Origin and Dissemination across the Colombian Andes Mountain Range of Sulfadoxine-Pyrimethamine Resistance in Plasmodium falciparum

Vladimir Corredor,1,7* Claribel Murillo,2 Diego F. Echeverry,2 Julie Benavides,2,3 Richard J. Pearce,4 Cally Roper,4 Angela P. Guerra,5 and Lyda Osorio2,6

Unidad de Parasitología, Departamento de Salud Pública, Facultad de Medicina, Universidad Nacional de Colombia, Bogotá, Colombia; Centro Internacional de Entrenamiento e Investigaciones Médicas—CIDEIM, Cali, Colombia; Departamento de Biología, Universidad de Nariño, Pasto, Colombia; London School of Hygiene and Tropical Medicine, Department of Infectious Tropical Diseases, London, United Kingdom; Grupo Bioquímica y Biología Celular, Instituto Nacional de Salud, Bogotá, Colombia; Grupo de Epidemiología y Salud Poblacional (GESP), Escuela de Salud Pública, Facultad de Salud, Universidad del Valle, Cali, Colombia; and Emory Vaccine Center, Yerkes National Primate Research Center, Emory University, Atlanta Georgia*

Received 9 January 2010/Returned for modification 17 February 2010/Accepted 14 May 2010

The therapeutic efficacy of sulfadoxine-pyrimethamine (SP) in treating uncomplicated Plasmodium falciparum malaria is unevenly distributed in Colombia. The Andes mountain range separates regions in the west where malaria is endemic from those in the east and constitutes a barrier against gene flow and the dispersal of parasite populations. The distribution of dhfr and dhps genotypes of 146 P. falciparum samples from the eastern Amazon and Orinoco basins and Northwest and Southwest Pacific regions of Colombia was consistent with the documented levels of therapeutic efficacy of SP. The diversity of four dhfr- and dhps-linked microsatellites indicated that double- and triple-mutant alleles for both resistance loci have a single origin. Likewise, multilocus association genotypes, including two unlinked microsatellite loci, suggested that genetic exchanges between the eastern Orinoco and Northwest Pacific populations has taken place across the Andes, most probably via migration of infected people.

Enzymes involved in folate metabolism are targeted by the antifolate antimalarial drugs. Pyrimethamine targets the enzyme dihydrofolate reductase (DHFR) and, in combination with sulfadoxine, which targets the enzyme dihydropteroate synthase (DHPS), has been widely used as first-line treatment for uncomplicated Plasmodium falciparum malaria worldwide. In South America, pyrimethamine was introduced and used as a mass treatment in the 1950s in Venezuela (10) until drug-resistant cases were detected (16). By 1968, pyrimethamine-resistant/sulfadiazine-sensitive parasites were documented in Brazil, Venezuela, and Colombia (28). The sulfadoxine-pyrimethamine (SP) combination was introduced in parts of South America in the 1970s and was used until 1981 in Colombia as an alternative to chloroquine, to which resistance was widespread (8). Soon after its introduction, treatment failure was reported, and resistance rapidly disseminated in the Amazon and Orinoco basins (12). In Colombia, SP resistance is unevenly distributed. High SP resistance levels (80%) have been consistently reported from the Colombian Amazon basin (21), while moderate levels (6% to 24%) are observed in the Caribbean, the Cauca Valley, and northwestern regions (3, 4).

This contrasts with regions on the southern Pacific coast where the SP combination is still efficacious (13).

P. falciparum resistance to pyrimethamine is acquired by the progressive accumulation of mutations at the dhfr locus. The S108N substitution is initially required for the acquisition of the resistant phenotype; in this genetic context, resistance increases with the accumulation of additional mutations at position N51I or C59R, and still higher levels of resistance are reached with the further acquisition of mutations at positions C50R and I164L (7, 14, 15). Resistance to sulfadoxine also depends on the progressive accumulation of mutations in the dhps locus at codons 436, 437, 540, 581, and 613. The A437G mutation is initially required and is followed by mutations at codons A581G, S436A, K540E, and A613S, which confer incrementally higher levels of resistance (14).

The current pattern of P. falciparum SP resistance in Colombia raises questions as to the nature, frequency, and origin of the circulating dhfr and dhps genotypes. Analysis of microsatellite markers linked to dhfr and dhps has shown that some mutant resistance alleles in Asia and Africa have become globally dispersed (27), while unlinked polymorphic loci on different chromosomes have suggested that the clonal expansion of a few resistant parasite genotypes underlies the spread of drug resistance in South America (6). Interestingly, in Colombia, the Andes mountain range separates the Pacific coast in the west from the Amazon and Orinoco basins in the east and may constitute a barrier against genetic exchanges. Here, we analyze the genetic nature and origin of P. falciparum SP resistance in Colombia. The results will help to the design strat-

* Corresponding author. Mailing address: Unidad de Parasitología, Departamento de Salud Pública, Facultad de Medicina, Universidad Nacional de Colombia, Bogotá, Colombia. Phone: (57)(1) 316-5000, ext. 15033. Fax: (57)(1) 316-5000, ext. 15073. E-mail: vcorredore@unal.edu.co.
† Supplemental material for this article may be found at http://aac.asm.org/.
‡ Published ahead of print on 24 May 2010.
ergies to prevent the spread of drug resistance in Colombia and eventually other Andean countries.

MATERIALS AND METHODS

P. falciparum samples. Blood samples from subjects with thick-smear-confirm- uncompliated P. falciparum malaria who took part in epidemiological studies carried out by Centro Internacional de Entrenamiento e Investigaciones Médicas (CIDEIM) and Instituto Nacional de Salud (INS) during the years 1999 to 2007 were stored on filter paper. A total of 146 samples representing four geographical regions of Colombia were selected: (i) Northwestern Pacific coast, Quibdó, Department of Chocó (n = 30; lat 5.694722, long −76.661111), (ii) Southwestern Pacific coast, Tumaco, Department of Nariño (n = 30; lat 1.798611, long −78.815556); (iii) Amazon basin, Department of Amazonas (n = 43, distributed as n = 14 from Leticia, lat −4.215278, long −69.940556; n = 16 from Tarapacá, lat −2.876811, long −69.74167; n = 3 from La Pedrera, lat −1.3002778, long −69.5638889; and n = 10 from an unknown locality in Amazonas Department); and (iv) Orinoco basin, Departments of Meta and Guaviare (n = 31 from Meta, lat 3.01236, long −73.154; and n = 12 from Guaviare, lat 2.570833, long −72.640278) (see Table S1 in the supplemental material). This study was approved by the Ethical Review Committee of CIDEIM.

DNA extraction, dhfr and dhps genotyping, sequencing, and microsatellite analysis. Genomic DNA for PCR-restriction fragment length polymorphism (RFLP) analysis from P. falciparum samples stored in filter paper was extracted and sequenced as described previously (29). Polymorphic positions in the dhfr (positions 51, 59, 108, and 164) and dhps (positions 437, 540, and 581) genes were genotyped as described previously (25; http://medschoo.UMARYLAND.edu/CVD/2002_pcr_astr.asp). To confirm the restriction enzyme analysis, 11 samples were sequenced for dhfr, 8 for dhps, and 10 for both genes after nested PCRs as described previously (23, 26). Genotyping of 6 microsatellite loci was performed in all samples: 2 dhfr-linked microsatellite loci (mDHFR and MA1, located 0.3 kb and 5.3 kb upstream, respectively, from codon 108 of the dhfr gene in chromosome 4 [chr4]), 2 dhps-linked microsatellites (m0.8 and m4.3, located 0.8 and 4.3 kb downstream, respectively, of the gene in chr8) (26), and 2 unlinked microsatellites (PPK2 and Polys), located in chr12 and chr4, respectively (1). The microsatellites were amplified using a semi-nested PCR, and the products were analyzed as previously described (26). When more than one allele was present at a locus, the allele with the higher peak was scored if the minority peak was less than 50% of the majority peak. If the minority peak exceeded the 50% cutoff value, the sample was not taken into account for the locus. Twenty-two samples out of 147 showed mixed infections, 14 of which were discarded for a given locus.

Resistance alleles, resistance genotypes, allelic haplotypes, and multilocus associations. Combinations of dhfr and dhps resistance alleles were defined as the resistance genotype. Allelic haplotypes comprised dhfr and its linked microsatellite alleles (at mDHFR and MA1 microsatellite loci) or dhps and its linked microsatellite alleles (at m0.8 and m4.3 microsatellite loci). When linked microsatellites were combined with unlinked microsatellites, this was defined as the multilocus association genotype. Different multilocus association genotypes representing various degrees of linkage were taken into account as follows: (i) complete multilocus association genotype (i.e., allelic linkage at all six microsatellite loci) and (ii) partial multilocus associations (i.e., allelic linkage in two of the three above-mentioned pairs of microsatellite loci). Haplotypes and multilocus association genotypes were coded with numbers, and there are listed in full in Table S1 in the supplemental material.

Statistical analysis, linkage disequilibrium, and population differentiation. Allelic, haplotypic, and multilocus association genotype diversity was calculated as Nei’s unbiased gene diversity estimate (19) as implemented in the Microsatellite toolkit (22). Statistical analysis of population differentiation was performed using Arlequin 3.1 software (9). Fst indexes (Wright’s Fs index) were calculated for neutrally evolving, nonlinked microsatellite loci (Fst) as pairwise comparisons between the four populations.

Allele frequency maps. Resistance allele frequencies were calculated using the Microsatellite toolkit (22) and were used as input for maps (the frequencies are presented in Table S2 in the supplemental material). Samples missing an allele, because of either failed sequencing reactions or restriction enzyme digestion, were not included in the maps. Frequencies were arranged into those originating before 2006 and those originating in 2006 or after. Specific localities and dates of sampling are presented in Table S1 in the supplemental material.

RESULTS

dhfr and dhps circulating genotypes. Of the 146 samples processed, 135 yielded interpretable results for both genes. There were 5 failures in the dhfr locus only and 2 in the dhps locus only. Figure 1 shows that in southwest and northwest Colombia the circulating dhfr alleles are wild-type NCSI (51/59/108/164) single NCSI (mutations are underlined) and double ICNL mutants, while in eastern Colombia (in the Amazon and Orinoco basins), single NCSI, double ICNL, and triple ICNL mutant alleles were found. Likewise, at dhps there were regional differences between eastern and western sites. Western sites were characterized by a mixture of sensitive AKA (437/540/581) and single GKA mutants, while eastern populations had double GKG and triple GEG mutants, as well. In the Southwestern Pacific coast, where SP treatment is still efficacious, there is a high prevalence of the wild-type genotypes dhfr NCSI (19/30; 63.3%) and dhps AKA (100%). In the Northwest Pacific coast, the majority of circulating parasites were dhfr ICNI double mutants (27/29; 93.1%) and wild type (AKA) (10/19; 52.6%) and single GKA dhps mutants (19/29; 65.5%). In the Orinoco basin, the majority of genotypes were dhfr ICNI double (37/39; 94.8%) and single GKA dhps (31/36; 86.1%) mutants. Only a single sample (from Vistahermosa, Meta, 2005) displayed the triple dhps GEG genotype. The dhfr dou-
able ICN/dhps single GKA allele was the most common in both the northwest (18/29; 62.1%) and Orinoco (29/35; 83%). In the Amazon, a high proportion (13/16; 81.2%) of dhfr ICNL triple mutants was observed before 2006, but they disappeared in 2006 or after as a result of an increase of NCNI single (2/16, 12.5%, to 7/26, 26.9%) and ICNI (from 1/16, 6.3%, before to 19/26, 73.1%, in 2006 or after) double mutants. Similarly, dhps double GKG mutants disappeared in 2006 and after (6/17; 35.3%, before 2006), while triple dhps GEG mutants increased slightly (8/17, 47%, before 2006 to 15/26, 57.7%, in 2006 or after). Sequencing of eight dhfr ICNL triple-mutant samples showed that all bear the “Bolivian repeat” (BR) (a 15-bp repeat between codons 30 and 31 with no apparent effect on drug resistance). In addition, two sequenced samples from Amazonas had a mutation at codon 50 in the dhfr gene, and in both cases, it was found in the RICNI (50R/51I/59C/108N/164I) triple-mutant allele.

**Microsatellite diversity and origins of resistance genotypes.**

Twelve dhfr-linked microsatellite haplotypes (coded 1 to 12) were identified (see Table S1 in the supplemental material). dhfr ICNI double mutants were found in all four regions associated with 5 different dhfr-linked microsatellite haplotypes (“2,” “4,” “5,” “11,” and “12”); haplotype “4” was highly prevalent in all four regions (from 38.5% in the southwest to 95.7% in the northwest). Except for haplotype “12,” which can be explained as the result of a recombination event with other circulating parasites, the haplotypes were represented by a single sample and were related to the prevalent haplotype either by their close allele sizes or by sharing one of the two alleles (Table 1). This suggests that they all have a common origin. dhfr ICNL triple mutants were associated with three different dhfr-linked microsatellite haplotypes (“1,” “2,” and “7”) two of which (“1” and “7”) were found in a single sample.

<table>
<thead>
<tr>
<th>Origin</th>
<th>No. of samples</th>
<th>dhfr allele</th>
<th>mDHFRb (0.3 kb)</th>
<th>MAIc (5.3 kb)</th>
<th>Haplotype code</th>
</tr>
</thead>
<tbody>
<tr>
<td>SWPa</td>
<td>2</td>
<td>ICNI</td>
<td>98</td>
<td>201</td>
<td>4</td>
</tr>
<tr>
<td>NWPb</td>
<td>20</td>
<td>ICNI</td>
<td>98</td>
<td>201</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>ICNI</td>
<td>96</td>
<td>201</td>
<td>2</td>
</tr>
<tr>
<td>Orinoco</td>
<td>21</td>
<td>ICNI</td>
<td>98</td>
<td>201</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>ICNI</td>
<td>104</td>
<td>201</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>ICNI</td>
<td>98</td>
<td>203</td>
<td>5</td>
</tr>
<tr>
<td>Amazonas</td>
<td>1</td>
<td>ICNI</td>
<td>125</td>
<td>221</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>ICNI</td>
<td>98</td>
<td>201</td>
<td>4</td>
</tr>
<tr>
<td>Amazonas</td>
<td>18</td>
<td>ICNL</td>
<td>102</td>
<td>221</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>ICNL</td>
<td>100</td>
<td>221</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>ICNL</td>
<td>96</td>
<td>199</td>
<td>1</td>
</tr>
</tbody>
</table>

**TABLE 1. dhfr-linked microsatellite haplotypes at double- and triple-mutant dhfr alleles**

<table>
<thead>
<tr>
<th>Origin</th>
<th>No. of samples</th>
<th>dhps allele</th>
<th>m0.8 (−0.8 kb)</th>
<th>m4.3 (−4.3 kb)</th>
<th>Haplotype code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Orinoco</td>
<td>3</td>
<td>GKG</td>
<td>115</td>
<td>104</td>
<td>19</td>
</tr>
<tr>
<td>Amazonas</td>
<td>1</td>
<td>GKG</td>
<td>113</td>
<td>102</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>GKG</td>
<td>115</td>
<td>104</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>23</td>
<td>GEG</td>
<td>115</td>
<td>104</td>
<td>19</td>
</tr>
</tbody>
</table>

**TABLE 2. dhps-linked microsatellite haplotypes at double- and triple-mutant dhps alleles**

The prevalent haplotype (“2”) shares a common allele with haplotype “1” at the MAI locus; the two alleles at the DHFR locus are closely related in size (Table 1). Furthermore, sequencing of a random sample of 7 individuals, including haplotypes “7” and “8” and representing three different resistance genotypes, showed that they all contain the “Bolivian repeat,” indicating that the majority of dhfr triple mutants have a single origin.

Eighteen dhps-linked microsatellites haplotypes (coded 13 to 30) were identified (see Table S1 in the supplemental material). dhps GKG double mutants were associated with 2 dhps-linked microsatellite haplotypes (“17” and “19”) and GEG triple mutants with only one (“19”), indicating a single origin for the dhps GKG double and GEG triple mutants (Table 2). Single dhfr NCNI and dhps GKA mutants are related to 3 and 12 different linked microsatellite haplotypes, respectively. The large haploptic diversity of dhfr- and dhps-linked microsatellites (average Hz, 0.813; n = 42) and the presence of eastern- and western-specific haplotypes suggest that single-mutant genotypes for both resistance loci have arisen on several occasions.

**Genetic differentiation and migration.** The largest number of shared multilocus association combinations occurred across the Andes, between northwestern and eastern Orinoco populations (see Table S3 in the supplemental material). Three complete multilocus associations (i.e., dhfr- and dhps-linked microsatellite haplotypes and the PolyA and PfPK2 unlinked microsatellite combination) were shared between the northwest and Orinoco, two between Amazonas and Orinoco, and one between the northwest and southwest populations. If only the combination of neutral loci (PolyA and PfPK2) is taken into account, four combinations are observed to be shared between the northwest and Orinoco, two between Amazonas and Orinoco, and two between the northwest and southwest populations.

It is expected that the introduction of new alleles into a population increases the genetic diversity and diminishes the differentiation between populations. The highest diversity values (expected heterozygosities) for the unlinked-microsatellite combination and the complete multilocus association genotype were consistently observed in the northwestern population (Table 3). Diversity values for all individual loci are presented in Table S4 in the supplemental material. Similarly, pairwise
differentiation indices showed evidence of significant population structure. The Orinoco/northwest and the northwest/ southwast populations were the less differentiated populations ($F_{ST}$ values: 0.094 and 0.070, respectively). In contrast, the most differentiated populations were the Amazon and Orinoco subpopulations ($F_{ST}$, 0.252) and the Tumaco population (average $F_{ST}$ value, 0.235) relative to the eastern subpopulations (Table 4).

**DISCUSSION**

The distribution of $dhfr$ and $dhps$ genotypes found in this study is consistent with the observed geographical differences in the therapeutic efficacies of SP against uncomplicated *P. falciparum* malaria in Colombia. Where SP is still efficacious (southwestern Colombia), the majority of circulating genotypes were wild type for both $dhfr$ and $dhps$ genes; where there is moderate resistance to SP (northwestern Colombia), the majority of circulating genotypes were $dhfr$ double (ICN1) and $dhps$ single (GKA) mutants. In the Amazon basin, where 80% SP treatment failures have been documented (21), quadruple-, quintuple-, and sextuple-mutant resistance genotypes were observed (Fig. 1). The absence in the Amazon of the highly resistant $dhfr$ ICNL genotypes and the presence of the $dhps$ wild type (AKA) after 2006 suggests the relaxation of drug pressure in the area. However, SP was removed from the national antimalarial drug policy only at the end of 2007. Accordingly, it is important to consider, due to the high human mobility in the area, the impact of Peru’s abandonment of SP to treat *P. falciparum* malaria in their Amazon region in 2001 (20).

A number of genotypes associated with SP resistance have been described from South America (see reference 18 for a review). Our results concur with previous observations from South American samples suggesting a common origin for the so-called secondary antifolate resistance mutants. Samples bearing the $dhfr$ ICNL triple-mutant allele from Amazonian Peru and Bolivia have the same or very similar $dhfr$-linked microsatellite allele sizes, as well as the Bolivian repeat, suggesting a common origin (2,6,30). Contrary to observations in other places in South America (2,30), our data also suggest a single origin for Colombian $dhfr$ ICN double mutants. The triple mutant RICNI, with an apparently different origin and conferring midlevel resistance, has been reported from Brazil, Bolivia, Peru, Suriname, and Venezuela (2,5,24).

Likewise, as previously observed, $dhps$ double GKG and triple GEG mutants seem to have a single origin in the Amazon and Orinoco basins (2,24,30) and have not disseminated into western Colombia. The single $dhps$ GKA mutant has been reported in Brazil, Colombia, Peru, Suriname, and Venezuela; the double mutant GKG from Brazil, Peru, and Venezuela; and the triple mutant GEG in Amazonian Brazil, Bolivia, Peru, Suriname, and Venezuela. Analyses of microsatellites closely linked to the resistance genes have shown that double and triple mutants seem to have a common origin (2,6,11,17,30).

The number of shared haplotypes and multilocus association genotypes and the high levels of genetic diversity values for the Northwest Pacific (0.915) population, as well as the low $F_{ST}$ values (0.094) between the Northwest Pacific and Orinoco populations, indicate the occurrence of genetic exchanges across the Andes mountain range between specific eastern (Orinoco) and western (Quibdo) localities in Colombia, most likely due to human migration. There is also evidence for genetic exchanges within each of the eastern and western populations. Since it is expected that the probability of a particular genotype invading another population increases with its frequency, genotypes with the highest frequency values are thought to have the highest probability of invading. This, as well as the time of emergence of a particular haplotype or multilocus association genotype, should give some indication as to the direction of migration. Following these criteria, our data indicate that migration has taken place in both directions (east and west of the Andes). Relatively high-frequency haplotypes in the Northwest Pacific population before 2006 are observed at relatively high frequencies in Orinoco and the Southwest Pacific after 2006, suggesting a flow of migration from the Northwest Pacific to the Southwest Pacific and Orinoco. Similarly, putatively clonal, complete multilocus association genotypes observed at a high frequency (60%) in Orinoco before 2006 are observed in the Northwest after 2006 (a detailed account of the dates of collection for each of the samples used in this study can be found in Table S1 in the supplemental material). The highest diversity values for the multilocus association genotypes occur in the Northwest Pacific population (Table 3). This suggests that although migration occurs in both directions, the flow may be higher from Orinoco to the Northwest Pacific.

In conclusion, these findings support the hypothesis of the expansion of *P. falciparum* SP-resistant populations and high-
light its importance for the formulation of multiregional antimalarial drug policies to deter the spread of drug resistance. Mapping human migration routes and understanding human migration patterns east and west of the Andes would be useful to decide whether malaria control strategies targeted to internal migrants are feasible. Further studies are required to determine the consequences of SP withdrawal in each of the regions and the magnitude and directions of genetic flow between populations on a larger scale (e.g., South America).

ACKNOWLEDGMENTS

This work was supported by COLCIENCIAS contract number 366-2006 and WHO/TDR contract number T16/181/658 ID no. A41403. We thank the managers of the malaria control programs in Amazon, Chocó, Narino, Meta, and Guaviare for their collaboration in the sample collection, as well as Santiago Nicholls for support with the acquisition of samples.

REFERENCES