Towards an Atlas of Human Helminth Infection in sub-Saharan Africa: The Use of Geographical Information Systems (GIS)

S. Brooker, M. Rowlands, L. Haller, L. Savioli and D.A.P. Bundy

The value of a geographical perspective to infectious disease epidemiology and control has long been recognized. However, the labour required to produce maps, and keep them up to date, has inhibited the development of this area, and very little is currently known about the spatial distribution of parasitic infections other than malaria, trypanosomiasis and onchocerciasis. A recent initiative by an international group of collaborators is attempting to redress the absence of detailed spatial information on the major helminth infections of humans. In this article, Simon Brooker and colleagues describe progress made by this initiative in mapping helminth infections in sub-Saharan Africa, highlighting the value as well as the limitations of this empirical mapping approach.

Single-dose drugs have provided effective and safe tools for the control of morbidity caused by infections with schistosomes and intestinal geohelminths[1,2]. However, the appropriate targeting of chemotherapy requires information on the distribution of infection prevalence within countries to identify high-risk areas that might benefit most from control. The need for schistosomiasis and intestinal geohelminth control is greatest in sub-Saharan Africa[1,3] and this area provides the greatest challenge for geographically targeted control and the most immediate need for accurate distribution maps. To date, a geographical approach has yielded valuable information for malaria[4,5], onchocerciasis[6] and African trypanosmiasis[7]. However, systematic information on helminth infections is not currently available for sub-Saharan Africa because national surveys have been conducted in only a few countries[8–13,*], and, for most countries, information on prevalence is widely scattered across the literature and not catalogued.

Previous approaches to describing infection prevalence in Africa have typically been made at the national level, using prevalence data from the few studies available within a country, which are then extrapolated to the country as a whole[14–16]. Although such an approach has proved very effective for advocacy, as demonstrated by the seminal 'This Wormy World' article by Stoll[17], it is of limited practical relevance to the targeting of control efforts. In 1987, the Division of Control of Tropical Diseases of the WHO published the *Atlas of the Global Distribution of Schistosomiasis*[18], which mapped the occurrence of schistosomiasis in 76 countries worldwide, along with tables of observed prevalence for each study location. Although an important contribution, the traditional cartographic approach has the disadvantage that the maps cannot be updated easily, and comparisons between areas shown on different maps are difficult[19]. Modern geographical information systems (GIS) offer solutions to both of these problems[20].

Here, we describe the work of an initiative, launched by the WHO with its Collaborating Centre for the Epidemiology of Intestinal Parasites at the University of Oxford, which attempts to collate the available survey data in a single database with the twofold aim of describing, where possible, the prevalence across the African continent and highlighting areas for which further information is required.

The WHO collaboration. Information is currently being sought from the international and national literature on the prevalence of infection with the two major schistosome species in sub-Saharan Africa (*Schistosoma haematobium* and *S. mansoni)* and with the major geohelminths (hookworms, *Ascaris lumbricoides* and *Trichuris trichiura*). The project intends to develop prevalence maps, using GIS as both a database and a graphical tool, for the whole of sub-Saharan Africa. Where possible, these maps are being prepared at the second administrative level – here defined as the

Box 1: The Contents and Structure of the Database

Sources of information included

Published surveys identified through computerized literature searches and complemented by reference tracing 'Grey literature' from unpublished sources collected from a manual search of material in the WHO archive in Geneva, which included unpublished reports, WHO personnel trip reports and direct contact with research institutions and non-governmental organizations (NGOs).

Personal communications from a network of contacts of WHO and the Partnership for Child Development

Inclusion criteria

Independent cross-sectional surveys (not hospital data)
Surveys after 1970
Sample size greater than 30

Information included in the database

Source – authors, title and journal information
Location of survey (recording whether it was rural or urban)
Date of survey
Method of sample recruitment
Diagnostic technique used for stool and urine examination
Population sampled (including age groups, sexes and occupation)
Number of individuals examined
Numbers found positive for each of the helminth species

*Ndir, O. *et al.* (1996) *Enquete Nationale sur la Bilharziose au Senegal.* Universite Cheikh anta Diop de Dakar.

Simon Brooker, Megan Rowlands, Laurence Haller and Don Bundy are at the WHO Collaborating Centre, Wellcome Trust Centre for the Epidemiology of Infectious Disease, University of Oxford, South Parks Road, Oxford, UK OX1 3FY. Lorenzo Savioli is at the World Health Organization, 1211, Geneva 27, Switzerland.
Tel: +44 1865 281232, Fax: +44 1865 281245,
e-mail: simon.brooker@ceid.ox.ac.uk

 0169-4758/00/$ – see front matter PII: S0169-4758(00)01687-2

Table 1. An African helminth database for sub-Saharan Africa[a,b]

Country	No. refs	No. surveys	No. of admin units in country	% of admin units for which there are data		
				Schistosomiasis	Geohelminths	Both
Angola[c]	3	39	48	10.4	2.1	0.0
Benin	6	60	77	16.9	2.6	1.3
Botswana	10	179	10	80.0	30.0	30.0
Burkina Faso	19	75	301	15.6	0.3	0.3
Burundi	8	16	15	26.7	13.3	13.3
C.A.R.	6	50	51	45.1	45.1	45.1
Cameroon	28	143	49	100.0	100.0	100.0
Congo	20	80	46	17.4	4.3	4.3
DR Congo	7	95	38	18.4	13.2	13.2
Equatorial Guinea[c]	2	5	7	0.0	14.3	0.0
Ethiopia	50	234	94	20.7	12.4	12.4
Eritrea[d]	2	2	38	0.0	0.0	0.0
Gabon	13	47	46	15.2	2.2	0.0
Gambia	6	14	37	2.7	8.1	2.7
Ghana[c]	11	57	10	30.0	70.0	20.0
Guinea	4	25	33	36.4	36.4	36.4
Guinea Bissau	1	1	37	0.0	2.7	0.0
Ivory coast	17	76	50	36.0	18.0	12.0
Kenya	45	88	47	29.8	25.5	21.3
Liberia	10	83	54	7.4	24.1	3.7
Madagascar	28	286	111	21.6	15.3	15.3
Malawi	4	42	24	83.3	4.2	0.0
Mali	24	579	46	78.3	4.3	4.3
Mauritania	8	93	44	56.8	4.5	4.5
Mozambique	4	4	128	4.7	0.7	0.0
Namibia[c]	6	28	26	11.5	11.5	11.5
Niger	16	57	35	31.4	5.7	5.7
Nigeria[c]	54	156	31	48.4	48.4	41.9
Rwanda[c]	3	5	10	10.0	20.0	10.0
Senegal	15	300	30	80.0	10.0	10.0
Sierra Leone	9	23	15	53.3	13.3	13.3
Somalia[c]	13	39	16	25.0	25.0	18.8
Sudan	10	50	18	27.8	5.6	5.6
Tanzania	55	110	97	38.1	37.1	25.8
Tchad	3	9	14	42.9	28.6	28.6
Togo[c]	4	66	5	80.0	20.0	20.0
Uganda	26	91	46	71.7	56.5	60.9
Zambia	23	97	97	19.6	4.1	4.1
Zimbabwe	14	119	57	94.7	5.3	5.3
Total	**583**	**3486**	**2304**	**29.7**	**14.9**	**12.6**

[a] This work in progress has currently identified data from 39 countries. Where possible, the prevalence data are associated with a specific district within a country, but see[c].
[b] Countries in North Africa were not included because of the concentration of disease burden as a result of helminth infection in sub-Saharan Africa. South Africa, Swaziland and Lesotho are temporarily excluded while we await publication of a separate GIS analysis of helminths for these countries.
[c] The administrative unit is the region.
[d] Not possible to assign survey data to specific admin units.

district. This level of geographical stratification was selected because it has direct administrative relevance for the increasingly decentralized government systems of Africa, and because it was considered the lowest level of stratification for which data could reasonably be expected to be available.

Mapping infection prevalence

The data included in the current atlas were abstracted from surveys incorporated into a standardized database (Box 1). Some publications reported a single survey, whereas others dealt with several surveys. Data were recorded for individual helminth species, except in the case of hookworm because differential diagnosis of *Necator americanus* and *Ancylostoma duodenale* from stool samples is difficult.

The initiative sought to allocate the location data of each survey to a specific district. The boundaries were taken from databases provided by HealthMap at the WHO. In Botswana, Ghana, Namibia, Somalia and Equatorial Guinea, digital district boundary maps have not yet been located, and so data were aggregated to the next level of administrative unit, typically a region (which can contain between three and 44 districts). Regions were also used for data for Rwanda, Burundi, Nigeria and Togo because the average districts in these countries were <600 km^2 and it proved difficult to locate surveys precisely within such small areas.

The current database incorporates 583 references, comprising 3486 independent cross-sectional surveys conducted since 1970 (see Box 1 for selection criteria). It was not possible to assign 224 surveys (6.5%) to a specific district, either because the information provided was inadequate or because the data were aggregated in some other way than at the district level. An example of this is the extensive survey carried out in

Table 2. Independent estimates of prevalence of the major helminth infections in sub-Saharan Africa[a] over the past 50 years

	S. haematobium	*S. mansoni*	*A. lumbricoides*	*T. trichiura*	Hookworm[b]
Stoll (1947)[17]	26.5	15.5	39.9	18.9	33.1
Crompton and Tulley (1987)[14]			32.3		
Utroska *et al.* (1989)[15]	25.5[c]				
Bundy *et al.* (1998)[3]			20.3	16.6	26.9
Weighted current estimates[d]					
Prevalence (%)	25.9	18.3	26.7	20.9	33.0
Total number surveyed	827 986	672 057	361166	317 359	394 983
Prevalence in school-aged children (%)	27.4	14.0	30.2	29.5	32.1
Number of school-aged children surveyed	514 679	330 533	165 599	136 331	190 932
Cases (million)[e]	131	98	161	100	192

[a] Defined as the Global Burden of Disease Study[3], but excludes South Africa, Swaziland and Lesotho.
[b] Both *Necator americanus* and *Anclyostoma duodenale*.
[c] Combined *Schistosoma haematobium* and *S. mansoni*.
[d] Estimates are for both sexes combined. Estimates are based on prevalences for each country in the database.
[e] Based on 2000 population estimates [US Census Bureau (1999). *International Data Base*].

Zambia in the 1970s*, where data are presented according to 'ecological zones', which have no reference to administrative boundaries.

Table 1 summarizes the database by country. The number of surveys available for each country varies considerably. In most countries, most surveys were conducted in the same few districts. Only Cameroon, Senegal, Mali and Zimbabwe have published national surveys available at the district level[8–13]*. Survey data appear to be particularly sparse for Guinea-Bissau, Equatorial Guinea, Angola, Mozambique, Eritrea and Rwanda.

Helminth infection in sub-Saharan Africa

Although it is possible to calculate national average prevalences based on the district prevalence figures, it is clear that such figures are potentially misleading. In particular, in certain countries, very few prevalence surveys have been undertaken, and these have typically chosen and examined areas of known high transmission. This suggests that such extrapolated national infection prevalences are more often a reflection of the numbers of surveys conducted and their location, rather than a reliable indication of prevalence. This illustrates the potential inaccuracy of prevalence estimates based on only a few studies within a country and then extrapolated to the country as a whole, as this belies the geographical heterogeneity within countries.

The present data have been mapped within defined administrative boundaries to simplify data handling, and because control approaches are increasingly implemented at the district level. However, it is unclear whether district units are an appropriate unit of comparison, because such assignment belies the finer-scale spatial heterogeneity of infection. Districts vary in size, geography and demography, so a more useful unit of analysis might be based on the number of people in a given area, or based on ecological zones without regard to administrative boundaries. Clearly, there is a trade-off between what is desirable and what is realistic.

Despite these limitations, the current data provide a crude estimate of overall disease burden for sub-Saharan Africa (Table 2). A comparison of the current estimates with those of Stoll[17] in 1947 shows a remarkable and, from a public health perspective, disappointing similarity. The only species for which there has been a substantial change in prevalence over the past 50 years is *A. lumbricoides*, where there appears to have been a fall from 39.9% to 26.7%. This might indicate that the current estimate of *A. lumbricoides* prevalence is too low, particularly because it is also lower than the estimate of 32.2% made by Crompton and Tulley[14] in 1987. However, their study reviewed 330 references; only 198 of those included in the final analyses were based on epidemiological surveys and the 58 based on hospital data might have tended to raise the apparent prevalence.

Schistosomiasis data are available for 684 (29.7%) of the 2304 administrative areas included in the atlas under construction, and geohelminthiases data are available for 343 (14.9%) areas. Data are available for both schistosomiasis and geohelminthiases in 290 (12.6%) administrative areas. This apparent lack of data is misleading, however, because many of the districts for which there is no information are also areas with very low population density, and thus may not be priority areas for initiative control. A map of average population density was constructed by calculating the average population density for each administrative area derived from a interpolation population density rater map (http://grid2.cr.usgs.gov/globalpop/africa). This layer masks 336 administrative areas (14.6%) or 948 administrative areas (41.1%) with, respectively, an average population density of <5 or <20 people per km^2. Calculated on this basis, the atlas provides information on schistosomiasis for 33.0% of administrative areas with population density of >5 people per km^2, and for geohelminths in 16.5%. The distributions of crude infection prevalence in Africa, in the absence of age stratification, are shown for schistosomiasis and intestinal helminth infections with the population masks in Fig. 1.

Sources of variation

When Norman Stoll asked over 50 years ago 'just how much human helminthiasis is there in the world?' he identified some major hurdles to an accurate estimation[17]:

> 'I need scarcely remind you of what some of those hurdles are: so many parasitological surveys of but small numbers of people, frequently by other design than to represent fair samples of an area, done by workers of varying aims and by techniques

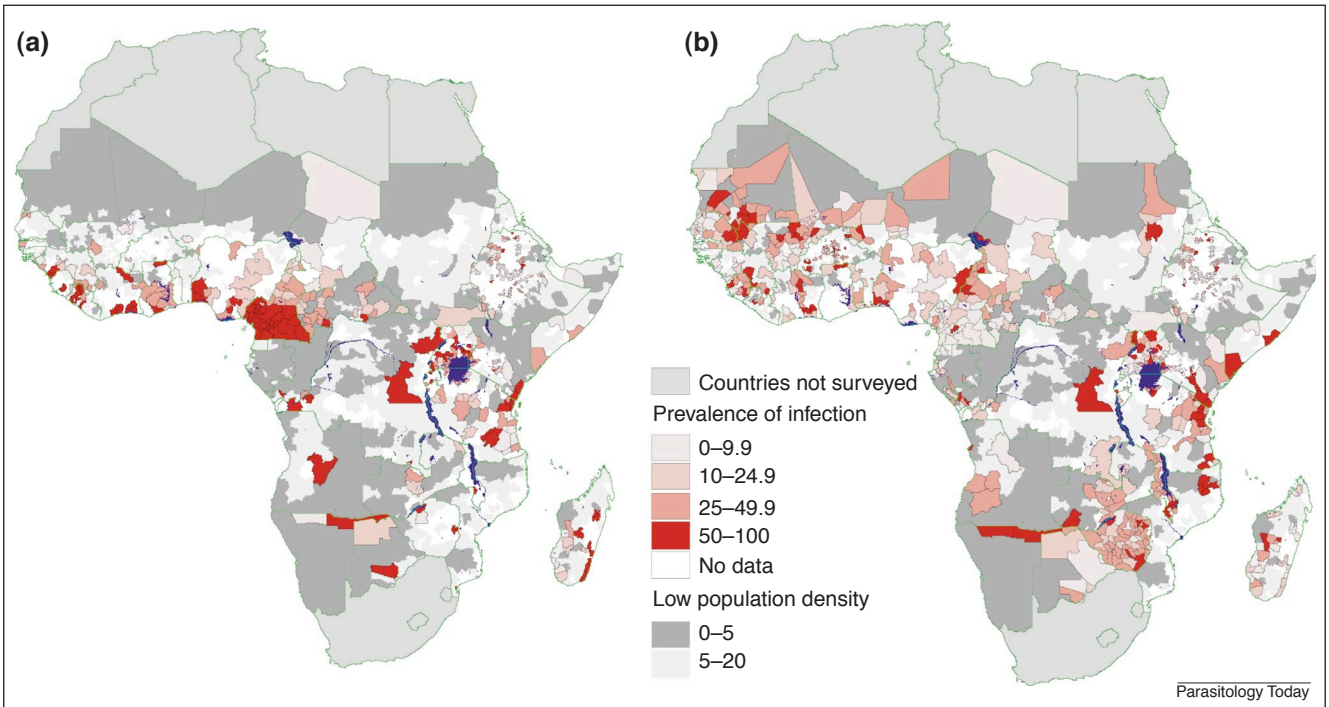

Fig. 1. Maps are shown detailing geohelminths in Africa: *Ascaris*, *Trichuris* and hookworm by district (a) and schistosomiasis in Africa: *Schistosoma mansoni* and *S. haematobium* by district (b). The geohelminth species with the highest prevalence is taken to represent the prevalence of geohelminth infection. Similarly, the schistosome species with the highest prevalence is taken to represent the total prevalence of schistosome infection. The maps show data for sub-Saharan Africa only. Of the 2304 administrative units (usually districts) in 39 countries, data are available for schistosomiasis in 684 (29.7%) and for geohelminths in 343 (14.9%). However, the population in Africa is not homogeneously distributed, and many of the districts for which no data are currently available also have low-density populations. To highlight the high population density areas of Africa we have created a template that masks out all districts with population density <5 people per km^2 and between 5 and 20 people per km^2. The atlas is also available at http://www.ceid.ox.ac.uk/child/.

of even more variable efficiency in relation to the task at hand.'

These remarks are entirely relevant to the survey results presented here. Differences in survey methods and techniques and variation in the timings of surveys produce variation in the prevalences, thereby reducing the comparability of the data and the potential of the maps to represent infection prevalence precisely in every district.

The absence of detailed age stratification is also a potentially major source of variability in the maps. The age dependency of the prevalence of infection for each species means that comparisons of crude prevalences between countries and districts are subject to error. Note that most of the surveys (56.7%) were carried out in schoolchildren, and so provide estimates for a moderately well-defined age group, which enhances the comparability of the data. Furthermore, the prevalence of infection in the school-aged population can be used to predict prevalence among preschool children and adults, and therefore among the community as a whole[21]. Thus, by mapping prevalence in school-aged children only, we can enhance confidence in age stratification which, when overlaid with population data, allows a more reliable estimate of the total population size infected[22].

Potential for application

Despite their limitations, the maps present the most detailed data currently available on the geographical distribution of helminth infection in Africa, covering 33.0% of administrative areas. However, there are many areas of the continent for which little or no data were found. The current data are thought to represent almost all of the surveys published in the international literature, and a substantial body of 'grey' literature. However, it is clear that further information exists within countries. Mapping malaria transmission in Kenya for example, Omumbo *et al.*[5] identified 682 new and independent surveys from information within published and unpublished locally archived sources. This experience demonstrates the potential value of reports and surveys from the national literature, and it is intended that the atlas will be expanded by including such information.

The information included in the atlas can help identify where current information is lacking, solicit feedback from endemic countries, or serve as a stimulus to collect new or additional data at the district level. Major efforts are now being launched by the WHO, the World Bank and other agencies to control morbidity as a result of helminth infections. Because rapid epidemiological assessment is part of the advocated control package, it is hoped that the geographical distribution of helminth infections will be updated as control progresses. The WHO is further developing and promoting user-friendly information and mapping system called the HealthMapper[23,24] (http://www.who.int/emc/healthmap/healthmap.html), which will enable this information to be managed at the country or district level, relate it to population distribution, environmental data and basic health and social infrastructure, and allow it to be regularly monitored and updated using national health statistics and survey results.

Acknowledgements
The atlas is a 'work in progress' and hence the estimates provided are preliminary. Revision and improvement require further information on the prevalence of infection within countries. If you know of relevant data that could be included, or if you would like to be a partner in this initiative, then please contact the WHO in Geneva. Collaborating partners include the Strategy Development and Monitoring for Parasitic Diseases and Vector Control, Communicable Disease Control, Prevention and Eradication, World Health Organization, 1211, Geneva 27 (Dirk Engels, Lester Chitsulo and Antonio Montreso); the WHO/UNICEF Joint Programme on Health Mapping and GIS, HealthMap (Kathy O'Neill, Jean-Pierre Meert and Isabelle Nutall); and the Partnership for Child Development and the International School Health Initiative of the World Bank. The Partnership is supported by the United Nations Development Programme, the Rockefeller Foundation, the Edna McConnell Clark Foundation, the James S. McDonnell Foundation, the Wellcome Trust, the World Bank, UNICEF, and the WHO. SB is in receipt of a Wellcome Trust Prize Studentship. DAPB acknowledges the financial support of the Wellcome Trust. We are grateful to Andrew Hall, Simon Hay and Edwin Michael for useful discussions and comments, and thank Michael Beasley and Jonathan Toomer for invaluable assistance in data preparation.

References

1 Savioli, L. *et al.* (1997) Control of schistosoniiasis – a global picture. *Parasitol. Today* 13, 444–448
2 Warren, K.S. *et al.* (1993) Helminth infections, in *Disease Control Priorities in Developing Countries* (Jamison, D.T. *et al.*, eds), pp 131–160, Oxford University Press
3 Bundy, D.A.P *et al.* Intestinal nematode infections, in *Health Priorities and Burden of Disease Analysis: Methods and Applications from Global, National and Sub-national Studies* (Murray, C.J.L. and Lopez, A.D., eds), Harvard University Press for the WHO and the World Bank (in press)
4 Snow, R.W. *et al* (1999) A preliminary continental risk map for malaria mortality among African children. *Parasitol. Today* 15, 99–104
5 Omumbo, J. *et al.* (1998) Mapping malaria transmission intensity using geographical information systems (GIS): an example from Kenya. *Ann. Trop. Med. Parasitol.* 92, 7–21
6 Ngoumou P. *et al.* (1994) A rapid mapping technique for the prevalence and distribution of onchocerciasis: a Cameroon case study. *Ann. Trop. Med. Parasitol.* 88, 463–474
7 Rogers, D.J. and Williams, B.G. (1993) Monitoring trypanosomiasis in space and time. *Parasitology* 106 (Suppl.), S77–S92
8 Wenlock, R.W. (1979) Prevalence of hookworm and *S. haematobium* in rural Zambia. *Trop. Med. Hyg.* 29, 415–421
9 Ratard, R.C. *et al.* (1990) Human schistosomiasis in Cameroon. 1. Distribution of schistosomiasis.·*Am. J. Trop. Med. Hyg.* 6, 561–572
10 Ratard, R.C. *et al.* (1991) Ascariasis and trichuriasis in Cameroon. *Trans. R. Soc. Trop. Med. Hyg.* 85, 84–88
11 Ratard, R.C. *et al.* (1992) Distribution of hookworm infection in Cameroon. *Ann. Trop. Med. Parasitol.* 86, 413–418
12 Brinkman, U.K. *et al.* (1988) Experiences with mass chemotherapy in the control of schistosomiasis in Mali. *Trop. Med. Parasitol.* 39, 167–174
13 Taylor, P. and Makura, O. (1985) Prevalence and distribution of schistosomiasis in Zimbabwe. *Ann. Trop. Med. Parasitol.* 79, 287–299
14 Crompton, D.W.T. and Tulley, J.J. (1987) How much ascariasis is there in Africa? *Parasitol. Today* 3, 123–127
15 Utroska, J.A. *et al.* (1989) *An Estimate of the Global Needs for Praziquantel within Schistosomiasis Control Programmes.* Geneva: WHO, Schsito/89
16 Bundy, D.A.P. *et al.* (1991) The epidemiological implications of a multiple-infection approach to the control of human helminth infections. *Trans. R. Soc. Trop. Med. Hyg.* 85, 274–276
17 Stoll, N.R. (1947) This wormy world. *J. Parasitol.* 33, 1–18
18 Doumenge, J.P. *et al.* (1987) *Atlas of the Global Distribution of Schistosomiasis.* Bordeaux: Universit de Bordeaux
19 Yoon, S.S. (1995) Geographical information systems: a new tool in the fight against schistosomiasis, in *The Added Value of Geographical Information Systems in Public and Environmental Health* (de Lepper, M.J.C. *et al.*, eds), pp 201–213, Academic Publishers and WHO
20 Nuttall, I. *et al.* (1998). Systemes d'information geographique et lutte contre les maladies tropicales. *Med. Trop.* 58, 221–227
21 Guyatt, H. *et al.* (1999) Can prevalence of infection in school-aged children be used as an index for assessing community prevalence. *Parasitology* 118, 257–268
22 Brooker, S. *et al.* practical application of an index to estimate the number of cases of intestinal nematode infections and schistosomiasis within a country: Cameroon as a case study. *Bull. WHO* (in press)
23 Thomson, M. *et al.* (2000) Environmental information for prediction of epidemics. *Parasitol. Today* 16, 137–138
24 Thomson, M.C. *et al.* (1997) Mapping malaria risk in Africa – what can satellite data contribute? *Parasitol. Today* 8, 313–318

Microsatellite Markers and Genetic Mapping in *Plasmodium falciparum*

M.T. Ferdig and X-z. Su

Whole-genome methods are changing the scope of biological questions that can be addressed in malaria research. In the rich context provided by Plasmodium falciparum *genome sequencing, genetic mapping is a powerful tool for identifying genes involved in parasite development, invasion, transmission and drug resistance. The recent development of a high-resolution* P. falciparum *linkage map consisting of hundreds of microsatellite markers will facilitate an integrated genomic approach to understanding the relationship between genetic variations and biological phenotypes. Here, Michael Ferdig and Xin-zhuan Su provide an overview for applying microsatellite markers and genetic maps to gene mapping, parasite typing and studies of parasite population changes.*

Michael Ferdig and **Xin-zhuan Su** are at the Laboratory of Parasitic Diseases, National Institutes of Health, 4 Center Drive 0425, Bethesda, MD 20892-0425, USA. **Tel: +1 301 496 4023, Fax: +1 301 402 0079, e-mail: xsu@helix.nih.gov**

The burgeoning data set emerging from the *Plasmodium falciparum* genome sequencing project is providing a foundation for advances against this human malaria parasite. Knowing gene sequences is a crucial step towards understanding the molecular mechanisms underlying such essential biological processes as development, transmission, disease pathogenesis and drug resistance, as well as identifying new drug and vaccine targets; however, even the completed genome of the *P. falciparum* line 3D7 currently being sequenced will not translate easily into a comprehensive recognition of genes and their roles in parasite biology. The *P. falciparum* genome sequence is not only that of the 3D7 isolate, but a collection of variant sequences from all *P. falciparum* parasites that share a common genomic theme. Therefore, complementary methods are required to define relationships between genetic variations (ie. sequence differences) and their consequent functional effects.

 0169-4758/00/$ – see front matter PII: S0169-4758(00)01676-8