Following the path of most resistance: *dhps* K540E dispersal in African *Plasmodium falciparum*

Inbarani Naidoo1,2 and Cally Roper1

1London School of Hygiene and Tropical Medicine, Keppel Street, London, WC1E 7HT, UK
2Malaria Research Programme, Medical Research Council, Durban, South Africa

Chloroquine resistant malaria (CQR) emerged in East Africa during the late 1970s and then spread westward. A molecular marker only became available in the late 1990s, and by that time CQR had permeated throughout Africa. By contrast, resistance to sulphadoxine-pyrimethamine (SPR) has emerged during an era of molecular surveillance, and the changing prevalence of SPR conferred by point mutations in the *dhfr* and *dhps* genes has been recorded in hundreds of sites across Africa. We have collated and mapped reports of the *dhps* K540E mutation, a uniquely informative marker of SPR, and used these to describe the geography of its dispersal through time. Like CQR, *dhps* K540E appeared first in East Africa and spread west. We discuss whether there are common principles governing resistance dispersal in Africa and how these might guide surveillance in future.

Drug resistance in African *Plasmodium falciparum*

Artemisinin combination therapies (ACT) such as arte-mether–lumefantrine, artesunate–mefloquine, artesunate-SP or artesunate–amodiaquine are now the recommended first-line treatment for *falciparum* malaria throughout the world. Unfortunately the future efficacy of ACTs has been overshadowed by reports of the emergence of artemisinin resistant *P. falciparum* in Pailin, western Cambodia, near the border with Thailand [1]. Although the mechanism of resistance has not yet been elucidated, there is a real concern that artemisinin resistance might escape this limited regional focus [2] and spread globally, as happened previously with resistance to chloroquine and pyrimethamine. In light of these recent developments, it is timely to take stock of what is currently understood about resistance dispersal.

The molecular basis of resistance to key malaria drug treatments of the past are now well understood, and the major genetic determinants of chloroquine and sulphadoxine–pyrimethamine (SP) resistance have been characterised. The major genetic determinant of chloroquine resistance are point mutations in the *pfcrt* gene [3], whereas point mutations in *dhps* and *dhfr* genes confer resistance to sulphadoxine and pyrimethamine respectively [4–7].

Focusing on resistance dispersal in African *P. falciparum*, we investigate the K540E mutation in the *dhps* gene. This marker has the potential to be uniquely informative because unlike *dhfr* and *pfcrt* resistance, mutations at this locus have emerged and spread during the era of routine molecular resistance surveillance. We have mapped all published reports of this mutation and used these to chart the dispersal of resistance in Africa. We then compare the molecular dispersal map with maps charting the first reports of chloroquine treatment failure across Africa from 1978 to 1989 and determine if there are general principles underlying the dispersal of resistance, which might be applied for improvement of surveillance and management of resistance.

Chloroquine resistance

The first confirmed clinical reports of chloroquine resistance (CQR) occurred in Pailin, the same area of western Cambodia where artemisinin tolerant parasites are now emerging [1,8]. A global review of CQR reports was compiled by Payne (D. Payne, Ph.D. thesis, University of London, 1989) [9]. Two initial foci of confirmed resistance in Asia and South America in 1960 were described as well as their subsequent dispersal in all epidemiologically feasible directions [8,9]. The first appearance of CQR in Africa occurred 17 years later.

A summary of CQR reports in Africa by Charmot and colleagues [10] described the incremental spread of resistance across the continent from east to west during a 12 year period (Figure 1). The authors recognized the key role of human migration in the dispersal of resistance and observed that increasing urbanisation during the 1970s and 1980s had exerted an influence on migrant labour. They noted that early reports of CQR in West Africa were from major centres of population, for example Lambarene in Gabon, Luanda in Angola, Malabo in Equatorial Guinea, Cotonou in Benin, Ibadan in Nigeria and Kaolack in Senegal [9], and they proposed that it was from these initial foci that resistance then spread [10].

Molecular analysis of CQR conferring mutations in the *pfcrt* gene has confirmed the observation by Payne that resistance radiated outwards from the initial Asian and South American foci [3,11], and analysis of markers in the *pfcrt* flanking sequence has confirmed speculation [10] that Asian CQR was imported into Africa [3,11] and then dispersed across the continent [12]. Although resistant *pfcrt* was probably introduced into Africa on multiple occasions, (it was abundant throughout Asia long before 1978) the...
clear inference of the geospatial data in Figure 1 is that CQR emerged in East Africa and dispersed from there, and this interpretation is widely accepted [13]. The first appearance of CQR in East Africa was shown to be co-incident with its first arrival in India (D. Payne, Ph.D. thesis, University of London, 1989). A plausible (but now untestable) hypothesis is that resistant parasites were first introduced to East Africa via the international ports of the Indian Ocean.

**Pyrimethamine resistance**

Pyrimethamine monotherapy was used widely in the 1950s and 1960s for treatment and for mass prophylaxis [14]. In contrast with CQR, the pyrimethamine resistance trait emerged almost immediately in a multifocal manner in both East and West Africa. Early reports showed rapid emergence of pyrimethamine resistant malaria in Kenya (1954), Tanzania (1954), Sudan (1955), Cameroon (1958), Burkina Faso (1958), Nigeria (1960), Ghana (1962), Gambia (1963), Senegal (1966), and Liberia (1980) (Figure 2) [14].

Dispersal of resistance from these early foci was considerable. In Mkuzi, Tanzania, pyrimethamine resistance was documented 100 miles from the original site of selec-
samples, they have confirmed the existence of triple mutant dhfr in Cameroon as early as 1985 [25], and of double and triple mutant dhfr alleles in Kenya in 1988 [19]. These early records of the triple mutant allele pre-date the arrival of SP treatment failure in East Africa in 1994 to 1995 [27,28] and lend support to the hypothesis that the triple mutant dhfr was introduced into Africa alongside pfcrt [22]. Pyrimethamine resistance was widespread in Asia during the 1970s [29], making a double importation feasible.

Inferring dispersal from reports of resistance
The pan-African distributions of Asian derived pfcrt and dhfr lineages today gives no clue to their original site(s) of introduction or their routes of dispersal. The passage of pfcrt can be inferred from the geospatial pattern of chloroquine treatment failure reports (Figure 1) [10], but an equivalent trail of pyrimethamine resistance is not recorded because the successful combination of SP masked the presence of pyrimethamine resistant parasites. It was only when sulphadoxine resistance appeared and combined with pre-existing pyrimethamine resistance, that cases of SP treatment failure began to be reported [30].

Sulphadoxine resistance
Resistance to sulfadoxine is acquired by the progressive accumulation of mutations in the dhps gene. The emergence of resistant dhps began in 1993 and heralded the arrival of SP treatment failure in Africa. In particular, the dhps double mutant (A437G + K540E) combined with the dhfr triple mutant (i.e. quintuple mutant or full house) is predictive of early SP treatment failure [31–33]. The A437G mutation can occur alone or in combination with K540E, but the K540E mutation is almost invariably found in association with A437G [30]. The occurrence of A581G, in addition to A437G or A437G + K540E confers higher levels of resistance but is still comparatively scarce in Africa (http://www.drugresistancemaps.org/maps/dhps581/). We have collated all published reports of sequence variation at amino acid 540 to map its geospatial distribution and determine how that changed through time.

Molecular surveillance coverage
In the review of the published literature (Box 1 and Table 1), we identified 212 surveys in 148 unique geographic localities of 37 countries Africa, where a total of 21 362 P. falciparum isolates were tested for the K540E mutation. The K540E mutation was present in 40% of the isolates. The sites of these survey data points are presented in Figure 3. The 106 references which describe them are listed in the Supplementary Data S1, and a Google map of their exact location is available at http://www.drugresistancemaps.org/maps/dhps540/, where each study is flagged and a link to the original reference is embedded. Figure 3 maps the 212 survey data points, showing they are distributed in 37 of

Figure 2. Ten countries with pyrimethamine resistance between 1954 and 1980 in East and West Africa. The years (blue numbers) indicate the year of the first report of pyrimethamine resistance in a particular country. The map is based on data from Peters [14].
the 50 malarious countries in Africa. The malarious countries without any published studies of K540E data are: Algeria, Botswana, Burundi, Cape Verde, Chad, Eritrea, Mauritius, Mayotte, Sierra Leone, Somalia and Togo.

Molecular surveillance methodology
In all the studies identified, PCR was used to amplify a dhps sequence fragment, generally from small quantities of parasite DNA, collected by finger-prick blood sampling. A variety of molecular methods were used to detect resistance mutations in the amplified sequence. Of the 104 articles reporting K540E prevalence, 50% used PCR-RFLP, 23% used direct sequencing methods, whereas probe hybridization (11%), primer extension (4%) and mutant specific PCR amplification (10%) accounted for the remainder. A small proportion (2%) used other methods.

Molecular genetic data are not reported in a standardised format, but the most common statistic reported is the prevalence of a given mutation. Prevalence is the proportion of infections analysed in which the mutation was detected. Since people are often simultaneously infected with multiple genotypes, it is not unusual for both the

<table>
<thead>
<tr>
<th>Online database</th>
<th>Search terms</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pubmed</td>
<td>Malaria, 540, dhps, pfDHPS, dihydropteroatesynthase</td>
<td>516</td>
</tr>
<tr>
<td>African Journal Online</td>
<td>malaria, 540, Dhps, pfDHPS</td>
<td>22</td>
</tr>
<tr>
<td>Bioline</td>
<td>dihydropteroatesynthase</td>
<td>2</td>
</tr>
</tbody>
</table>
Crossing the East–West divide

The time sequence of reports confirming the presence or absence of the K540E mutation between 1993 and 2008 is illustrated in Figure 4. The data spans 15 years and is subdivided into 5 year blocks. In the first 5 years between 1993 and 1998, the mutation was confirmed in East Africa and absent in West Africa. The earliest records of K540E came from East Africa. In Kenya, it was found in samples collected from 1993 to 1995 [35], in Tanzania during 1995 [36] and in Malawi from 1995 to 1996 [37]. Within four years, the East African distribution of K540E had extended significantly to include South Africa, where it was absent in 1995 and 1996 [18] but present in 1998 [38]. It was found in Uganda in 1999 [32], and Mozambique and Zambia in 2000 [39,40]. After ten years the range expanded to 15 countries, and after 15 years it had been confirmed in 27 countries. A pan-African study of microsatellite polymorphism flanking the dhps gene [30] has shown that the K540E found in parasite populations throughout Tanzania, Kenya, Uganda, eastern Democratic Republic of Congo, Zambia, Mozambique and South Africa are descendants of the same ancestral A437G + K540E double mutant lineage, referred to as SGE1. The same study identified small numbers of the SGE1 lineage in the Central and West African parasite populations of Gabon, Congo, Guinea and Namibia. This demonstrates that the arrival of the K540E mutation in Central and West Africa is attributable to dispersal of the SGE1 lineage from East African populations rather than de novo occurrence of the mutation in those populations.

There is an apparent gulf between East and West African parasite populations, which is perhaps attributable to the difficulty of travel between towns and cities in Central Africa. The first analysis in the Democratic Republic of Congo (DRC) was of parasites collected during 2002 [41], which found that K540E was present in the east of the country at a high prevalence. A study comparing parasites from multiple sites in the DRC during 2004 found the prevalence of 540E was 13.3% to 19.3% in eastern sites and 0.9% to 3.9% in western sites [42]. In the same year, K540E was found in 0.8% of P. falciparum infections in Republic of Congo [30], 0.2% in Cameroon [30] and 23.8% in Nigeria [43]. By 2005, small numbers of these mutations had appeared in Guinea [44], and by 2007 in Gabon [30] and Mali [45]. The cumulative reports confirming K540E show its gradual spread westward. Of the 37 countries for which data exists, 27 have now been confirmed to have K540E, leaving ten countries where recent surveys have been carried out, but K540E has still not been recorded: Burkina Faso [30,46,47], Gambia [30,48], Guinea-Bissau [49], Ivory Coast [50], Liberia [51], Madagascar [52], Mauritania [53], Niger [54], Senegal [30,55,56] and Equatorial Guinea [57] (Figure 4) Box 1.

Independent origins in northeast and southeast African populations

There are two major lineages of K540E, as well as the first lineage, designated SGE1, which has the wide geographical range described above, there is a second independently evolved lineage, designated SGE2. This emerged over the same time period as SGE1 but has been found exclusively in parasites from northeast Africa [30]. In northeastern Sudan, K540E was absent in 1993 but had appeared by 1998 [58]. In neighbouring countries, the K540E mutation has been confirmed; samples collected in Djibouti in 1998 and 1999 [59] and Ethiopia in 2004 [30] show K540E to be
thoroughly established. The Ethiopian and Sudanese K540E was confirmed to be of SGE2 lineage [30]. We assume that Djibouti might be the same, but this needs to be tested. The predominance of a different resistance lineage in the horn of Africa suggests a degree of separation from other parasite populations in East Africa, which was exacerbated during the 1990s because of political instability and civil war in southern Sudan.

Macro-geographic distribution and local prevalence of K540E

The K540E was clearly abundant in the southeast African region from 1995 onwards, and it is clear that an overriding determinant of the presence or absence of K540E mutations in African parasite populations is their place in the continent-wide distribution of the mutant lineages. Pearce et al. [30] described five regional parasite gene pools defined on the basis of dhps resistance allele sharing. Within these five regional clusters, the same mutant dhps lineages were common to all survey sites. The southeast African gene pool encompasses Tanzania, Zambia, Kenya, South Africa, Uganda, eastern DRC and Mozambique. The northeast African gene pool includes Ethiopia and Sudan, whereas the west African consists of Burkina Faso, Ghana, Guinea, Nigeria, and Senegal. Southwest Africa includes the Republic of Congo, Namibia, Angola, and Gabon. Cameroon was different again, and for the purposes of this study we termed this the central African parasite gene pool. We sorted 540E prevalence reports into the five mainland regions in order to examine trends in prevalence over time. Reports of K540E from island populations were put into a sixth separate category (Figure 5).

Graphs of the prevalence (%) of K540E with 95% confidence intervals were plotted according to year from 1988 to 2008 for each region (Figure 5). The 540E mutation was very clearly most abundant in the southeast African region.

Figure 4. The expanding geographical range of K540E mutation in Africa between 1993 and 2008. The data spans 15 years. (a) In the 5 years between 1993 and 1998, the mutation was confirmed in six countries. (b) By 2003 (ten years), this had increased to 15 countries. (c) By 2008 (15 years), K540E was recorded in 27 African countries. The mutation is still absent in ten countries, mainly in West Africa. There were nine data points in which a study year was not reported. Of these three studies, Mali and Equatorial Guinea did not record any 540E data, and six studies which did were from Angola, Kenya, Mozambique, Uganda and Tanzania, reporting prevalences ranging from 11.8% to 88%. We included the prevalence of 11.8% reported in Angola in map C, judging that the survey took place prior to the publication date.

(a) 1993-1998

(b) 1993-2003

(c) 1993-2008

Key:
- K540E absent
- K540E present
- Malarious countries with no published K540E data

Within these five regional clusters, the same mutant dhps lineages were common to all survey sites. The southeast African gene pool encompasses Tanzania, Zambia, Kenya, South Africa, Uganda, eastern DRC and Mozambique. The northeast African gene pool includes Ethiopia and Sudan, whereas the west African consists of Burkina Faso, Ghana, Guinea, Nigeria, and Senegal. Southwest Africa includes the Republic of Congo, Namibia, Angola, and Gabon. Cameroon was different again, and for the purposes of this study we termed this the central African parasite gene pool. We sorted 540E prevalence reports into the five mainland regions in order to examine trends in prevalence over time. Reports of K540E from island populations were put into a sixth separate category (Figure 5).

Graphs of the prevalence (%) of K540E with 95% confidence intervals were plotted according to year from 1988 to 2008 for each region (Figure 5). The 540E mutation was very clearly most abundant in the southeast African region.
Figure 5. Prevalence (%) of K540E with 95% confidence intervals plotted according to year from 1988 to 2008 for each region. Each blue square represents a survey from a particular site in a particular year, and the 95% confidence intervals (generated by binomial exact) reflect the size of the survey sample. The mainland surveys were sorted into the five genetically distinct regions [31], whereas surveys on islands were treated separately as a sixth category. In the southeast African region, 108 surveys in 78 sites were carried out in the countries listed. In the Northeast, 27 surveys were carried out in 13 survey sites in the countries listed. In the West, 43 surveys were carried out at 34 unique survey sites in the countries listed. In the Southwest, there were 15 surveys carried out at 14 unique survey sites. In the central African region, there were 14 surveys carried out in 7 unique survey sites. On Comoros, Madagascar and Sao Tome/Principe, there were five surveys available for three unique survey sites. Graphs were generated in Intercooled Stata 9.2, showing the prevalence of K540E with 95% exact binomial confidence intervals.
from 1995 onwards. The 108 data points record the gathering abundance of K540E over the 20 year surveillance period throughout the southeast African region with the exception of island populations such as Madagascar and the Comores, where 540E is still rare or absent. Other regions had fewer data points but the rising prevalence of the 540E mutation in northeastern countries was clear: (i) 70% of the 27 data points had less than 25% K540E between 1990 and 2002, (ii) six sites had greater than 70% K540E in 2003 to 2004. In the western region, the highest prevalence reported was 24%, which was in Nigeria in 2004. The majority of the west African data points (78%) did not have K540E. In Southwest and Central Africa, the K540E was absent completely until 2004, after which low prevalence was reported in Namibia, Congo, Gabon, Cameroon and Central African Republic.

Three common principles governed resistance dispersal in the past
Parasite populations are linked through networks of human circulation
Once a resistance mutation is established at a site in one of the major regional gene pools it can be disseminated rapidly to other populations in the same regional gene pool. Circulation of infected people between the major regions and the mainland and island populations is less frequent and consequently exchange of resistance genes is more stochastic.

Rates of dispersal on the African mainland
CQR and K540E arrived in the same east African region independently and perhaps via the same route (see below). Having arrived, they each spread rapidly throughout the southeast African region within a four to five year period, and they each appear to have reached much of Central and West Africa within 15 years. Although the data is patchy, and there are different detection sensitivities for the reporting of mutations versus treatment failures, the progression of the SGE1 lineage across Africa during 1993 to 2008 is broadly similar to the geographic progress of chloroquine resistance between 1978 and 1990.

Starting in East Africa
Pearce et al. [30] reported that SGE1 has a matching microsatellite haplotype to a cultured parasite line known as ‘K1’ which was originally isolated from a patient in Thailand. This indicates that the SGE1 double mutant, like the dhfr triple mutation and pfcrt mutation, originated in Southeast Asia. The geographic dispersal of K540E outwards from East Africa follows a similar path to that of CQR in Africa (Figure 1). It is tempting to speculate that there was a common mechanism by which these resistance adaptations were transported globally.

Predicting resistance dispersal in the future?
Will artemisinin resistance follow the same path as CQR and the dhps mutation K540E? Since Pailin was the crucible for both CQR and artemisinin resistance, it now perhaps seems probable that the same patterns of dispersal will play out all over again. There are a number of significant factors that have changed or are changing: (i) patterns of human circulation change according to the dictates of war, trade and transport infrastructure. New trade partnerships have been established between Africa and the rest of the world since the 1970s, and the volume of travel, with its associated risks of disease dispersal via the worldwide airline network are considerably increased [60]. (ii) Rates of infection have been reduced significantly in many regions of Africa, and the numbers of migrants likely to be infected with malaria is proportionately less. This will reduce the mobility of resistance. (iii) Improved access to treatment and improved diagnostic practices will increase the proportion of infections treated and with it the strength of drug pressure in Africa. Elevated drug selection through higher treatment coverage increases the likelihood of de novo resistance mutations occurring in Africa. Conversely, an improved formulation and dosing in treatment, for example through the combination of artemisinin with an effective companion drug [61] and appropriate dosing of children [62], will both reduce the selection pressure, and with it the likelihood of de novo resistance emerging. (iv) Artemisinins have been successfully used in Cambodia for more than 30 years, which indicates that resistance mechanisms do not quickly arise by de novo mutation; however, we do not yet know the molecular basis of artemisinin resistance, and this will become clearer once the genetic basis of resistance is discovered. Analysis of pfmdr shows that if copy number variation were the underlying basis for changes in artemisinin sensitivity, multiple independent foci of resistance can be expected [63]. On the contrary, if single nucleotide polymorphisms are the genetic basis for resistance, as with pfcr t, dhfr and dhps, resistance alleles are likely to arise infrequently, and dispersal will play the central role in the introduction of resistance to new geographic regions [22].

Conclusions
Although local resistance levels are moderated by local drug use, it is now clear that the availability of resistance genes to a local parasite population is largely determined by its geographical location. This is because a mutation’s distribution is governed by pathways of parasite circulation which describe large, natural populations and their regional boundaries. With this insight, can molecular surveillance be made more efficient? The molecular surveillance coverage of Africa illustrated in Figure 3 is remarkable. Although data are absent in a few areas, most regions are represented by at least one study, and this level of coverage was achieved through the independent activities of many research groups. With new understanding of the dynamics of parasite circulation and by the coordination of monitoring through regional networks, the gaps in coverage can be addressed. Vigilance against new multi-resistant strains can protect malaria treatment in the coming years but only if underpinned by a contemporary geography of migration and malaria.

Appendix A. Supplementary data
Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.pt.2010.05.001.
References