEDGMENTS

Field studies in Hai were funded by a DfID grant (R7950 to D. Chandramohan et al.), and those in the Pare mountains were funded by a grant from MRC UK (G0901439 to B. M. Greenwood, E. M. Riley, et al.). C.R. and R.J.P. are supported by an Advanced Training Fellowship from The Wellcome Trust (ref. 060714) awarded to C. Roper. We are grateful for logistical support from W. M. M. Nkya and H. Reyburn and for technical support from E. Nyale, F. Laizer, and M. Mvula.

This study was conducted under the auspices of the Joint Malaria Programme, a collaborative research initiative between The London School of Hygiene and Tropical Medicine, The Tanzanian National Institute for Medical Research, The Kilimanjaro Christian Medical College, and the Centre for Medical Parasitology, University of Copenhagen.

REFERENCES


sation of the dihydrofolate reductase-thymidylate synthetase gene from hu-

man malaria parasites highly resistant to pyrimethamine. Gene 76:41–52.


sitol. 36:253–262.