Physica A 387 (2008) 6137–6150

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Global analysis of the immune response

Leonardo C. Ribeiro [a], Ronald Dickman [a], Américo T. Bernardes [b,*]

[a] Departamento de Física, ICEx, Universidade Federal de Minas Gerais, 30161-970, Belo Horizonte - MG, Brazil

[b] Departamento de Física, ICEB, Universidade Federal de Ouro Preto, 35400-000, Ouro Preto - MG, Brazil

ARTICLE INFO

Article history:
Received 19 June 2008
Available online 27 June 2008

PACS:
87.10.+e
87.16.Yc
87.16.Ac
05.10.-a

Keywords:
Immune system
Scale free networks
Global response
Hurst exponent

ABSTRACT

The immune system may be seen as a complex system, characterized using tools developed in the study of such systems, for example, surface roughness and its associated Hurst exponent. We analyze densitometric (Panama blot) profiles of immune reactivity, to classify individuals into groups with similar roughness statistics. We focus on a population of individuals living in a region in which malaria endemic, as well as a control group from a disease-free region. Our analysis groups individuals according to the presence, or absence, of malaria symptoms and number of malaria manifestations. Applied to the Panama blot data, our method proves more effective at discriminating between groups than principal-components analysis or super-paramagnetic clustering. Our findings provide evidence that some phenomena observed in the immune system can be only understood from a global point of view. We observe similar tendencies between experimental immune profiles and those of artificial profiles, obtained from an immune network model. The statistical entropy of the experimental profiles is found to exhibit variations similar to those observed in the Hurst exponent.

1. Introduction

In recent decades, theories dealing with the complexity of the immune system have been proposed [1,2]. The focus in these works is on modeling the immune response. Despite the remarkable specificity of immune receptors to their targets, there is much evidence that the behavior of the immune system is not restricted to one-to-one antigen-antibody responses. Collective immune phenomena, such as tolerance, autoimmune diseases, and memory, among others, are far from well understood [3–9]

The immune system is formed of a huge number of components, with diverse interactions between them. Networks of interacting elements can be used to model some features of its dynamics. In these models, the elements (representing, for example, lymphocytes), interact via a set of well defined rules. It is important to note that most of the activity of the system is elicited not by foreign agents, but by self agents, i.e., molecules native to the organism.

The immune system may be seen as a complex system, that is, one composed of a large number of elements, with specific interactions. The interactions are often simple, though this need not always be the case. The dynamics of the system as a whole exhibit features (“emergent properties”) not found in the behavior of each element. Emergent properties and self-organized behavior are common characteristics of complex systems.

Many complex systems may be characterized by trajectories or, in some cases, fractal, n-dimensional surfaces in a given space; a key global property of such a surface is its roughness statistics [10,11]. One hopes, by analyzing the roughness, to classify different regimes of behavior. Analysis of surfaces generated by apparently unrelated processes, may also help to identify underlying similarities in their governing equations or dynamical rules [12–14].

* Corresponding author.
E-mail address: atb@iceb.ufop.br (A.T. Bernardes).

0378-4371/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.physa.2008.06.031

To apply this approach to global properties of the immune response, one requires an overall picture of the immune repertoire. We do this by studying a large set of reactivities against another (large) set of antigens, derived from tissue extracts. The experimental technique generating information on the set of reactivities is called *immunobloting*. The results of this technique is usually analyzed through multiparametric statistical analysis.

One of the problems in describing the immune system as a multiconnected network, is how to deal with its global connectivity. A bold attempt in this direction was made by the inventors of the Panama blot assay [15,16]. A modified immunoblot assay, the Panama blot, is not aimed at the detection or quantification of a specific antigen-antibody reaction. Rather, it permits comparison of patterns of global reactivity of natural serum immunoglobulins with complex antigenic mixtures. Antigenic extracts consisting of whole organs (brain, liver, muscle, etc.) or whole bacterial cultures (such as Escherichia coli), have been used to generate these mixtures. The reaction of serum immunoglobulins with these mixtures results in activity profiles, which are used to compare patterns of different individuals, or the history of a given individual, rather than to detect or measure individual immunological reactions, as is customary in serology.

Traditionally, those profiles are described by points in a n-dimensional space, where the set of coordinates corresponds to the rescaled peaks of each profile. The dimensionality of the space is then reduced, using principal component analysis (PCA), and clusters of similar individuals are identified via statistical analysis. In the present work, we introduce an approach to study the densitometric profiles, and arrange individuals in groups, based on roughness statistics. We focus on a population of individuals from a region in which malaria is endemic. The individuals of this population can be classified into several groups, according to the presence or absence of malaria symptoms, number of malarial manifestations, etc. A control population, from a malaria-free region, is used to define reference characteristics of the reactivity profiles.

In our approach, the densitometric profile is represented as a function $h(x)$, where the variable x represents position on the gel (a function of protein molecular weight), and h reflects the reactivity at this position. The functions $h(x)$ corresponding to each individual are then analyzed, using several techniques: PCA, superparamagnetic clustering, and analysis of roughness, leading to calculation of the Hurst exponent of the profile. We show that the Hurst exponent, calculated from single parametric statistical analysis, separates individuals from malaria-endemic and malaria-free regions. We compare our method with the traditional one and with a clustering technique introduced by Domany et al [17], based on phase transitions in a Potts model, with many states and distance-dependent interactions. Our method shows better discrimination than the other two studied. More importantly, the success of profile-based methods highlights the utility of global analysis of the immune system, and suggests the presence of collective phenomena in this system.

2. Materials and methods

Immune reactivity profiling, via the Panama Blot method, was performed on blood collected from a set of 78 individuals having similar lifestyles, living in a region of the State of Mato Grosso, Brazil, in which malaria is endemic. Individuals living in this region are continuously exposed to the malaria parasite. For comparison, blood from a group of 10 individuals residing in Belo Horizonte (State of Minas Gerais, Brazil), a region in which malaria is not endemic, was also analyzed. The individuals in the endemic region are classified into groups A–F, according to their history of malaria exposure:

A: those whose blood shows no evidence of the parasite;
B: those that have presented symptoms of malaria, between 1 to 10 times, but were asymptomatic at the time of blood collection;
C: those that have presented symptoms more than 10 times, but were asymptomatic at the time of blood collection;
D: those symptomatic at the time of blood collection;
E: those with the parasite in their blood, but asymptomatic for the disease at the time of collection.

The individuals from the non-endemic region are characterized as:

F: those never having been exposed to the parasite.

Details about the experimental procedures can be found in Ref. [18].

2.1. Experiment: Panama Blot

Immune profiles were obtained through reactions between a population of immunoglobulins, and an extract containing a large number of proteins or peptides (protein fragments). In the Panama Blot method, the protein extract is generally prepared from a cell culture or tissue; in the present case it was prepared from human liver or brain (from cadavers). The extract is dissolved in a pH-adjusted solution, so that the proteins acquire electric charge, necessary for electrophoresis [19]. The solution is poured over a polyacrylamide gel confined between glass plates, and subject to a constant electric field in the plane of the plates. The proteins then migrate within the gel, with velocities depending on their molecular weight and charge.

Once the extract proteins have been separated via electrophoresis, they are transferred, by application of an electric field perpendicular to the plates, to a nitrocellulose membrane, for subsequent reaction with serum (Fig. 1). The membrane, with the proteins adsorbed on its surface, is placed on a mold bearing a series of narrow, parallel channels. The channels are then loaded with serum samples, each from a different individual, and allowed to react with the extract proteins for a period of four hours.

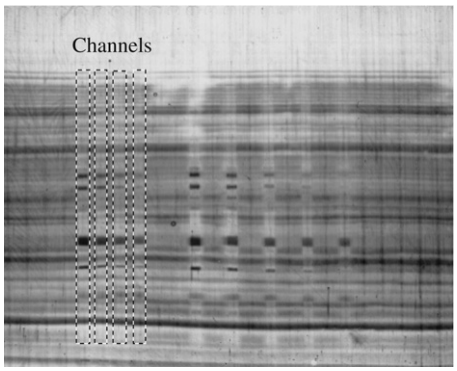

Fig. 1. Stained proteins following separation and fixation in the nitrocellulose membrane (from Ref. [18]).

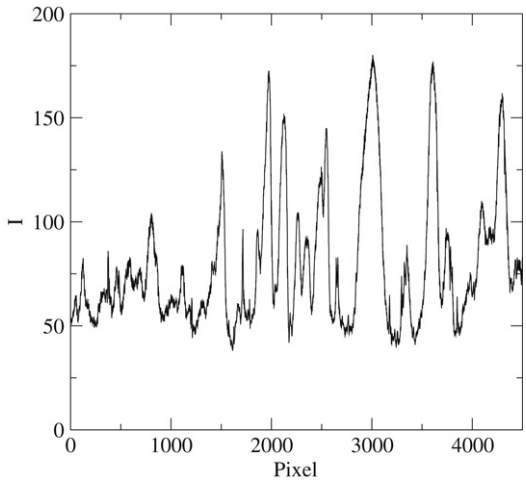

Fig. 2. Digitized signal representing the gray scale of a single channel. Large (small) *y* corresponds to dark (light) color in the channel, while *x* represents horizontal position, which depends on molecular weight as described in text.

Following the reaction, the membrane is washed, leaving behind only the immunoglobulins that reacted with the extract. In order to determine the position and intensity of reactions, the immunoglobulins adsorbed to the membrane are stained. This is done by allowing proteins with an affinity for IgM, and bearing an attached alkaline phosphatase, to associate with the adsorbed Ig. Following this, a stain that attaches to the alkaline phosphatase is added, and the sample is incubated for three to five minutes. (The choice of a marker protein with affinity to IgM is because this Ig class embodies the primary immune response to the malaria parasite.) Once stained, the positions and intensities of IgM-extract protein reactions in each channel are evident visually; the channels are scanned for quantitative analysis of the reaction profile. Finally, all extract proteins are stained using *Protogold*, as is customarily done in Panama-blot experiments. The light-and-dark pattern to be analyzed therefore consists of the sum of two signals: (1) local IgM concentration (the alkaline phosphatase stain, which is dark purple), and (2) local extract-protein concentration (the Protogold stain, which is rose colored). (Note however that the Protogold signal is essentially the same for each of the channels.)

All the steps described above were performed by the Immunology Group of the ICB/UFMG, cited in Ref. [18].

2.2. Digitization

In order to apply data analysis techniques, we digitize the dark-and-light pattern of each channel, using a scanner (Hewlett–Packard model Scanjet 8350), with high resolution (2400 d.p.i.) and color depth. This device avoids the distortions that typically arise in common scanners, due to interpolation and to poor discrimination of light intensities. (The scanner used in this study has a color depth of 4.6.) The scanner generates a file in which the light intensity of each pixel is given as a number between 0 and 255, where 0 corresponds to white and 255 to black. The intensities of the pixels corresponding to a given position in a channel are then averaged (Fig. 2). (Each channel is approximately 100 pixels wide.) This profile, obtained via digitization, is somewhat noisy, due to agglomeration of proteins on the membrane, and to the intrinsic noise of the scanner. We therefore smooth the profile by performing a running average over windows of ten pixels.

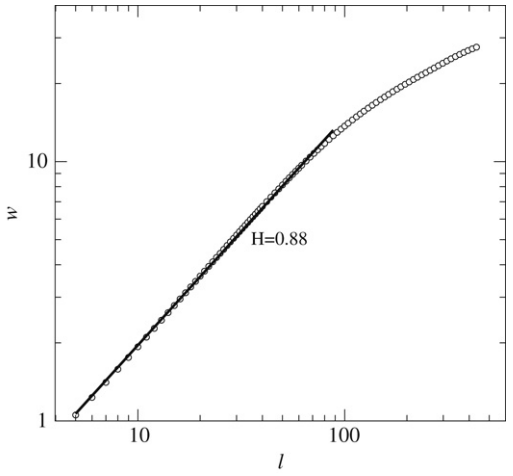

Fig. 3. Roughness $w(\ell)$ calculated from smoothed profiles.

3. Data analysis

3.1. Hurst exponent

In this analysis, we treat the reactivity profile h_j as if it represented the height of a one-dimensional surface, and analyze it using an approach developed in the study of scaling properties of random surfaces [20,21]. The first step is to calculate the roughness w_k in a window of length ℓ, starting at pixel k, defined via

$$w_k^2 = \sum_{j=k}^{k+\ell-1} (h_j - \overline{h})^2, \tag{1}$$

where $\overline{h}$ is the mean height over the ℓ points in the window. The mean of w_k^2 over all starting points k defines $w^2(\ell)$, characterizing the roughness on a scale ℓ. As is common in analysis of rough interfaces or random signals, we find that for small ℓ, the roughness follows a power law:

$$w^2(\ell) \propto \ell^2 H \tag{2}$$

with H usually referred to as the Hurst exponent. Typical results for the roughness $w(\ell)$ are shown in Fig. 3.

A fractal curve is a profile following the scaling of Eq. (2) over an unlimited range of scales ℓ. In practice, power-law scaling over two or more decades is generally interpreted as evidence of fractal behavior. In the present case, we observe power-law behavior of $w(\ell)$ for ℓ between about 10 and 100. Although the profile h_j should not, therefore, be characterized as a fractal, we may nevertheless hope to extract useful information from the value of the Hurst exponent.

In particular, we study the values of H found for each group of individuals, classified according to their history of malaria exposure, as discussed above. The range of Hurst exponent values for the various groups is displayed in Fig. 4. In this set of studies, the reactivity profiles were obtained using a protein extract derived from human liver. In another set of studies, using a different set of individuals, a protein extract derived from human brain was used; the results for the Hurst exponent are shown in Fig. 5.

It is evident, in Figs. 4 and 5 that the Hurst exponent values for individuals never exposed to the malaria parasite, differ significantly from those of the other groups, in which all individuals have been exposed to the parasite. The exposure to the parasite in some way modifies the reactivities of the immune system in relation to the proteins present in the extract, in a manner that manifests itself through an altered value of the Hurst exponent.

In an effort to relate the values of the Hurst exponent with changes in the reactivity profiles, we applied the roughness analysis for a set of artificial profiles, $\tilde{h}_j$. The latter are generated by summing a number of Gaussian functions:

$$\tilde{h}_j = \sum_{i=1}^{M} A_i \exp\left[-\frac{(j-\mu_i)^2}{2\sigma_i^2}\right] + \xi_j \tag{3}$$

with mean μ_i and variance σ_i^2. The means μ_i are uniformly distributed over the profile ($1 \le \mu_i \le 4500$), while the variances are uniformly distributed on the interval [0, 120]. The maximum A_i of each Gaussian is proportional to its variance, $A_i = c\sigma_i^2$. The "noise" terms ξ_j are independent Gaussian random variables with zero mean and variance 11, much smaller than the amplitudes A_i. (The noise is included as a test of our smoothing procedure.) The number M of Gaussian terms, the range of

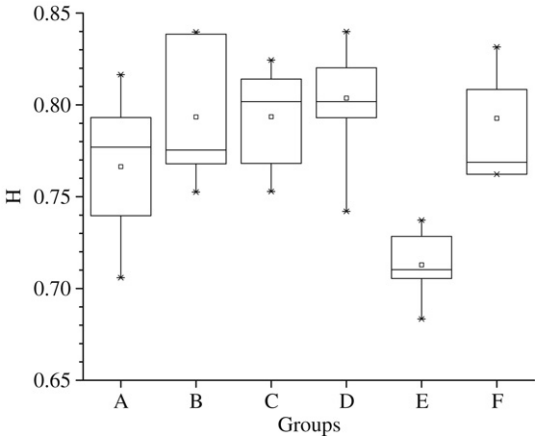

Fig. 4. Hurst exponent of the profiles obtained using liver extract, separated according to history of malaria exposure. Horizontal bars denote the largest and the smallest exponent values; boxes indicate the interval containing 75% of the data points. The sample is composed of 7 individuals from group A; 8 from B; 8 from C; 8 from D; 9 from E and 7 from F.

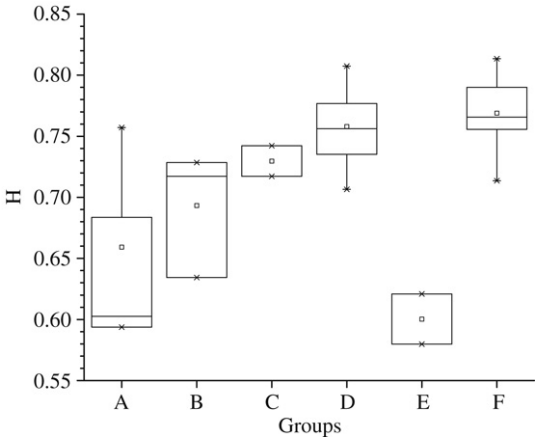

Fig. 5. Hurst exponent of the profiles obtained using brain extract, as in Fig. 4. The data are composed of 4 individuals from group A; 5 from B; 5 from C; 10 from D; 12 from E and 3 from F.

the variance and the proportionality constant c are chosen to reproduce the characteristics of a typical digitized profile, as judged by eye (Fig. 6).

To test our smoothing procedure, we calculate (for a set of 50 artificial profiles) the "processed" noise, obtained via the smoothing and subtraction procedure described above. The probability distribution of this processed noise (represented by a histogram), is then compared with the original Gaussian distribution. With a proper choice of smoothing interval, the processed noise distribution shows good agreement with the original one (Fig. 7), allowing us to be confident that only noise is removed in the smoothing process, leaving the immunological information unchanged.

We calculate the scale-dependent roughness $w(\ell)$ of smoothed artificial profiles, as was done for the experimental profiles. The experimental and artificial functions w exhibit the same qualitative features: linear growth for small ℓ and saturation at larger scales, $\ell > ell_s$ (Fig. 8).

Varying the interval over which the variances σ_i^2 are distributed, and the value of coefficient c, we find that the saturation scale $\ell_s \propto \langle\sigma_i\rangle$, the mean of the standard deviations of the Gaussians used to construct the profile. (This tendency was verified in hundreds of generated profiles.) ℓ_s is a measure of the typical width of peaks in the profile. Note that the Hurst exponent is calculated using the data for $\ell < \ell_s$.

3.2. Principal components analysis

In immunology, the method known as principal components analysis (PCA) is commonly applied to reactivity profiles [22]. Starting with a d-dimensional data set, PCA identifies the set of directions (i.e., linear combinations of the d original components), exhibiting the greatest correlation; the data are then projected onto this space of reduced dimensionality. PCA facilitates subsequent statistical analysis and/or visualization of the data in a manner that highlights correlations.

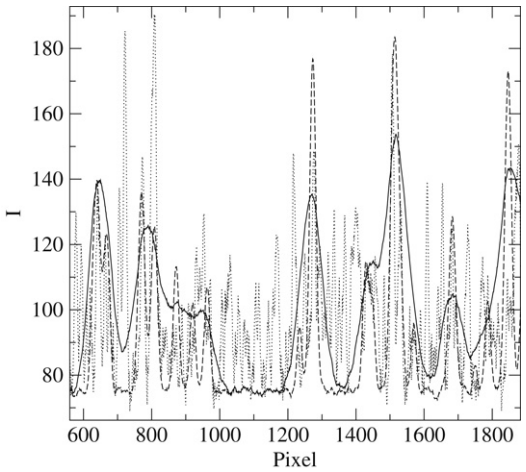

Fig. 6. Three artificial profiles, generated using different average variances: 30 (dotted line), 60 (dashed line) and 90 (continuous line). The curves have been smoothed via the same procedure as applied to experimental profiles.

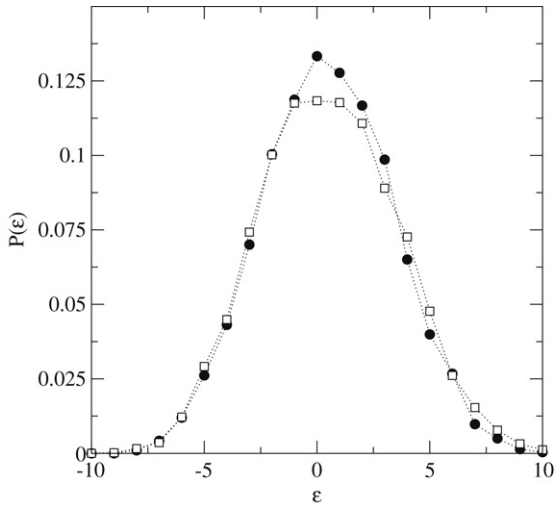

Fig. 7. Comparison of probability distributions of processed (squares) and original noise (circles) variables in artificial profiles.

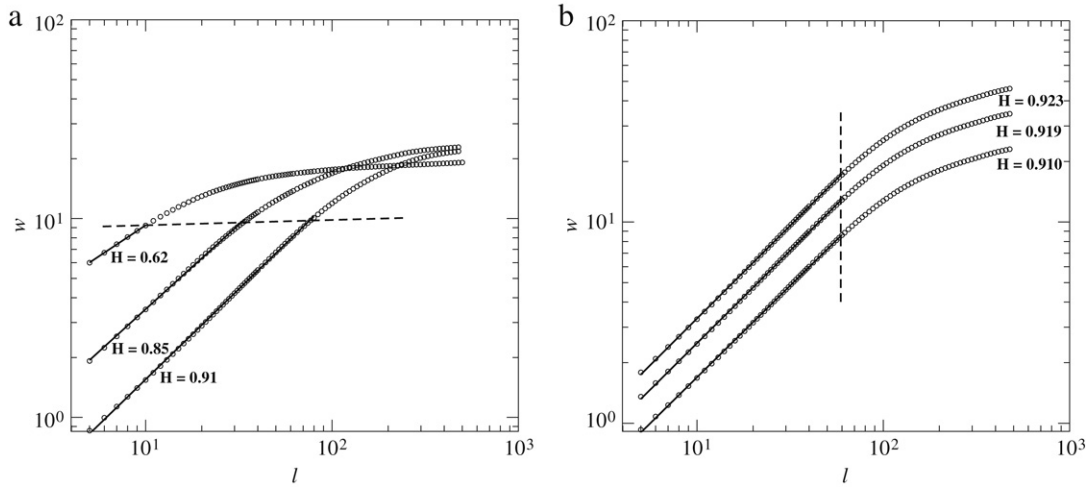

Fig. 8. (a) Roughness for profiles with different variances (30, 60 and 90); (b) roughness for profiles with different amplitudes (200, 600 and 900).

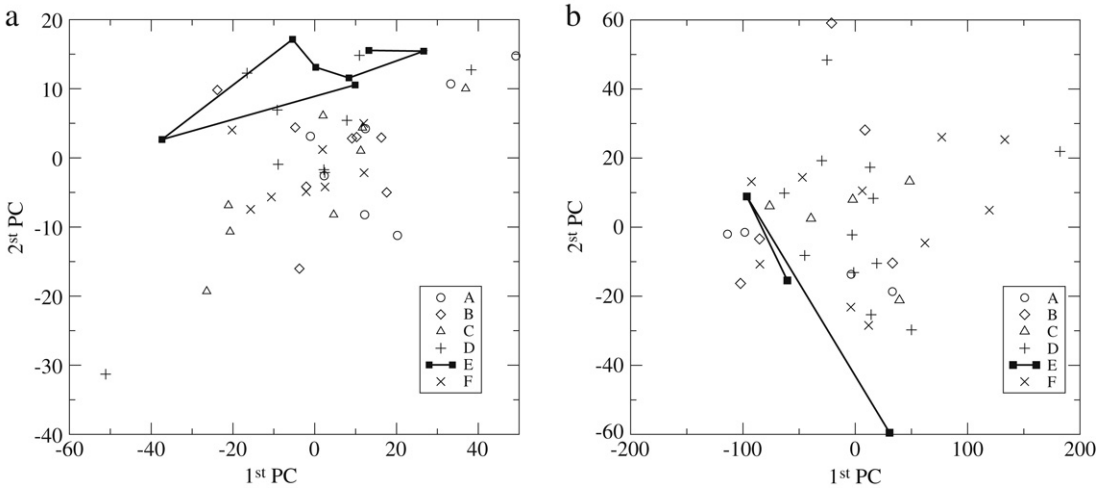

Fig. 9. (a) Projection of points from profiles using liver extract, onto first and second principal axes of the PCA analysis. (b) similar projection of points from profiles obtained using brain extract. Labels A–F identify the groups as discussed in text. The lines are guides to facilitate visualization of the set of points belonging to the same group.

Given a set of N observations, represented as d-dimensional vectors, $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N$, we form the $d \times d$ *covariance matrix*,

$$\begin{pmatrix} \mathrm{Cov}(\mathbf{x}_1, \mathbf{x}_1) & \mathrm{Cov}(\mathbf{x}_1, \mathbf{x}_2) & \ldots & \mathrm{Cov}(\mathbf{x}_1, \mathbf{x}_d) \\ \mathrm{Cov}(\mathbf{x}_2, \mathbf{x}_1) & \mathrm{Cov}(\mathbf{x}_2, \mathbf{x}_2) & \ldots & \mathrm{Cov}(\mathbf{x}_2, \mathbf{x}_d) \\ \vdots & \vdots & & \vdots \\ \mathrm{Cov}(\mathbf{x}_d, \mathbf{x}_1) & \mathrm{Cov}(\mathbf{x}_d, \mathbf{x}_2) & \ldots & \mathrm{Cov}(\mathbf{x}_d, \mathbf{x}_d) \end{pmatrix}. \tag{4}$$

Since the covariance matrix is Hermitian, it can be diagonalized. The directions defined by the eigenvectors associated with the largest eigenvalues then represent the greatest degree of correlation. The "principal components" of the original vectors are then projected onto this set of directions. (Typically, one projects onto the two or three directions having the largest eigenvalues.)

To apply PCA to the reactivity profiles, we represent each profile as a point in $\mathbb{R}^d$, where the ith component corresponds to the height of ith peak. Thus d is the number of identifiable peaks in the profile; the number of peaks in the experiments using liver extract was 19, while for brain extract the number was 15. These data points are projected onto the plane associated with the two largest eigenvalues of the covariance matrix. In this projection (see Fig. 9) the points associated with individuals from the endemic and non-endemic regions are well separated in both the experiments, using liver and brain extracts.

3.3. Superparamagnetic clustering

The super-paramagnetic clustering (SPC) technique is based on the physical properties of an inhomogeneous ferromagnet [17,23,24]. Unlike other clustering methods, SPC does not require any assumptions regarding the underlying probability distribution. In this method, to each data point, $\mathbf{r}_i \in \mathbb{R}^d$ is assigned a Potts spin variable s_i (an integer in the set $\{1., , , .q\}$). In this way we define a q-state Potts model, with Hamiltonian

$$H = -\sum_{1 \le i < j \le N} J_{ij} \delta_{s_i s_j}, \tag{5}$$

with an interaction $J_{ij} \ge 0$ between points i and j, whose strength decreases with distance $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$, as specified below.

Blatt, Wiesman and Domany have shown that for typical data sets, this model exhibits three phases: a disordered, high-temperature phase, a low-temperature, ferromagnetic phase, with a majority of spins sharing the same state, and, at intermediate temperatures, a *superparamagnetic* (SP) phase, characterized by an elevated susceptibility. The correlation, in the superparamagnetic phase, between spins at two different sites determines whether the associated data points are placed in the same or distinct clusters. To see this, consider the (unlikely) case in which all pairs of data points lie at roughly the same distance, one from another. (Recall that there is no intrinsic minimum distance.) In this case, each spin interacts essentially equally with all others, and we expect a transition directly from the ordered to the disordered phase, as temperature is increased. Consider now a case in which the data points fall in two or more well separated regions, each containing many data points, such that points in the same region have a separation $\sim r_0$, but the distance between distinct regions is $\gg r_0$. As we increase the temperature, the alignment of different regions will be lost, but we expect the alignment within each region to persist until some higher temperature is reached. This is the superparamagnetic phase: spins belonging to each region are highly correlated, but the orientations of the various regions are uncorrelated.

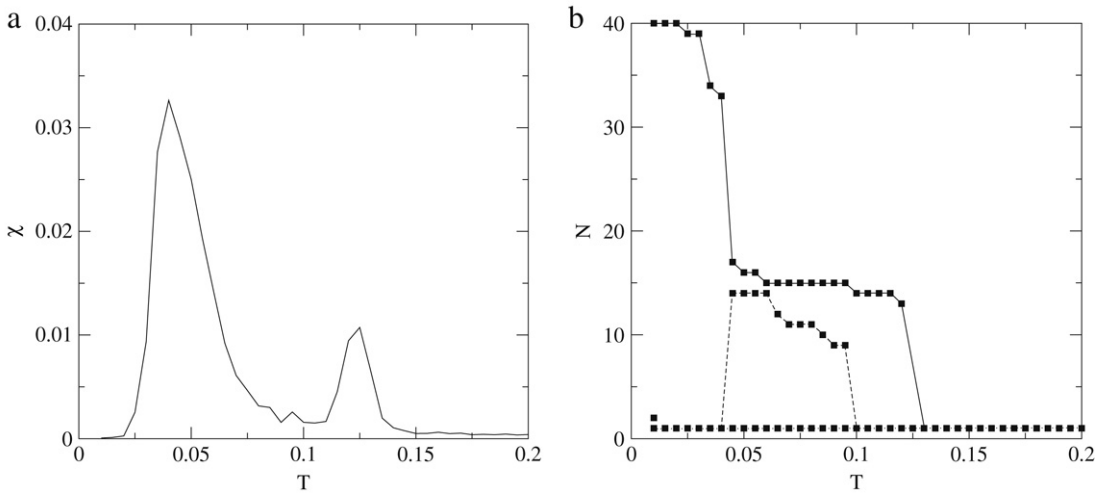

Fig. 10. (a) Susceptibility observed using SPC applied to the brain extract data. (b) Sizes of the three largest groups generated by SPC. During the phase transition, the largest group splits in two, which persist throughout the super-paramagnetic phase.

The onset of the SP phase is signaled by changes in the susceptibility, χ. At very low and very high temperatures, χ tends to zero; the ferromagnetic-superparamagnetic transition is marked by a susceptibility maximum. For temperatures within the SP phase, susceptibility remains relatively high, as correlated regions function as "super-spins". We use the spin–spin correlation function $C_{ij} = \langle s_i s_j \rangle$, (calculated in the SP phase, in a Monte Carlo simulation using the Swendsen–Wang algorithm [25]), to partition the spins into clusters. In the super-paramagnetic phase, the correlation between spins belonging to the same group tends to $1 - \frac{2}{q}$, whereas the correlation between spins belonging to different groups is $\simeq \frac{1}{q}$ [23]. One this basis, pairs of points (i, j) with high correlation, $C_{ij} > C_0$, are assigned to the same cluster. (Here C_0 is some reasonably large correlation; in the present study we use $C_0 = 0.9$) A point j with $C_{ij} < C_0$ for all $i \neq j$ forms a "cluster of one."

To implement SPC, one must define the interactions J_{ij}. While it is in principle possible to allow nonzero interactions between all pairs of points, in practice it is much more efficient to limit the number of neighbors a given point interacts with. Following Ref. [17], to define the neighbors of point i, we first determine the set $\mathcal{K}_i$ of K points nearest i. Then for each point $j \in \mathcal{K}_i$, we construct the set $\mathcal{K}_j$. Points i and j are neighbors if $i \in \mathcal{K}_j$ *and* $j \in \mathcal{K}_i$. We implement this *mutual neighbor* scheme using $K = 10$, which is sufficient to generate a connected graph, when applied to our data sets.

The interaction J_{ij} between pairs of neighboring points depends on the distance d_{ij} separating them,

$$J_{ij} = \frac{1}{\bar{k}} \exp\left(-\frac{d_{ij}^2}{2\bar{d}^2}\right) \tag{6}$$

where d is the mean distance (over all pairs of neighboring points), and k is the average number of neighbors per point. (Studies using $J_{ij} \propto \exp(-d_{ij}/\bar{d})$ yielded qualitatively similar results.)

We apply the technique of super-paramagnetic clustering, using $q = 50$ Potts states, in the same d-dimensional space as the PCA analysis, where each point corresponds to a simplified profile. For the profiles obtained using brain extract (Fig. 10a), we find evidence of the super-paramagnetic phase in the form of a plateau in the susceptibility. In this phase, the data points are grouped into two correlated sets, as can be seen in Fig. 10b. In the case of the liver-extract data, we do not observe a SP phase; the susceptibility (Fig. 11a) does not show the plateau characteristic of this phase. This result appears to reflect the rather homogeneous distribution of points.

The two groups identified in the brain-extract data are visualized in Fig. 12, for temperature 0.095. The groups identified in the SPC analysis correspond roughly to those evident on the right- and left-hand sides of the PCA projection in this figure. The two groups do not however correspond to distinct sets of individuals classified according to the history of malaria exposure. There is evidence that the two groups identified via PCA and SPC techniques, in the brain-extract data, may be correlated to history of exposure to hepatitis B [26].

4. Model immune network profiles

Recently we proposed a model of B cell interactions on scale-free and random (Erdös–Renyi) networks [27]. We model the population dynamics of B cell clones, including a plausible set of effective interactions (which could arise through the action of T cells), among such clones. Each node of the network corresponds to a distinct B cell clone, or to a species of ligand (i.e., a molecule that binds to specific B cell receptors). We use this system to generate model immune profiles. (Details

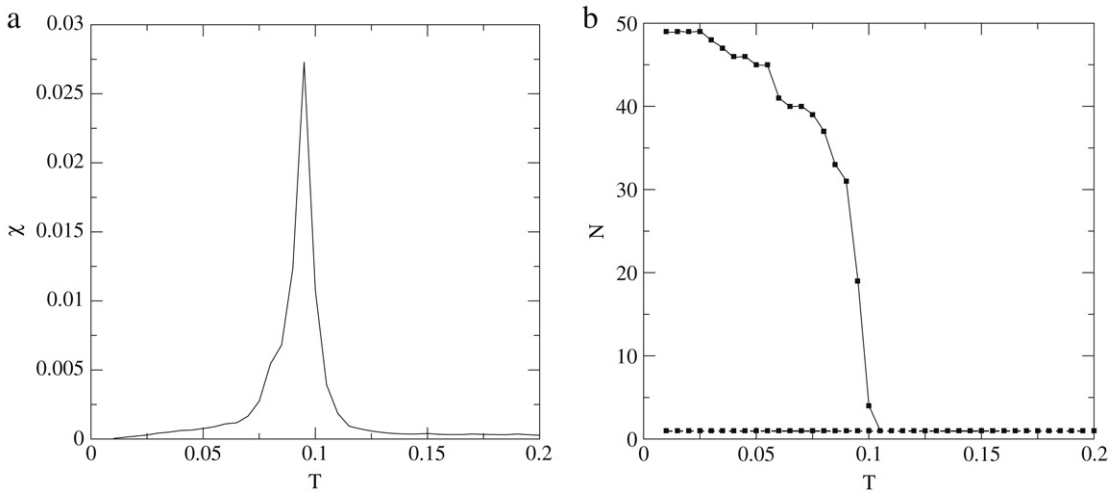

Fig. 11. (a) Susceptibility observed using SPC applied to the liver extract data. The peak signals a continuous transition between the ferromagnetic and paramagnetic phases; there is no evidence of a super-paramagnetic phase. (b) Sizes of the two largest groups. During the phase transition, the largest group (containing most of the sites), splits into many small groups.

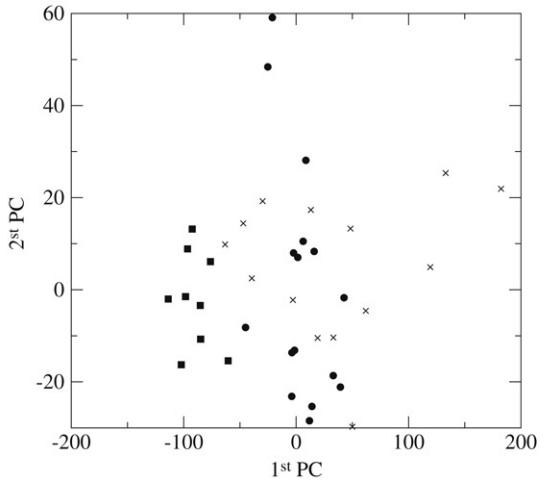

Fig. 12. PCA analysis: Projection of points corresponding to profiles obtained from brain extract onto the plane spanned by the first e second principal axes. The groups identified via SPC correspond roughly to the clusters of points evident at the left and right.

regarding the model may be found in Ref. [27].) The Hurst exponent of the model profiles, before and after perturbations of the network, are be compared with the corresponding exponent of the experimental profiles.

The network model is characterized by population variables B_i (if site i represents a B cell clone), and L_i (if the site represents a ligand species). Interactions between pairs of B cell clones, and between B cells and ligands, correspond to links in the network. (There are 4500 sites representing B cell clones and an equal number representing ligand species.) Interactions between a pair of B cell clones may be either mutually excitatory or mutually inhibitory, leading to clonal expansion or contraction, respectively. The B cell-ligand interaction is such that presence of the ligand stimulates expansion of the B cell population, while a B cell clone tends to reduce the quantity of the ligands that stimulate it.

The evolution of the model consists of the following steps:

- 1. A site is chosen at random. The associated population increases by β, regardless of whether it represents a ligand type or a B cell clone.
- 2. Another site i is chosen at random. If it is of type B, the corresponding population B_i reduced by λB_i. (Otherwise there is no change.)
- 3. A pair of connected sites are chosen at random.
  If both are of type L there is no change.
  In case one site is of type B (B_i, say) and the other of type L (species L_j, say), then B_i increases by the factor $1 + \alpha(\rho_{L_j})$, where ρ_{L_j} is the number of ligands j per B_i cell. At the same time L_j decreases by $\gamma_L B_i$.

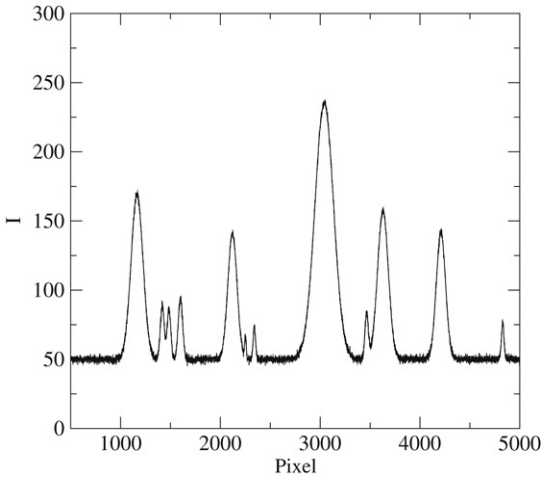

Fig. 13. Model profile generated from simulations on a scale-free network, with ten reactivity peaks, and ten peaks representing the effects of the protogold stain. Proportionality constant between the clone population and peak height $\beta = 1/30$.

If both are of type B and the interaction between them is excitatory, the population of each clone increases by the factor $1 + \alpha(\rho_B)$ where ρ_B is the concentration of complementary B cell population.

If both are of type B and the interaction between them is suppressive, the population of each clone decreases by the factor $1 - \gamma_b B$.

(The activation function is given by $\alpha(x) = 0.3\exp\{-[\log(x/1000)]^2/2\}$, as discussed in Ref. [27].)

Following the procedure of Ref. [27], once a stationary state is attained, we perturb the system by increasing the population of ten randomly chosen ligand species. Following the perturbation, the first-line clone populations (i.e., those B cell clones directly linked to the perturbed ligands) grow, so that the perturbation is suppressed, and the system again relaxes to a stationary state. The perturbation/recovery protocol is repeated several times. At each stationary state, we generate artificial reactivity profiles on the basis of the B clone populations.

To explain how the artificial profiles are generated, we return for a moment to the Panama Blot technique. In the experiment, the quantity of immunoglobulins bound to the protein extract are measured. Recall that the extract is a complex mixture of proteins, separated by molecular weight in the blot. Therefore, there is no evident correlation between the position(s) in the blot, at which a given immunoglobulin binds, and its specificity; for practical purposes the position is random.

Based on this consideration, we generate model profiles $h_{M,j}$ (j is the pixel index), by associating, with certain B cell clones, a reactivity function $f_j^{(i)}$ (i is the clone index), that is,

$$h_{M,j} = \sum_{i=1}^{10} f_j^{(i)} + \sum_{k=1}^{10} \tilde{f}_j^{(k)}. \tag{7}$$

The ten "reactive" B cell clones used in generating the profiles are chosen randomly at the start of the simulation, and are maintained fixed during the model evolution. The functions $f^{(i)}$ are Gaussians whose height and standard deviation are proportional to the population B_i of the associated clone. This choice is based on the observation that in the real profiles, higher peaks are also broader. The central position, $j_0^{(i)}$, of each Gaussian, is chosen at random.

In addition to the contributions associated with the B cell clones, further random Gaussian functions $\tilde{f}^{(k)}$, whose form and amplitude are fixed, independent of the B cell populations, are added to the profile, to represent contributions due to the protogold stain, which affects all the extract proteins. We determine the roughness $w(\ell)$ and associated Hurst exponent of the model profiles, as was done for the digitized experimental profiles.

Fig. 13 shows a model profile generated using a scale-free network, prior to any perturbation. (In simulations on a scale-free network we use parameters $\gamma_L = 1/25$, $\gamma_b = 1/50$, $\lambda = 1/100$.)

Fig. 14 shows the Hurst exponent as a function of the number n of perturbations. The Hurst exponent initially grows with n, and then tends to saturate. A similar behavior is observed in the Hurst exponent of the experimental profiles, if we order the groups according to the clinical history of malaria exposure (Fig. 15). Note however that this correspondence is only qualitative, since we cannot associate a specific number n of infections with each group.

For comparison we generate model profiles using a random (Erdös–Renyi) network, with the same mean number of connections per site 18.1 as the scale-free network discussed above. (Here the parameter values are $\gamma_L = 1/25$, $\gamma_b = 1/10$ and $\lambda = 1/50$, as discussed in Ref. [27].) The Hurst exponent follows the same trends as in the scale-free network, though with a smaller overall variation.

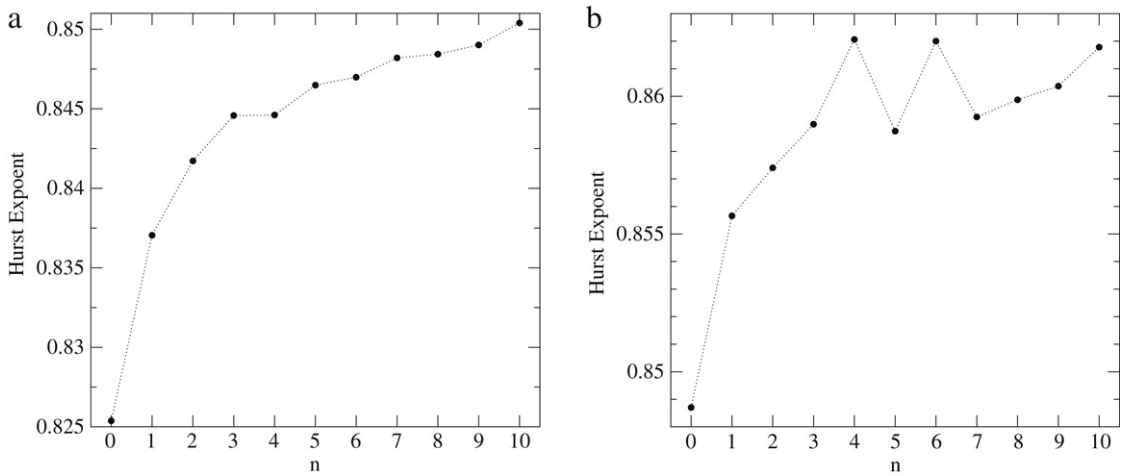

Fig. 14. Hurst exponent H of model profiles generated via simulation on (a) a scale-free and (b) an Erdös–Renyi network, as a function of the number n of perturbations.

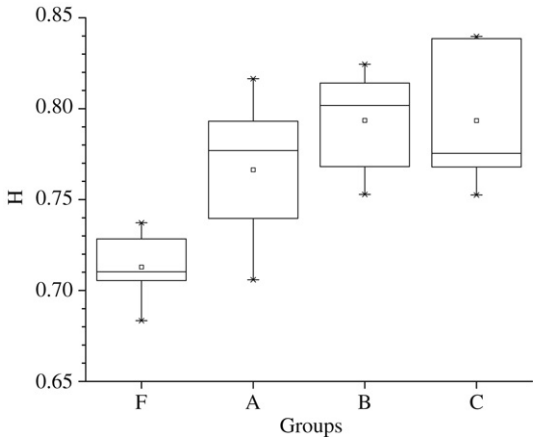

Fig. 15. Box-Plot of Hurst exponent H of profiles from the various groups, ordered by increasing number of malaria infections.

5. Shannon entropy

Another global characteristic of a reactivity profile, experimental or artificial, is its statistical entropy. Following Shannon's definition, the entropy associated with a probability distribution p_i is defined as $S = -\sum_i p_i \ln p_i$ [28]. In the experimental reactivity profiles, p_i represents the probability of grey level i (with i varying from 0 to 255); a similar discretization into 256 activity levels is performed on the artificial profiles, to define the probability distribution in an analogous manner. If exposure to a pathogen alters the global reactivity profile, such a change might be expected to affect the statistical entropy.

Ideally, the reactivity profile of each individual should represent an independent event, uncorrelated with any of the other profiles. In practice, however, the profiles were performed in batches of about 15 per membrane, so that membrane-to-membrane differences can create spurious correlations between individual signals. (We note that profiles representing individuals for all the different groups are present in each membrane, so that the systematic differences observed as a function of malaria exposure cannot be attributed to inter-membrane differences.)

Fig. 16 shows the entropy of each profile, grouped by membrane. These data reveal that the average entropy, and its dispersion, vary significantly between the three membranes. (Here we may note that a mere overall difference in gray level should not affect the entropy, while a difference in dynamic range, or the spread between the darkest and lightest regions, should change S. Such a difference in dynamic range presumably arises due to differences in gel preparation, incubation times, and/or processing.) Hence, it is not meaningful to compare entropy values of profiles belonging to different membranes, without normalizing them. (Note that the Hurst exponent is invariant under linear transformations of the profile, and so is much less sensitive to variations among membranes.) It is nevertheless useful to compare entropies of profiles in the same membrane (Table 1); this comparison reveals that the profiles of symptomatic individuals tend to have smaller entropy values.

Table 1
Entropy of the profiles, grouped by membrane

Membrane 1			Membrane 2			Membrane 3		
Order	Group	Entropy	Order	Group	Entropy	Order	Group	Entropy
1	F	4.13	1	D	3.22	1	D	3.92
2	E	4.13	2	A	3.25	2	D	4.01
3	D	4.23	3	D	3.38	3	B	4.09
4	D	4.28	4	D	3.53	4	F	4.09
5	B	4.28	5	A	3.62	5	B	4.15
6	F	4.30	6	F	3.62	6	A	4.18
7	C	4.30	7	B	3.79	7	F	4.23
8	A	4.33	8	E	3.81	8	A	4.27
9	A	4.33	9	E	3.99	9	F	4.31
10	E	4.34	10	C	4.02	10	E	4.33
11	A	4.38	11	C	4.07	11	C	4.49
12	B	4.39	12	E	4.23	12	C	4.50
13	C	4.43	13	B	4.33	13	E	4.53
14	F	4.46				14	E	4.53
15	C	4.55				15	C	4.57
16	E	4.59				16	B	4.58
17	B	4.67						

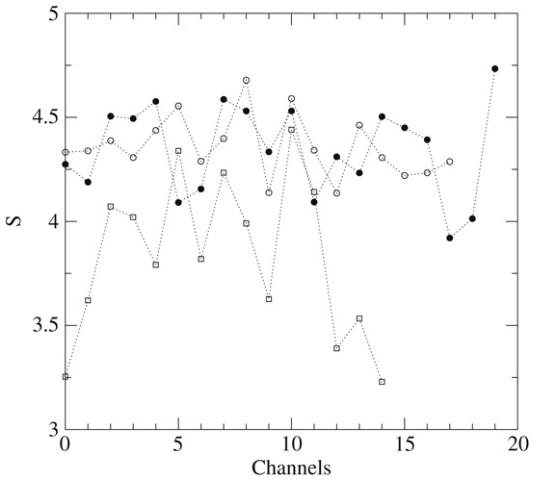

Fig. 16. Entropy S of experimental reactivity profiles, grouped by membrane. Open circles, open squares, and closed circles represent profiles from membranes 1, 2, and 3, resp.

For each membrane, let s be the number of symptomatics and m the position of the symptomatic with highest entropy (with $m = 1$ the position of lowest entropy). In membrane 1, for example, the two symptomatics have the third and fourth lowest values, so $s = 2$ while $m = 4$. For membrane 2, $s = 3$ and $m = 4$, while $s = m = 2$ in membrane 3. If the distribution of entropies were random, the probabilities of a given outcome (s, m) in a membrane with n individuals would be given by the hypergeometric distribution. The associated probabilities are 0.008, $p = 0.014$, and $p = 0, 007$ for membranes 1, 2, and 3, respectively. The low probabilities of occurrence ($p \approx 0.01$), strongly suggest that the profiles of symptomatic individuals exhibit lower statistical entropy than non-symptomatics. (Note that this tendency is opposite to that of the Hurst exponent, which is larger for symptomatics and others inhabiting the endemic region. A larger H corresponds to a higher degree of correlation in height variables, which in turn corresponds to smaller S, the entropy being maximum for a uniform probability distribution.)

Using the procedure described above, we calculate the model profile entropy. Figs. 17 and 18 show the entropy as a function of the number of perturbations for for scale-free, and random networks, respectively. We note an initial growth of the entropy followed by a tendency toward saturation, i.e., a behavior similar to that of the Hurst exponent. The initial growth of the entropy can be understood as follows. Before any perturbation, the reactivity profile has a few peaks, superimposed on a roughly uniform, low-level background. The growth of certain B cell clones following a perturbation has the effect of amplifying the associated peaks in the profile, and thereby making the distribution p_i somewhat more uniform, which corresponds to a higher entropy. Thus, in the model, entropy and Hurst exponent exhibit a qualitatively similar response to perturbations.

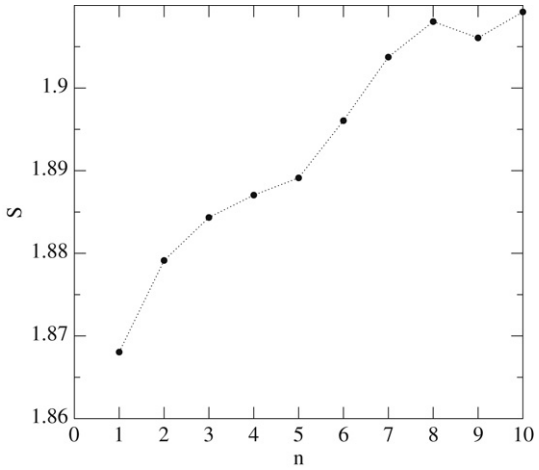

Fig. 17. Entropy of model profiles (scale-free network) versus number of perturbations.

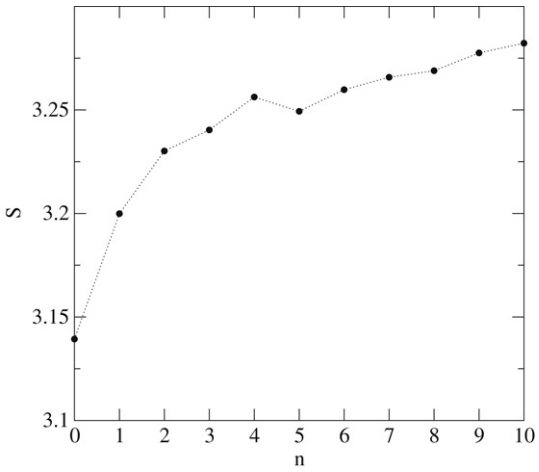

Fig. 18. Entropy of model profiles (Erdös–Renyi network) versus number of perturbations.

6. Conclusions

The possibility of multiple effective interactions between B lymphocyte clones motivates the study of the immune system as a complex system [27]. Studies of this nature are important for interpreting experiments related to global immune response, such as the Panama blot technique. Among the many possible global measures of activity, the immune profile roughness and statistical entropy are analyzed here, and compared with results of simulated profiles, generated using an immune network model [27].

The roughness w_i is characterized by the Hurst exponent H. (Although the immune profile does not appear to be fractal, we may nonetheless evaluate the associated Hurst exponent at small scales.) We find that the value of H discriminates, in a statistically significant manner, between individuals living in regions endemic for malaria, and those from malaria-free regions. One can argue that several factors might influence the reactivity profiles: age, previous diseases, genetic profile etc. It is important to cite, from Ref. [18] and references therein, that "reactivity patterns of IgM (...) natural antibodies to autologous tissue proteins appear established early in life, and remain remarkably stable throughout healthy living, but are capable of characteristic changes in autoimmune and other human diseases. Thus, such patterns are likely to reflect stabilized states of physiologic activation, shaped by polymorphic genes relevant for the immune system which can be selected in evolution". The calculation of Hurst exponent has been used in the characterization of states in many different biological problems [13,14].

We also find that the statistical entropy of the reactivity profile is highly correlated to malaria exposure, when the comparison is restricted to the same membrane. (Profiles of symptomatic individuals tend to have a lower entropy. Membrane-to-membrane variations hamper comparison of profiles on different membranes.) We find that the sorting by entropy value observed within a given membrane ($\simeq$15 profiles), has a rather small probability ($p \approx 0.01$) of occurring randomly. Unlike the Hurst exponent, which discriminates between individuals from endemic and non-endemic regions,

entropy discriminates between symptomatics and non-symptomatics, but does not separate individuals by region. More surprisingly, entropy is smaller for symptomatics than for the rest of the population. This is qualitatively different from what is observed in our model simulations, in which entropy grows with an increasing number of perturbations.

From our results on the Hurst exponent and statistical entropy, we infer that exposure to the malaria parasite modifies the state of immunological activity, in a manner that affects reactivity to a set of proteins that are not necessarily related to those of the parasite. In our view, the relation between H values and history of malaria exposure (and between entropy and history of malaria exposure), is likely due to one or more of the following factors: (1) the (human) protein extracts used in the Panama blot may contain some proteins in common with the parasite; (2) exposure to the parasite may elicit cross reactions to self-proteins; (3) exposure to the parasite may cause global changes in immune reactivity.

Representing each profile as a point in reactivity space (with a dimensionality equal to the number of protein bands in the extract, when separated by electrophoresis), we determine the plane associated with the highest degree of correlation via principal components analysis. Points associated with individuals from the two regions (endemic and nonendemic), tend to fall in distinct areas of this plane. The results of super-paramagnetic clustering are less clear. For the profiles obtained using the liver extract, there is no separation into distinct groups. By contrast, SPC does separate the set of brain-extract profiles into two groups, but these groups are not correlated to the history of malaria exposure.

As an aid to interpret the experimental profile results, we study reactivity profiles generated using an immune network model [27]. As the model is repeatedly perturbed with the same set of ligands, the Hurst exponent of the associated profiles increases. The increasing of H as a function of the number of perturbations is sublinear. Analogous behavior is observed in the Hurst exponent of the experimental profiles groups according to the intensity of malaria exposure.

Acknowledgments

We thank the Immunology Group of ICB/UFMG, and in particular, N. Vaz, A.S. Neto and L. Goulart, for providing us with the Panama-blot data, and for very helpful discussions, clarifying the experimental procedures. Their help was fundamental for the conclusion of this work. This work was supported in part by the Brazilian agencies CAPES, FINEP, CNPq, and FAPEMIG.

References

[1] F.J. Varela, A. Coutinho, Immunol. Today 5 (1991) 159–166.
[2] I.R. Cohen, Immunol. Today 12 (1992) 490–494.
[3] A.M.C. Faria, H.L. Weiner, Immunol. Rev. 206 (2005) 232–259.
[4] R. Duchman, E. Schimit, P. Knolle, K.M. Zum, H. Buschenfelde, M.M. Neurath, Eur. J. Immunol. 26 (1996) 934–938.
[5] S. Danese, S. Semeraro, A. Papa, I. Roberto, F. Scaldaferri, G. Fedeli, G. Gasbarrini, A. Gasbarrini, World J. Gastroenterol. 11 (2005) 7227–7236.
[6] J.J. Powell, P. Ashwood, R.S.J. Harvey, R. Wolstencroft, M.E. Gershwin, R.P.H. Thompson, J. Autoimmunity 14 (2000) 99–105.
[7] A. Coutinho, Int. J. Dev. Biol. 49 (2005) 131–136.
[8] A. Coutinho, M.D. Kazatchkine, S. Avrameas, Curr. Opin. Immunol. 7 (1995) 812–818.
[9] D. Stahl, S. Lacroix-Desmazes, L. Mouthon, S.V. Kaveri, M.D. Kazatchkine, J. Immunol. Methods 240 (2000) 1–14.
[10] J.G. Moreira, J.K.L. Silva, S.O. Kamphorst, J. Phys. A 27 (1994) 8079.
[11] J.A. Sales, M.L. Martins, J.G. Moreira, J. Phys. A 245 (1997) 461–471.
[12] C. Castelnovo, A. Podesta, P. Piseri, P. Milani, Phys. Rev. E 65 (2002) 021601.
[13] B.H. Havsteen, J. Theoret. Biol. 231 (2004) 39–48.
[14] V. Maxim, L. Xendur, J. Fadili, J. Suckling, R. Gould, R. Howard, E. Bullmore, NeuroImage 25 (2005) 141–158.
[15] A. Nobrega, M. Haury, A. Grandien, E. Malanchere, A. Sundblad, A. Coutinho, Eur. J. Immunol. 23 (1993) 2851–2859.
[16] M. Haury, A. Grandien, A. Sundblad, A. Coutinho, A. Nobrega, Scand. J. Immunol. 39 (1994) 79–87.
[17] M. Blatt, S. Wiseman, E. Domany, Neural Comput. 9 (8) (1997) 1805.
[18] C. Fesel, L.F. Goulart, A.S. Neto, A. Coelho, C.J.F. Fontes, E.M. Braga, N.M. Vaz, Malaria J. 4 (2005) 1–15.
[19] B. Stransky, Tese de Doutorado Instituto de Microbiologia, Universidade Federal do Rio de Janeiro, 2003.
[20] A.L. Barabasi, R. Albert, Science 286 (1999) 509–512.
[21] R. Albert, H. Jeong, A.L. Barabasi, Nature 406 (2000) 378–382.
[22] W.R. Dillon, M. Goldstein, Multivariate Analysis: Methods and Applications, John Wiley, New York, 1984.
[23] M. Blatt, S. Wiseman, E. Domany, Phys. Rev. Lett. 76 (1996) 3251.
[24] M. Blatt, S. Wiseman, E. Domany, Phys. Rev. E 57 (1998) 3767.
[25] R.H. Swendsen, S. Wang, A.M. Ferrenberg, The Monte Carlo Method in Condensed Matter Physics, Springer-Verlag, Berlin, 1992.
[26] L.F. Goulart, private communication.
[27] L.C. Ribeiro, R. Dickman, A.T. Bernardes, N.M. Vaz, Phys. Rev. E 75 (2007) 203191.
[28] C.E. Shannon, Bell Syst. Tech. J. 27 (1948) 379–423.