Neutral Clustering in a Simple Experimental Ecological Community

B. Houchmandzadeh

CNRS & Grenoble Universités, Lab. Spectrométrie Physique, BP87, 38402 St-Martin d'Hères Cedex, France

The spatial distribution of most species in ecosystems is nonuniform. New theories try to explain patterns observed at multiple scales in terms of neutral processes such as birth, death, and migration. We have devised an experimental, niche-free ecosystem where the amplitude of neutral patchiness can be precisely measured. Spatial distribution of species in this system is extremely clustered. We demonstrate that this clustering is entirely attributed to neutral causes and show that the most basic properties of life can provoke intricate spatial structures without clues from the environment.

　　　　　PACS numbers: 87.23.Cc, 05.40.−a, 89.75.Fb

Since the seminal paper by L. R. Taylor [1] who surveyed around 4000 samples from 100 species across different kingdoms, it is now common knowledge that the spatial distribution of species in ecosystems is far from random and organized into aggregates and clusters. The *patchy* distributions of species in ecological systems have been documented in a wide range of biological communities, from phytoplankton [2] to trees in tropical or Nordic forest [3,4], and even cancerous cells [5]. Understanding the spatial structures of species distribution is of fundamental importance to ecology and to our comprehension of population dynamics and biodiversity [6,7].

It is generally believed that patchy distributions cannot be random and will not exist in a perfectly homogeneous environment because random displacements of organisms or plants' seeds smoothen the profile of organisms densities. Therefore, the existence of complex spatial structures in ecosystems has been mainly attributed to exogenous causes such as the inherent heterogeneity of the environment (lakes, mountains, salinity and temperature gradients, ...) [8] or competition for food and mates. These causes are widely studied by ecologists and have been broadly termed "niche theories."

The nonuniform distribution of species is however so widespread that one could wonder if more general factors are not involved. During the last 7–8 years, alternative explanations called neutral theories [9] have come to the forefront which search the origin of intricate spatial structures of species distribution in the basic properties of life itself: birth, death, and migration (limited dispersal). The birth event has very specific properties: (i) it changes the number of individuals in integer (discrete) units; (ii) a newborn always appears close to a parent. In contrast, the death phenomenon can take place anywhere. Limited dispersal implies that an organism can explore only a small portion of the ecosystem during its lifetime. Even if it seems counterintuitive, these very simple facts can overcome the smoothening caused by random migrations. Keeping all environmental parameters constant and removing all individual interactions, it can be shown that an ecological system of randomly moving and replicating

organisms will display considerable patchiness [10,11]. For example, Fig. 1 displays the spatial distribution of *Dictyostelium Discoïdum* in a petri dish. As we will show, this distribution is an instance of neutral clustering.

Of course, it would be absurd to neglect the existence of landscape influence or social interactions on spatial distribution of species in natural ecosystems [12,13]. Exogenous causes however have to be weighted against neutral ones. In natural ecosystems, it is difficult to assess the relative importance of neutral versus exogenous sources of spatial clustering, first because it is difficult to collect a complete, statistically significant data set over an extended spatial and temporal range [14] and second because it is difficult to infer directly the influence of geographical and/or social interactions [15]. The absence of unambiguous proof has made neutral theories the subject of heated debates [16].

In order to address this issue, we have devised a laboratory ecosystem of *Dictyostelium Discoidum* (i) which is spatially and temporally extended, (ii) where the environment is homogeneous and controlling parameters can be varied easily, and (iii) where the position of *each* individual can be precisely measured at various times. This ecosystem
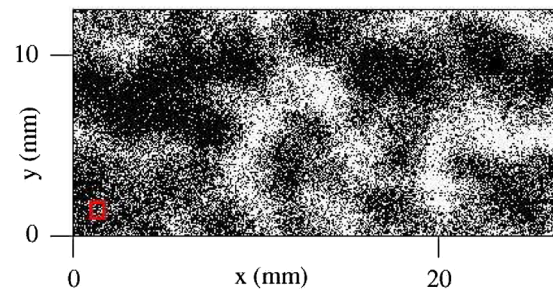

FIG. 1 (color online).　Spatial distribution of amoebae in an experimental ecosystem: $\approx 70\,000$ Axenic *D. Discodum* grown in a petri dish under controlled conditions after 112 hr (8.7 generation) are displayed. Each dot represents one *Dictyostelium* (~ 5–10 micron in diameter). The square corresponds to the size of one photograph of a sequence covering the area (see Fig. 2). The variance to mean ratio, computed for the number of organisms contained in quadrats the size of the square, is ≈ 45.

displays substantial patchiness (Fig. 1) which can be entirely attributed to neutral clustering and computed by means of mathematical modelling and numerical simulations, as we will demonstrate below.

The schematic of our experiment is shown in Fig. 2. Various parameters in the growth media can be changed to obtain different growth rates and mobilities for the amoebae (see Material & Methods [17]). We observe that in all of experimental ecosystems, the initial homogeneous distribution of species rapidly becomes a patchy distribution (Fig. 1).

Dictyostelia are complex organisms, capable of displaying subtle behaviors such as chemotaxis to find bacteria or transition to a form of multicellularity when starved. One could suppose *a priori* that the aggregation of the amoebae is caused by some hidden factors such as cell communication. The conditions used in these experiments are far from starvation, and the growth is always exponential. The clustering phenomena we observe does not have any resemblance to the aggregation of *Dictyostelia*, where individual cells fuse to form a multicellular slug; in our experiments, cells conserve their shape and individuality throughout the experiment. Furthermore, we have also used a strong mutant (*pdsA-*) which lacks the chemotactic cell movement during starvation [8], and the results for these cells are no different from the nonmutant type (Fig. 4).

However, we cannot exclude *a priori* the presence of other hidden, unsuspected factors. To demonstrate that the aggregation phenomena is only due to neutral causes, we should be able to relate *quantitatively* all the clustering to basic, neutral parameters of the ecosystem. The only two neutral causes involved here are birth and random migration and both rates can be precisely measured. In this experimental ecosystem, the movements of *Dictyostelia* are Brownian (random) and thus can be characterized by a unique diffusion coefficient D [Figs. 2(c) and 2(d)]; moreover, the population increase is exponential and characterized by the growth rate α [Fig. 2(e)] as long as their number is far from saturation.

To quantify the patchiness of the ecosystem and to compare it to a neutral model, we use two sets of measurements: the pair correlation function and the variance to mean ratio.

The pair correlation function provides the probability of finding an organism at distance r of another one. Purely random distributions of organisms will have a flat pair correlation function; clustered distributions (such as those in Fig. 1) display a peak for short distances. A quantity closely related to the pair correlation function and used in ecological literature is the β-diversity [3], the rate at which similarity between quadrats decreases as a function of distance [18,19]. The pair correlation function is one of the most complete pieces of information about a system, and we are able to measure it precisely. To give an idea, in the final stages of each ecosystem, positions of about 50–70 000 amoebae are measured, and the experimental pair
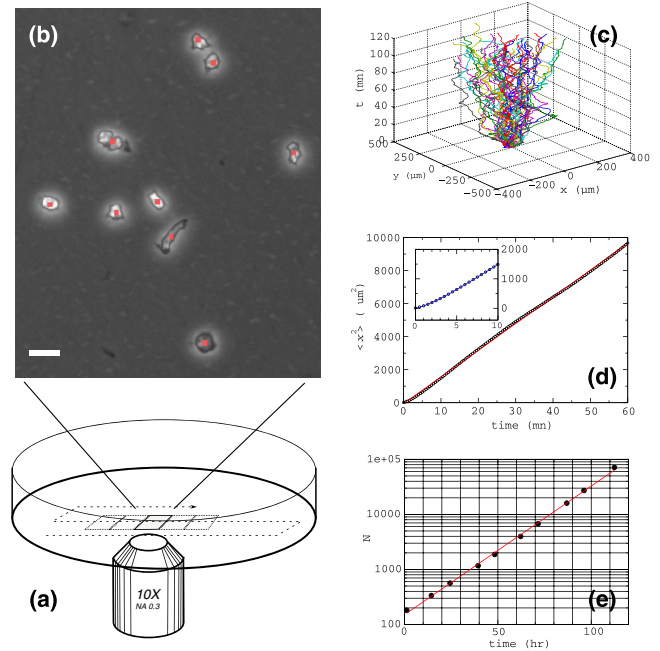

FIG. 2 (color online). Scheme of one experiment. Axenic *Dictyostelia* are spread and put to growth in a 50 mm petri dish under controlled conditions. (a) At regular time intervals, the petri dish is scanned by a computer controlled microscope objective and contiguous photographs (660×832 μm) are taken. Paving the area of Fig. 1 requires 600 photographs. (b) A homemade image analysis program detects the position of each *Dictyostelium* in each photograph. The displayed figure is a portion of a photograph (bar = 20 μm) where detected *Dictyostelia* are marked by squares. A global map such as Fig. 1 is reconstituted by combining *Dictyostelia* positions from all photographs in a sequence. By repeating the operation (4 to 12 hour intervals, depending on the growth rate), a spatio-temporal map is constituted. One experiment can require up to 10 000 photos. (c) To measure the Brownian diffusivity, the movements of *Dictyostelia* in a microscope frame are recorded and the cells trajectories are reconstituted. The figure shows the trajectories of ≈ 50 *Dictyostelia* over two hours. (d) The mean square linear displacement function $\langle x^2 \rangle$ of all trajectories is computed for various time intervals t. The diffusion coefficient D is recovered from the linear regression $\langle x^2 \rangle = Dt$ (see M&M [17]) (inset). The small deviation from the straight line at the origin is due to persistence in amoebae movement (2–3 minutes). (e) The growth rate α is computed by reporting the total number N of *Dictyostelia* in each map versus time: $N = N_0 \exp(\alpha t)$.

correlation function is a statistic over $\approx 10^9$ measured distances.

The pair correlation function of a neutral model where organisms move randomly and replicate can be exactly computed. Denoting the average density of cells in the petri dish by $c(t)$, the exponential growth is $c(t) = c_0 \exp(\alpha t)$ where c_0 is the initial concentration [Fig. 2(e)]. By using techniques introduced by Glauber [20] (see for example [11]), the normalized pair correlation function (see M&M [17]) $g(\mathbf{r}, t)$ can be shown to obey

$$\partial g / \partial t = D\nabla^2 g + (2\alpha/c)\delta(\mathbf{r}) \quad (1)$$

where $\delta(\mathbf{r})$ is the Dirac delta function. The three parameters appearing in this equation, D, α, and c_0 are measured from the experiment [see Figs. 2(c)–2(e) and M&M [17]]. Neutral clustering is a competition between birth and diffusion. Birth events increase short pair distances because individuals always appear close to their parent. On the other hand, random diffusion tends to smoothen overrepresented distances and make the distribution even. Equation (1) captures this competition.

In order to test the validity of this neutral model, different laboratory ecosystems are realized in which both the growth rate and the diffusion coefficient are varied through the composition of the growth medium. In order to group all the results, it is more convenient to use the natural scales of each ecosystem. The natural time scale is the generation time $1/\alpha$; the natural scale of space is the diffusion length $\lambda = \sqrt{2D/\alpha}$, i.e., the average distance travelled by a Brownian organism during one generation. In these units, the solution of the Eq. (1) takes the simple form

$$g(\mathbf{r}, t) = 1 + \frac{1}{\pi c_0} \int_0^t \frac{1}{s} \exp\left(-\frac{r^2}{2s}\right) \exp(s - t) ds. \quad (2)$$

Figure 3 is one of the key results of this article; it shows the comparison at different times between Eq. (2) and pair correlation functions extracted from experimental ecosystems of *Dictyostelia*. There is an excellent agreement between the neutral theory of patchiness and the strong clustering observed in actual ecosystems. In fact, *all* the patchiness of Fig. 1 can be attributed to neutral clustering and no other cause such as interactions between *Dictyostelia* is needed. Note also that the theory has not been corrected for small deviations due the finite size of the experimental ecosystems.

Pair correlation function provides an invaluable information about the underlying spatial distribution, but its measure in natural ecosystems, necessitating the determination of each individual's position, will be extremely difficult. On the other hand, field ecologists routinely use tools such as fluorescence or quantitative PCR to measure the *number* of individuals in a given patch. A robust metric to measure the degree of patchiness and more adapted to ecologists need is the variance to mean ratio (VMR). VMR is obtained by dividing the space into contiguous quadrats (squares) of area $\ell \times \ell$, counting the number N_i of individuals in each quadrats, computing the mean μ and variance σ^2 of these numbers and forming their ratio. A VMR = 1 is the signature of a pure (Poissonian) random distribution; VMR > 1 indicates patchiness and aggregation; VMR < 1 is seldom encountered in natural ecosystems and reflects ordered distributions such as those encountered in human plantations [21]. VMR depends in general on the size of quadrats ℓ, except in the case where the spatial distribution is purely random. The dependence of VMR on ℓ informs us of the underlying nature of the
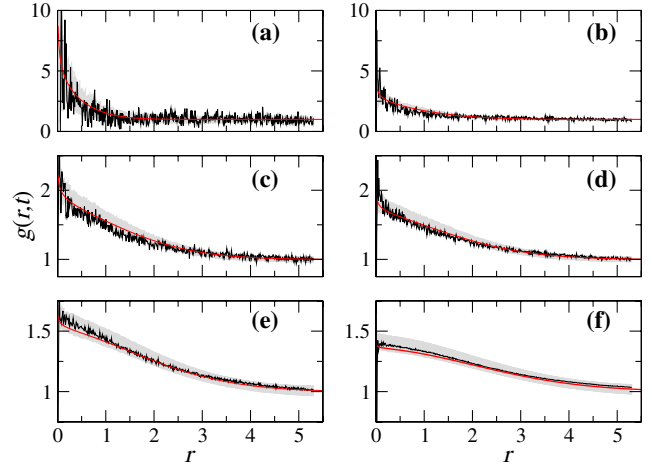

FIG. 3 (color online). Pair correlation function of an experimental growing population $g(r, t)$ as a function of reduced distance r ($= R/\lambda$ where R is the physical distance) for 6 different reduced times t ($= \alpha T$ where T is the physical time). Rough (black online) curves represent experimental data and are computed from spatial maps of *Dictyostelia* distribution in the petri dish. Smooth (red online) curves are the theoretical values given by expression (2). Gray shaded areas are estimations of the variability of experimental pair correlation functions (because of the finite number of organisms); they are obtained by numerical simulation of 26 replicates of the ecosystem and represent the average pair correlation function plus and minus 1 standard deviation. For this set of experiment, $D = 160$ $\mu m^2/mn$, $\alpha = 9 10^{-4}$ mn^{-1}, $\lambda = 600$ μm, $c_0 = 0.19$. (a) $t = 0.78$ ($T = 14$ hr); (b) $t = 2.1$ (39 hr); (c) $t = 3.4$ (62 hr); (d) $t = 3.9$ (72 hr); (e) $t = 4.7$ (87 hr); (f) $t = 6.1$ (112 hr). Figure (f) corresponds to the autocorrelation function of the spatial map displayed in Fig. 1.

causes of patchiness. Theoretically, the VMR is obtained by integration of the pair correlation function (2) (see M&M [17]) and reads, in the natural units mentioned above

$$\text{VMR}(\ell, t) = 1 + \frac{4}{\pi} \ell^2 \int_0^t \frac{e^s}{s} h^2(s/\ell^2) ds \quad (3)$$

where

$$h(u) = \int_0^1 (1 - x) \exp(-x^2/2u) dx.$$

The above formula displays explicitly the dependence of the VMR on the growth time t *and* the quadrat size ℓ. Figure 4 shows the experimental measurement of VMR($\ell = 1, t$) for five different ecosystems and its comparison to the theoretical prediction Eq. (3). For the growing population of *Dictyostelia*, the predictive capabilities of the theory seems outstanding. At large times, the spatial structure of the population is highly clustered (VMR $\gg 1$), *regardless* of the scale of the study.

In conclusion, we have shown that living and diffusing organisms show a natural tendency to aggregate because of neutral causes. We have achieved that by using an experimental ecosystem of amoebae where we have stripped out
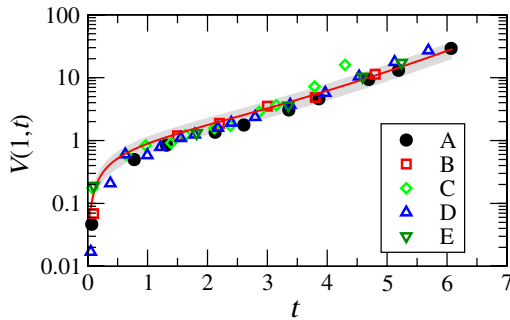
FIG. 4 (color online). The $V(\ell, t)$ function (variance to mean ratio minus one) (a) as a function of time, for fixed quadrat size $\ell = 1$; The gray shaded area represents the average ± 1 standard deviation of 26 numerical simulations and is used as an estimation of the fluctuation of $V(\ell, t)$ for finite samples. Curves are the exact analytical description of the function V. Symbols represent actual data for various experimental ecosystems (A–D: Axenic AX2 cells, E: *pdsA*-mutants). α and D parameters, in hr^{-1} and μm^2/mn are measured as A: 0.054, 160; B: 0.069, 370; C: 0.030, 60; D: 0.024, 68; E: 0.072, 170.

all environment heterogeneity. This simple ecosystem displays highly clustered spatial structure, and we have shown that its statistical properties can be precisely predicted by the neutral model we have developed. To our knowledge, this is the first time the existence and importance of neutral causes has been demonstrated experimentally. Random migration is often associated with smoothening of rough distributions and normally causes uniform distribution. What sets apart a living ecosystem however is that correlation creation at short distances due to duplication (i.e., birth close to a parent) cannot be smoothened by random diffusion and provokes clustering.

The clustering we have described here applies mostly to growing populations such as plankton blooms, populations extending their geographical range [22] or more generally to situations where there are successive (or cyclic) events of extinction and recolonization. For this latter case, the amount of clustering will be higher at each new cycle because the initial condition at the start of each new cycle would itself display an amount of clustering.

More generally, the inclusion of neutral death phenomena only enhances the clustering effect: birth always enriches correlations at short distances, and death removes correlations at *all* distances, thus contributing to a sharper peak in the pair correlation function. More precisely, when death with rate β is included, the source term in Eq. (1) will read $2(\alpha + \beta) \exp[(\beta - \alpha)t]\delta(r)$. At the limit $\alpha = \beta$, correlations diverge with time [10,11]. The clustering persists when various forms of density dependence growth are included [23,24].

The aim of the present Letter is not to deny or diminish the role of environmental factors, but to stress that the observation of patchy spatial distributions of species (VMR $\gg 1$) in natural ecosystems should not be considered surprising or nonrandom by itself. As we have seen, even in the most simple ecosystems, this is the rule. In studying natural ecosystem, one has to evaluate the relative importance of neutral versus environmental causes.

The present work could be further extended to model environmental factors provoking nonuniform (in time and space) growth and diffusion rates through, for example, temperature fluctuations and local substrate modification. This will constitute a model ecosystem to assess the relative importance of neutral versus exogenous causes.

[1] L. R. Taylor, I. P. Woiwod, and J. N. Perry, J. Anim. Ecol. **47**, 383 (1978).
[2] A. Martin, Philos. Transact. Ser A Math Phys. Eng. Sci. **363**, 2873 (2005).
[3] J. C. Nekola and P. S. White, Journal of Biogeography **26**, 867 (1999).
[4] R. Condit *et al.*, Science **295**, 666 (2002).
[5] I. González-García, R. V. Solé, and J. Costa, Proc. Natl. Acad. Sci. U.S.A. **99**, 13085 (2002).
[6] *Ecology, Genetics and Evolution of Metapopulations*, edited by I. Hanski and O. E. Gaggiotti (Academic Press, New York, 2004).
[7] *Spatial Ecology*, edited by D. Tilman and P. Kareiva (Princeton University Press, Princeton, NJ, 1997).
[8] R. Sucgang *et al.*, Dev. Biol. **192**, 181 (1997).
[9] G. Bell, Science **293**, 2413 (2001).
[10] W. R. Young, A. J. Roberts, and G. Stuhne, Nature (London) **412**, 328 (2001).
[11] B. Houchmandzadeh, Phys. Rev. E **66**, 052902 (2002).
[12] S. P. Hubbell, Functional ecology **19**, 166 (2005).
[13] M. A. Leibold and M. A. McPeek, Ecology **87**, 1399 (2006).
[14] J. H. Brown, G. C. Stevens, and D. M. Kaufman, Annual Review of Ecology and Systematics **27**, 597 (1996).
[15] M. G. Turner, Annual Review of Ecology, Evolution, and Systematics **36**, 319 (2005).
[16] V. Gewin, PLoS Biology **4**, e278 (2006).
[17] See EPAPS Document No. E-PRLTAO-101-021832 for experimental and theoretical details of the submitted article. For more information on EPAPS, see http://www.aip.org/pubservs/epaps.html.
[18] M. J. Doubell *et al.*, Appl. Environ. Microbiol. **72**, 4475 (2006).
[19] R. Muneepeerakul *et al.*, Nature (London) **453**, 220 (2008).
[20] R. J. Glauber, J. Math. Phys. (N.Y.) **4** 294 (1963).
[21] A. Okubo and J. G. Mitchell, Springer, 2001.
[22] R. J. Wilson, C. D. Thomas, R. Fox, D. B. Roy, and W. E. Kunin, Nature (London) **432** 393 (2004).
[23] B. Bolker and S. W. Pacala, Theor. Popul. Biol. **52**, 179 (1997).
[24] B. Houchmandzadeh and M. Vallade, Phys. Rev. E **68**, 061912 (2003).