Revista da Sociedade Brasileira de Medicina Tropical
35(5): 527-530, set-out, 2002.

Genetic diversity and differentiation in natural *Plasmodium falciparum* populations inferred by molecular typing of the merozoite surface proteins 1 and 2

Diversidade genética e diferenciação em populações naturais de *Plasmodium falciparum* inferidas pela tipagem molecular das proteínas de superfície de merozoítos 1 e 2

Fabrício J.T. Pereira[1], José A. Cordeiro[2], Erika H.E. Hoffmann[3] and Marcelo U. Ferreira[1 3]

Abstract *Genetic diversity and differentiation, inferred by typing the polymorphic genes coding for the merozoite surface proteins 1 (*Msp-1*) and 2 (*Msp-2*), were compared for 345 isolates belonging to seven* Plasmodium falciparum *populations from three continents. Both loci yielded similar estimates of genetic diversity for each population, but rather different patterns of between-population differentiation, suggesting that natural selection on these loci, rather than the transmission dynamics of* P. falciparum, *determines the variation in allele frequencies among populations.*

Key-words: *Malaria.* Plasmodium falciparum. Msp-1. Msp-2. *Population genetics.*

Resumo *Estimativas de diversidade e diferenciação genética, obtidas com a tipagem dos genes polimórficos que codificam as proteínas de superfície de merozoítos 1 (*Msp-1*) e 2 (*Msp-2*) foram comparadas em 345 isolados pertencentes a sete populações de* Plasmodium falciparum *de três continentes. Os dois* loci *produziram estimativas semelhantes de diversidade genética para cada população, porém com padrões distintos de diferenciação entre populações, sugerindo que a seleção natural nestes loci, em vez da dinâmica de transmissão de* P. falciparum, *determina a variação de freqüências alélicas entre populações.*

Palavras-chaves: *Malária.* Plasmodium falciparum. Msp-1. Msp-2. *Genética de populações.*

The genes encoding the polymorphic merozoite surface proteins 1 (*Msp-1*) and 2 (*Msp-2*) have been extensively used as markers to investigate the genetic diversity and population structure of *Plasmodium falciparum*[3]. Here we examine how comparable are estimates of genetic diversity and differentiation inferred by *Msp-1* and *Msp-2* typing in the same *P. falciparum* populations.

Polymorphism at the *Msp-1* and *Msp-2* loci was examined in 345 isolates belonging to seven *P. falciparum* populations. Five of them were from different sites within the Brazilian Amazon Basin, one from southern Vietnam and one from north-eastern Tanzania (Table 1). Five variable domains or blocks (2, 4a, 4b, 6 and 10) of the *Msp-1* gene were typed with PCR using the allele-specific primers described elsewhere[11]. *Msp-1* alleles were defined as unique combinations of allelic types in variable blocks[11] Most typing results, except for those for 34 additional isolates from Vietnam, had been previously published[7 8 14]. The central region of *Msp-2* (blocks 2 and 3) was amplified by PCR and typed by hybridization with allele-specific probes[9]. *Msp-2* alleles were defined according to PCR fragment sizes and hybridization patterns. *Msp-2* typing results had been

1. Laboratório de Parasitologia Molecular. 2. Departamento de Epidemiologia e Saúde Coletiva da Faculdade de Medicina de São José do Rio Preto, São José do Rio Preto,SP. 3. Departamento de Parasitologia do Instituto de Ciências Biomédicas da Universidade de São Paulo, São Paulo, SP.
Financial support: Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), UNDP/World Bank/World Health Organization Special Program for Research and Training in Tropical Diseases and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). FJTP and EHEH are recipients of studentships from FAPESP, and MUF is a recipient of a research fellowship from CNPq.
Address to: Dr. Marcelo U. Ferreira. Depto. de Parasitologia/ICB/USP. Av. Prof. Lineu Prestes 1374, 05508-900 São Paulo, SP, Brazil.
Tel: 55 11 3091-7273, Fax: 55 11 3091-7417.
e-mail: muferrei@usp.br
Recebido para publicação em 10/12/2001.

Table 1 - Geographical origin and estimates of genetic diversity at the Msp-1 and Msp-2 loci for seven Plasmodium falciparum populations

			Msp-1		Msp-2	
Population (origin)	Years of collection	No. of isolates typed	No. of alleles	*H**	No. of alleles	*H**
Brazil						
RO1 (Rondônia)	1985-86	49	10	0.822	12	0.889
PA1 (Pará and Amapá)	1985-86	17	6	0.785	7	0.797
PA2 (Pará and Amapá)	1992-95	17	8	0.768	8	0.938
RO2 (Rondônia)	1995-98	46	9	0.786	16	0.896
AC (Acre)	1999	26	5	0.566	9	0.818
Vietnam						
VN (Bao Loc)	1996-97	126	19	0.884	32	0.942
Tanzania						
TZ (Tanga)	1996	64	13	0.855	35	0.953

*The average gene diversity index H is calculated as $H = [n/(n-1)]\,[1 - \Sigma p_i^2]$, where n is the number of isolates sampled and p_i is the frequency of each allele at a given locus[4].

previously published for RO1, RO2, VN and TZ[9]; 60 additional Brazilian isolates, which constitute the populations PA1, PA2 and AC, were analyzed for the present comparison (Table 1).

Both loci yielded comparable estimates of genetic diversity for each population, as judged by the number of different alleles and the average diversity index *H*[4], defined as the probability of randomly drawing two different alleles from the population sample (Table 1). In other words, both typing systems seemed to be equivalent in terms of their ability to discriminate between unrelated parasite clones or strains within the same population[10].

Different results were obtained when we used allele frequencies to assess the degree of genetic differentiation among the seven *P. falciparum* populations. Fixation indices (F_{ST}), which describe the proportion of overall genetic diversity that is attributable to differences between (instead of within) populations[13], were calculated for pair-wise comparisons of allele frequencies for both loci. AC was the population associated with the highest fixation indices for *Msp-1* (Table 2), suggesting that gene flow between AC and all other populations is relatively scant. Fixation indices for *Msp-1* suggest that more differentiation (less genetic admixture) has occurred between AC and RO2 (F_{ST} = 0.130), which were sampled in the 1990s in two neighboring regions of Brazil (distance between collection sites of about 600km), than between parasites on different continents (F_{ST} values between 0.010 and 0.119 in pairwise comparisons). The overall F_{ST} estimate for *Msp-1* was 0.095. In contrast, less genetic differentiation between populations was usually detected by *Msp-2* typing (Table 2), with an overall F_{ST} estimate of 0.072. No association between F_{ST} values and

Table 2 - Matrix of pair-wise F_{ST} values for Msp-1 (above diagonal) and Msp-2 (below diagonal) in comparisons of seven* Plasmodium falciparum *populations.*

	RO1	PA1	PA2	RO2	AC	VN	TZ
RO1	-	0.013	0.052	0.015	0.088	0.060	0.055
PA1	0.069	-	0.038	0.016	0.093	0.079	0.058
PA2	0.025	0.046	-	0.035	0.193	0.064	0.039
RO2	0.024	0.069	0.029	-	0.130	0.055	0.010
AC	0.068	0.073	0.046	0.062	-	0.119	0.092
VN	0.028	0.033	0.014	0.031	0.041	-	0.031
TZ	0.024	0.041	0.016	0.027	0.042	0.004	-

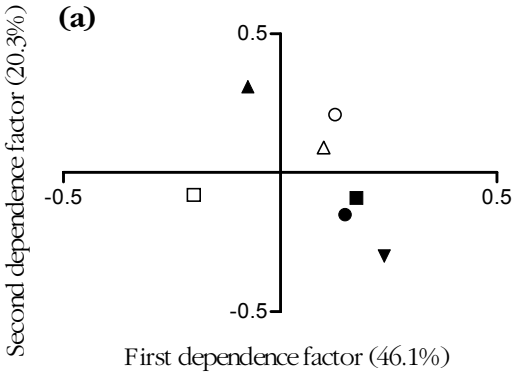
*The fixation index (F_{ST}), ranging between 0 and 1, estimates the proportion of the overall variation in allele frequencies that is due to variation *between* (instead of *within*) populations; high F_{ST} values indicate high levels of genetic differentiation (i. e., low levels of gene flow and genetic admixture) between pairs of populations[4].

geographical distance was apparent for either locus. These findings suggest that natural selection on these loci, rather than the transmission dynamics of *P. falciparum,* determines the observed patterns of between-population variation in allele frequencies. Thus, they argue against the use of non-neutral polymorphic markers (i. e., markers potentially under selection pressure, such as those encoding antigens) in studies of population genetics of malaria parasites.

An alternative way of looking at these data was based on analysis of dependence (ANADEP)[6]. ANADEP is a geometrical method of analysis. Allele frequencies of each population are represented as a point in the real space whose dimension is the total number of alleles detectable by the typing procedure[12]. ANADEP may be thus regarded as an equivalent to principal component analysis, a method of multivariate analysis that is extensively used in Biology and Epidemiology. In contrast with principal

component analysis, however, ANADEP is suitable for use with categorical variables. The first two factors of dependence, which together explain nearly two-thirds of the overall dependence among populations, were considered to describe the dispersion of populations, according to allele frequencies, in a two-dimensional space (Figure 1). *Msp-1* (but not *Msp-2)* allele frequencies clearly differentiated AC, VN and PA2 from all other populations. Conversely, RO1 and RO2 were well differentiated from all other populations by considering *Msp-2* (but not *Msp-1*) allele frequencies (Figure 1). A quite similar overall pattern emerged when the three first factors of ANADEP, which together explain between 74% and 77% of the overall dependence, where taken into account (data not shown). The overall ability to detect genetic differentiation among populations, as shown by their dispersion in Figure 1, was considerably higher for *Msp-1* than for *Msp-2*.

Describing the patterns of genetic differentiation among parasite populations is of considerable interest for population geneticists[1 5]. Here we illustrate some limitations that are inherent to studies of genetic relationships of malaria parasite populations based on non-neutral markers, such as the widely used *Msp-1* and *Msp-2* antigenic loci. In short, these markers may differ in their ability to discriminate *between* populations, even when they are equivalent in their ability to discriminate between clones and strains *within* the same population, especially when natural selection is a major source of variation in allele frequencies. Putatively neutral markers (i. e., polymorphisms located in non-coding DNA sequences, that are not under selection), such as most hypervariable microsatellite loci, however, provide geneticists with excellent tools to overcome these limitations[2].



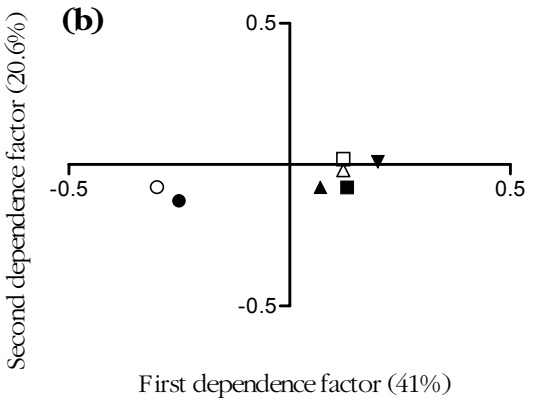


Figure 1 - Patterns of differentiation of seven Plasmodium falciparum *populations, as estimated by analysis of dependence of allele frequencies of* Msp-1 *(a) and* Msp-2 *(b). The first two dependence factors, which together explain about two thirds of the overall dependence, are considered in this analysis; the percentage of overall dependence explained by each factor is given in parentheses. Symbols for populations are as follows: RO1 ○; PA1 ■; PA2 ▲; RO2 ●; AC ▼; VN □, and TZ △. Populations are described in Table 1.*

REFERENCES

1. Anderson TJC, Haubold B, Williams JT, Estrada-Franco JG, Richardson L, Mollinedo R, Mockarie M, Mokili J, Mharakurwa S, French N, Whitworth J, Velez ID, Brockman AH, Nosten F, Ferreira MU, Day KP. Microsatellites reveal a spectrum of population structures in the malaria parasite *Plasmodium falciparum.* Molecular Biology and Evolution 17: 1467-1482, 2000.
2. Awadalla P, Walliker D, Babiker H, Mackinnon M. The question of *Plasmodium falciparum* population structure. Trends in Parasitology 17:351-353, 2001.
3. Babiker H, Walliker D. Current views on the population structure of *Plasmodium falciparum:* implications for control. Parasitology Today 13: 262-267, 1997.
4. Conway DJ. Population genetics. *In:* Thompson RCA (ed) Molecular Epidemiology of Infectious Diseases, Arnold, London, p. 45-58, 2000.
5. Conway DJ, Cavanagh DR, Tanabe K, Roper C, Mikes ZS, Sakihama N, Bojang KA, Oduola AMJ, Kremsner PG, Arnot DE, Greenwood BM, Mcbride JS. A principal target of human immunity to malaria identified by molecular population genetic and immunological analyses. Nature Medicine 6: 689-692, 2000.
6. Cordeiro JA. Analysis of Dependence. Technical Report, Instituto de Matemática da Universidade Estadual de Campinas, Campinas, 1987.
7. Ferreira MU, Liu Q, Kimura M, Ndawi BT, Tanabe K, Kawamoto F. Allelic diversity in the merozoite surface protein-1 and epidemiology of multiple-clone infections in northern Tanzania. Journal of Parasitology 84: 1286-1289, 1998.
8. Ferreira MU, Liu Q, Zhou M, Kimura M, Kaneko O, Thien HV, Isomura S, Tanabe K, Kawamoto F. Stable patterns of allelic diversity at the merozoite surface protein-1 locus of *Plasmodium falciparum* in clinical isolates from southern Vietnam. Journal of Eukaryotic Microbiology 45: 131-136, 1998.
9. Hoffmann EHE, Silveira LA, Tonhosolo R, Pereira FJT, Ribeiro WL, Tonon AP, Kawamoto F, Ferreira MU. Geographical patterns of allelic diversity in the *Plasmodium falciparum* malaria-vaccine candidate, merozoite surface protein-2. Annals of Tropical Medicine and Parasitology 95: 117-132, 2001.
10. Hunter PR, Gaston MA. Numerical index of the discriminatory ability of typing systems: an application of Simpson's index of diversity. Journal of Clinical Microbiology 26: 2465-2466, 1988.
11. Kaneko O, Kimura M, Kawamoto F, Ferreira MU, Tanabe K. *Plasmodium falciparum:* variation in the merozoite surface protein 1 gene in wild isolates from southern Vietnam. Experimental Parasitology 86: 45-57, 1997.
12. Lapenta AS, Bicudo HEMC, Ceron CE, Cordeiro JA. Esterase patterns and phylogenetic relationships of species and strains included in the *Drosophila buzzati* cluster. Cytobios 96: 95-107, 1998.
13. Nei M, Chesser RK. Estimation of fixation indices and gene diversities. Annals of Human Genetics 47: 253-259, 1983.
14. Silva NS, Silveira LA, Machado RLD, Póvoa MM, Ferreira MU. Temporal and spatial distribution of variants of merozoite surface protein-1 (MSP-1) in *Plasmodium falciparum* populations in Brazil. Annals of Tropical Medicine and Parasitology 94: 675-688, 2000.