

Ensemble de bactéries *Vibrio Harveyi* génétiquement identiques, d'environ 1,5  $\mu\text{m}$  de longueur. À gauche, photo en microscopie à contraste de phase. À droite, signal de luminescence, très variable d'une bactérie à l'autre.

# Importance et contrôle des fluctuations dans les systèmes biologiques

**Bahram Houchmandzadeh** (bahram@spectro.ujf-grenoble.fr) et **Irina Mihalcescu**  
Laboratoire de spectrométrie physique, UMR 5588 CNRS-UJF, 38402 Saint-Martin d'Hères Cedex

Les cellules vivantes sont des « usines » chimiques miniatures. Contrairement à leurs équivalents macroscopiques et plus précisément à cause du faible nombre de molécules impliquées, les réactions qui s'y déroulent ne sont pas entièrement déterministes.

La gestion des « probabilités » fait donc partie intégrante de la vie des cellules et laisse sa trace à toutes les échelles spatiales du vivant.

Nous allons passer en revue dans cet article quelques aspects de cette gestion du « bruit stochastique », qui a profondément changé notre vision des systèmes vivants durant ces dix dernières années.

Les fluctuations et l'approche probabiliste ont joué un rôle fondamental en physique à travers la physique statistique et la mécanique quantique. En biologie cependant, mise à part la théorie de l'évolution, le concept de probabilité n'avait pas un grand rôle jusqu'à une période récente. Pour la plupart des biologistes, les cellules étaient des machines déterministes, certes très complexes mais répondant de façon unique à une stimulation donnée. Mais à partir de la fin des années 1990, il est apparu de plus en plus clairement que les fluctuations jouent un rôle majeur en biologie, et un nombre très important de travaux y ont été consacrés. La variabilité des comportements des organismes génétiquement identiques en conditions externes identiques (appelée aussi « individualité non génétique ») était reconnue par les biologistes depuis longtemps et des expériences importantes y avaient été consacrées ; mais cette variabilité n'occupe le devant de la scène que depuis récemment. On s'est rendu compte que les organismes vivants doivent consacrer une partie de leurs ressources au contrôle des fluctuations, surtout pour les fonctions où la fidélité de réponse est cruciale à la survie de l'individu [1].

Dans cet article, nous allons passer en revue quelques aspects importants de ce thème : (i) pourquoi les fluctuations existent et quels outils les scientifiques ont élaboré pour les mesurer ; (ii) comment font les cellules pour les maîtriser ou (iii) les utiliser ; (iv) quelle est leur importance à des échelles spatiales plus grandes.

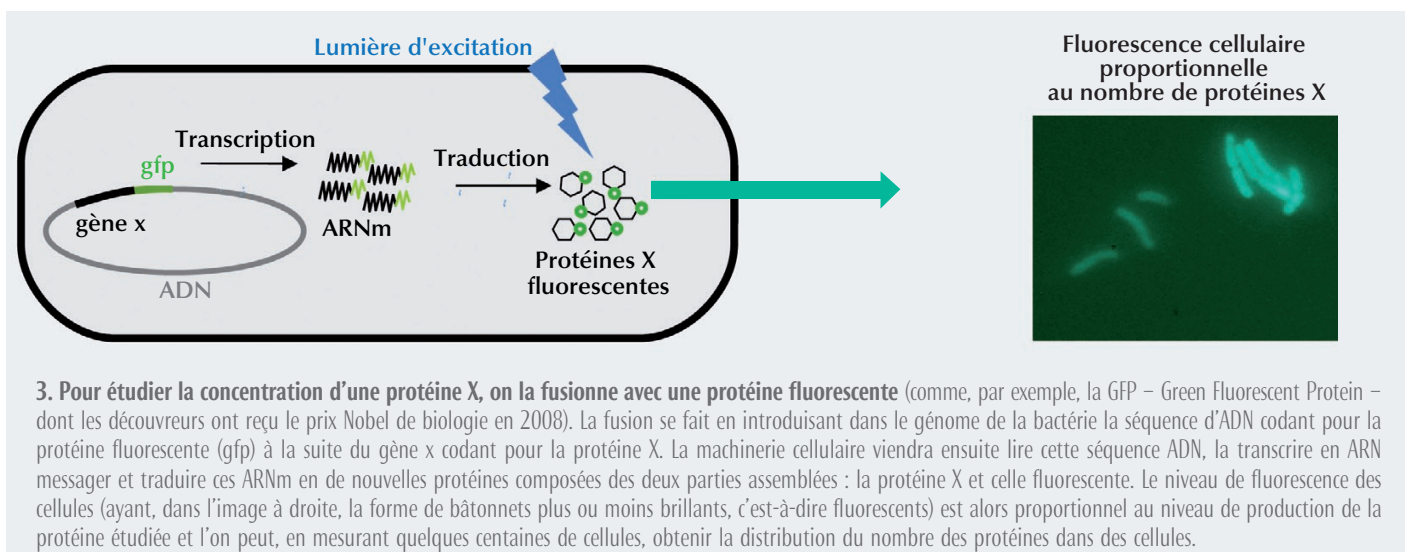
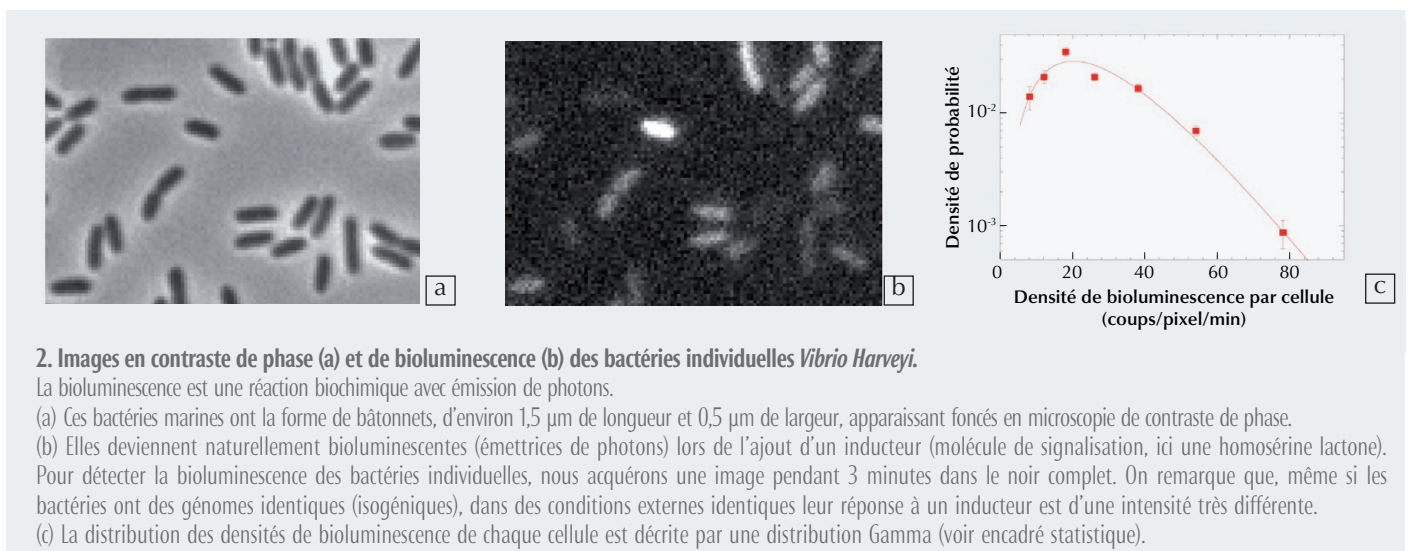
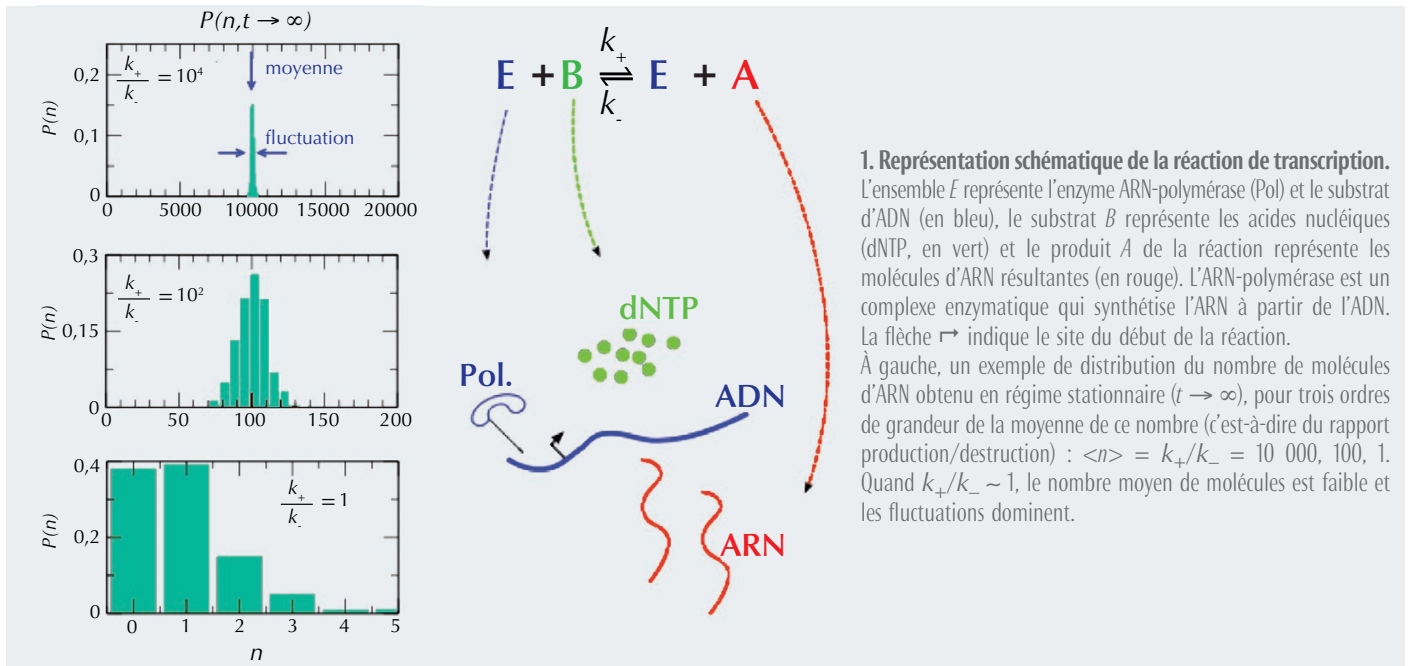
## Fluctuations dans les cellules

Une machine déterministe produit toujours le même résultat en réponse à une « entrée » donnée. Au contraire, une machine probabiliste produit des résultats différents, distribués plus ou moins largement autour d'une valeur moyenne. En pratique, nous pouvons considérer une machine probabiliste comme déterministe si la largeur de la distribution des résultats devient très faible comparée à la moyenne. Ceci est le cas en chimie macroscopique où l'information est essentiellement contenue dans la « moyenne », dont nous pouvons décrire l'évolution en utilisant les équations différentielles usuelles de la cinétique chimique.

Cependant, quand le nombre de molécules est faible, la moyenne est une quantité qui décrit mal l'observation, car les fluctuations deviennent importantes (voir encadré statistique, p. 20). Dans les bactéries, les réactions chimiques (assurées essentiellement par des protéines) impliquent souvent un très faible nombre de molécules (fig. 1). Un exemple classique est la bactérie *Escherichia coli*, dont 80% des gènes sont exprimés à moins de 100 copies de protéines par cycle cellulaire, menant à des amplitudes relatives de fluctuations  $\sigma / \langle \mu \rangle$  de 5% à 65%. Ainsi, observant un ensemble de bactéries, nous devons nous attendre à ce que le produit d'une réaction donnée soit très variable d'une bactérie à une autre (voir l'exemple de la bactérie *Vibrio Harveyi* en figure 2).

Une des premières et plus belles expériences qui a démontré l'existence et l'importance de telles fluctuations est celle de

>>>





Novick et Weiner en 1957, à une époque où l'on ne disposait d'aucun outil moléculaire et où le mécanisme même de la régulation génétique n'avait pas encore été découvert. En observant la capacité d'un ensemble d'*E. coli* à digérer le lactose (un sucre), ils ont démontré que les bactéries, toutes issues du même clone (et donc possédant le même génome) et baignant dans exactement le même environnement, formaient deux populations distinctes capables ou non de digérer ce sucre. Fait encore plus troublant, cette capacité, reliée à la concentration d'une enzyme ( $\beta$ -galactosidase) dans la cellule, pouvait se transmettre aux descendants de celle-ci. Nous savons maintenant que les produits des réactions métaboliques, bien qu'initialement sous le contrôle de l'ADN, peuvent avoir deux états stables. Cette bistabilité est due à l'existence de boucles de rétroactions positives. Le « réacteur chimique » bactérien tombe dans l'un ou l'autre de ces états par fluctuation et y demeure ensuite.

D'autres expériences ont continué ces observations ; mais ce n'est qu'avec l'arrivée des outils de la biologie moléculaire et des marqueurs fluorescents tels que la protéine GFP (Green Fluorescent Protein), que l'étendue de la variabilité a pu être caractérisée quantitativement. Par le biais de l'ingénierie moléculaire, on peut, par exemple, forcer une cellule à produire des protéines chimériques qui, d'une part, gardent leur fonction originale et, de l'autre, sont fluorescentes, permettant ainsi de lire optiquement le niveau d'une réaction chimique (fig. 3). À partir des observations au microscope, couplées à l'analyse d'images, on peut alors constituer une statistique sur des centaines, voire des milliers de cellules et avoir accès à la probabilité  $P(n,t)$  d'observer  $n$  protéines dans une cellule au temps  $t$  (voir encadré statistique).

Les différentes mesures (par des outils optiques ou biochimiques) ont permis de démontrer que les fluctuations sont extrêmement répandues ; des amplitudes de fluctuations très importantes ont été observées dans de nombreux systèmes, allant de l'expression de protéines chez les bactéries à des systèmes plus complexes, comme la maturation des ovocytes sous l'effet de la progestérone. Les fluctuations sont le résultat combiné de nombreux facteurs, comme la quantité d'enzymes présente dans l'ensemble de la cellule (voir l'exemple de l'ARN polymérase en figure 1) et le bruit propre de chaque réaction chimique.

Des expériences astucieuses ont permis dans certains cas de mesurer séparément l'influence de ces divers facteurs et d'obtenir parfois des lois de probabilité génériques. Par exemple, pour un simple processus de transcription/traduction, en régime stationnaire, si le nombre de protéines par cellule est suffisamment grand pour que l'on puisse approcher la concentration  $x$  de cette protéine par cellule par une variable continue, la distribution de la concentration  $P(x)$  est décrite par une distribution Gamma (fig. 2). Cette dernière est souvent observée quand une des réalisations d'un processus aléatoire déclenche un autre processus aléatoire.

## Contrôler les fluctuations

Pour fonctionner, un organisme vivant doit pouvoir conserver le même comportement dans un environnement donné et donc contrôler l'amplitude de ses fluctuations. Un nombre important de circuits de régulation génétiques et métaboliques sont dédiés à cette tâche, et nous commençons à comprendre les principes de régulation les plus simples. Pour réguler le niveau d'un certain signal, les ingénieurs humains utilisent la rétroaction négative<sup>(1)</sup> (contre-réaction). Ce mécanisme est également répandu dans le monde du vivant, et nous connaissons chez les bactéries un nombre important de gènes qui s'autorégulent négativement. Une simple autorégulation négative est seulement un des chaînons de la régulation dans le monde vivant ; quand la précision est très importante pour la survie, nous pouvons observer des bijoux d'ingénierie d'une précision inouïe.

Un exemple récent est celui du rythme circadien (oscillation nocturne/diurne) chez la cyanobactérie, responsable de l'apparition de l'oxygène<sup>(2)</sup> sur Terre il y a quelques deux milliards d'années. Cette bactérie, qui est l'organisme le plus simple à posséder un oscillateur biochimique auto-entretenu de période d'environ 24 heures, est capable de conserver la phase de son oscillation sans aucun indicateur extérieur (comme la lumière) pendant plusieurs mois [2]. Pendant longtemps, on a supposé que cette précision résultait d'un phénomène collectif, la communication entre bactéries permettant d'augmenter le nombre effectif de molécules en interaction. Par des expériences où deux populations de bactéries ayant des phases initiales différentes ont été mises ensemble et suivies sur de longues périodes de temps (fig. 4), nous avons

démontré qu'en réalité il n'en était rien [3]. Chaque population garde sa propre phase, donc chaque bactérie est capable, en utilisant peu de molécules, de conserver une excellente précision de phase, ce qui semble défier la thermodynamique.

Un autre exemple, où la précision est importante, est le développement embryonnaire. Lors de cette étape de la vie multicellulaire, différents tissus sont créés à partir des cellules souches et se différencient au fur et à mesure. Il est essentiel pour un animal de contrôler précisément la taille et surtout la proportion de ses membres. Le programme génétique couramment utilisé par la nature pour la différenciation cellulaire est le suivant : un gradient chimique d'une molécule (appelée morphogène) est créé à travers l'embryon ; chaque cellule « lit » ce signal et enclenche, de façon « tout ou rien », sa différenciation en fonction du niveau du morphogène présent à sa position. Chaque cellule différenciée peut à son tour produire un nouveau morphogène et provoquer d'autres cascades de différenciation. Ce programme simple est beaucoup trop sensible à toutes sortes de fluctuations et, *a priori*, ne garantit pas la précision des proportions chez les descendants. Le circuit génétique de base doit donc être complété par des circuits correcteurs d'erreurs. Nous pouvons seulement observer l'extrême efficacité de ces circuits, sans encore entièrement comprendre leur fonctionnement [4].

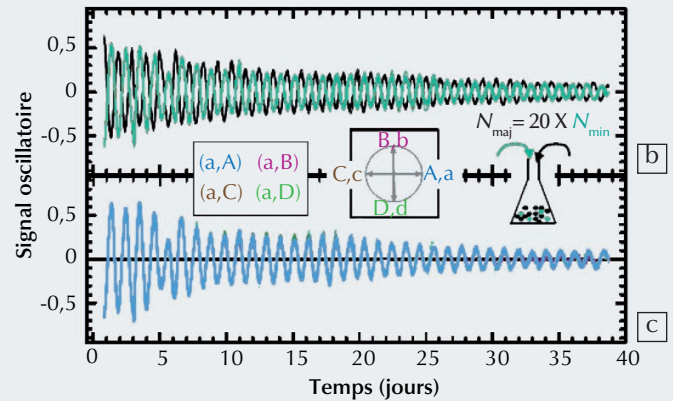
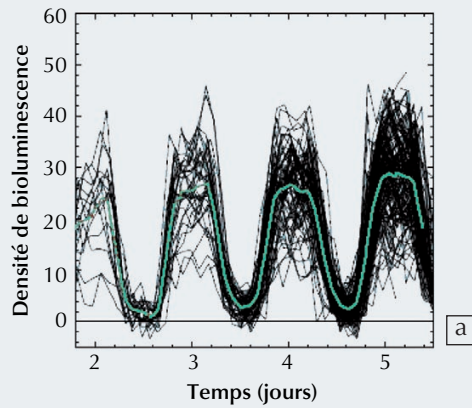
## Utiliser les fluctuations

Le monde vivant sait également exploiter les fluctuations aléatoires. L'exemple le plus évident est, bien sûr, l'évolution darwinienne, où les erreurs aléatoires de duplication d'ADN (les mutations) permettent à un organisme, *via* la sélection naturelle, de s'adapter à son milieu à travers les générations.

Le même phénomène peut également se rencontrer sur des échelles de temps beaucoup plus courtes et a été, par exemple, récemment mis en évidence dans le cas de la persistance bactérienne chez *E. coli*. Cette bactérie peut vivre, soit dans un état métabolique à multiplication rapide en puisant abondamment dans le milieu qui l'entoure, soit dans un état métabolique à multiplication lente en consommant très peu de ressources. Une commutation aléatoire entre les deux états maintient une faible proportion de la population dans l'état défavorable à multiplication lente ; cette faible proportion est en grande partie







#### 4. Exemple d'oscillation synchrone des bactéries individuelles et de leur descendance, suivie par microscopie de bioluminescence.

(a) Sous le contrôle de l'oscillateur central, la production de l'ensemble des protéines chez la cyanobactérie *Synechococcus elongatus* a une composante périodique. Nous suivons ici cette oscillation en utilisant comme rapporteur la bioluminescence. Chaque courbe noire représente le signal émis par une bactérie, la courbe verte la moyenne de ces dernières.

(b) L'oscillation circadienne normalisée de deux populations de cellules indépendantes (domaines vert et noir) ayant une phase initiale opposée garde cette opposition de phase pendant plusieurs semaines. Pour chaque couleur, 8 à 12 courbes expérimentales sont superposées. À noter que seule l'information de phase est pertinente, car l'amplitude est plus sensible à la variation du métabolisme cellulaire tout au long d'une expérience.

(c) Pour tester l'hypothèse d'un couplage entre oscillateurs, nous avons mélangé deux populations de cellules de phases initiales différentes : une « majoritaire » en nombre 20 fois plus grand que l'autre, « minoritaire ». Les mélanges d'une population minoritaire, avec 4 populations majoritaires différentes (domaines bleu, violet, marron et vert) n'affectent pas l'oscillation des minoritaires. (a,A), (a,B), (a,C) et (a,D) désignent respectivement le mélange d'une population majoritaire de phase initiale A, B, C ou D (séparées de  $90^\circ$ ), avec une population des minoritaires de phase A (désignée ici comme a). Chaque mélange a été fait en 8 à 12 exemplaires, tous représentés sur la figure. Dans ce cas, on ne détecte que la population minoritaire a (bleue), car ici la population majoritaire n'est pas munie d'un marqueur bioluminescent. (D'après [3] et [4]).

## Rappels de statistique

Les réactions chimiques sont probabilistes par nature. Considérons, par exemple, une réaction enzymatique élémentaire  $E + B \leftrightarrow E + A$ , où les concentrations de l'enzyme  $E$  et du substrat  $B$  sont à un niveau saturant (maintenues ainsi par une dynamique rapide, comparée à la synthèse des molécules  $A$ ). Ceci peut schématiser une version rudimentaire de la synthèse d'ARN (fig. 1, p. 18).

Pour décrire une telle réaction, nous devons calculer la probabilité  $P(n,t)$  d'observer  $n$  molécules  $A$  à l'instant  $t$ . Connaissant cette fonction, nous pouvons accéder à la moyenne  $\langle n(t) \rangle$  et à la déviation  $\sigma(t)$  que nous appelons *fluctuations* dans cet article. En général, de tels processus obéissent à une statistique de Poisson, et  $\sigma(t)/\langle n(t) \rangle = \langle n \rangle^{-1/2}$ . L'amplitude relative des fluctuations  $\sigma/\langle n \rangle \rightarrow 0$  quand  $\langle n \rangle \gg 1$ .

L'équation de cette réaction peut s'écrire en termes d'évolution de  $P(n,t)$  :  $\partial P(n,t)/\partial t = k_+ [P(n-1,t) - P(n,t)] + k_- [(n+1)P(n+1,t) - nP(n,t)]$ , où  $k_+$  et  $k_-$  sont les taux de production et de destruction des molécules  $A$ . Dans cette réaction, la probabilité de production des molécules  $A$  est indépendante de leur nombre et ne dépend que du substrat et de l'enzyme. Par contre, la probabilité de destruction des molécules  $A$  augmente linéairement avec leur nombre. C'est pourquoi ces deux termes ne sont pas symétriques dans l'équation maîtresse. Les variations temporelles de la moyenne,  $\langle n(t) \rangle = \sum_n n P(n,t)$ , et de la variance,  $\sigma^2(t) = \sum_n (n - \langle n \rangle)^2 P(n,t)$ , peuvent être directement déduites de cette équation :  $d\langle n \rangle/dt = -k_- \langle n \rangle + k_+$ ,  $d\sigma^2/dt = -2k_- \sigma^2 + (k_+ + k_- \langle n \rangle)$ , dont les solutions stationnaires ( $t \rightarrow \infty$ ) sont  $\langle n \rangle = k_+/k_-$ ,  $\sigma^2 = k_+/k_-$  et  $\sigma^2/\langle n \rangle = 1$ . Ceci est caractéristique d'une distribution de Poisson :  $P(n) = e^{-\lambda} (\lambda^n / n!)$ . Ici,  $\lambda = k_+/k_-$ .

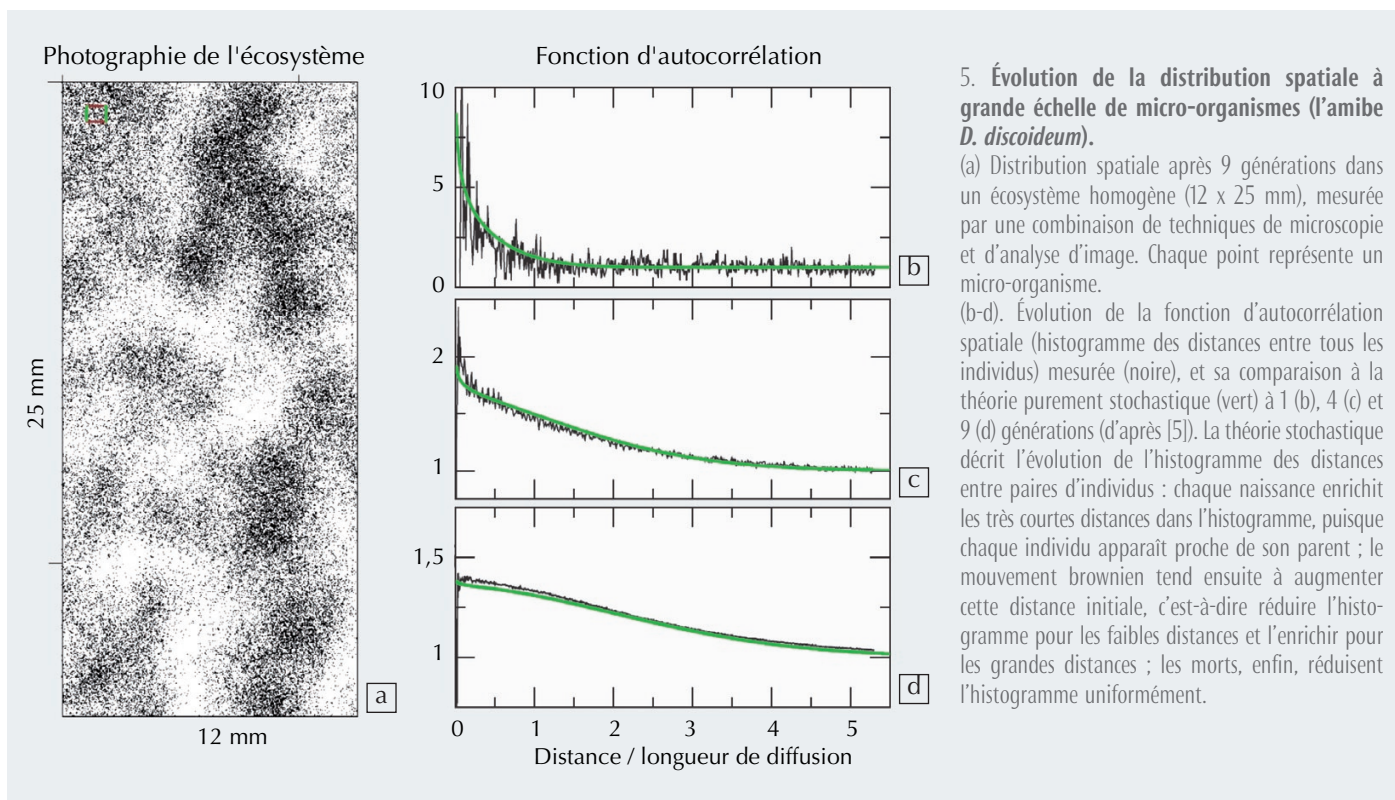
Dans les cellules, un grand nombre de réactions de ce genre sont mises en série. Par exemple, chaque molécule d'ARN produite pilote elle-même la production de plusieurs protéines par des réactions enzymatiques similaires. Le résultat est que la probabilité d'observer  $n$  protéines,  $P_{\text{protéines}}(n)$ , est la composition de deux lois de probabilité, qui souvent peut être approchée par une distribution Gamma :  $P_{\text{protéines}}(n) = n^{a-1} \exp(-n/b) / [b^{a-1} \Gamma(a)]$ . Dans cette fonction,  $n$  est le nombre de protéines (considéré grand),  $a$  le nombre moyen d'ARN généré pendant le temps de vie d'une protéine,  $b$  le nombre moyen de protéines produit pendant le temps de vie d'un ARN et  $\Gamma$  est la fonction Gamma. Quand le nombre moyen de protéines est grand, la variable  $n$  peut être considérée comme une variable continue.



le résultat de la différence entre les taux de multiplication dans les deux états. En présence d'antibiotiques dans le milieu, les bactéries se trouvant dans l'état à multiplication rapide meurent ; les autres, communiquant peu avec l'extérieur, survivent encore longtemps. Une fois l'antibiotique enlevé, une proportion de bactéries à multiplication lente bascule aléatoirement dans le mode rapide et repeuple la niche écologique. De façon générale, il apparaît que dans les environnements variables un gène dont le produit est plus aléatoire peut conférer à l'organisme un avantage.

## Fluctuations à grande échelle spatiale

Les phénomènes de fluctuations que nous venons de décrire à l'échelle de la cellule ou de l'organisme peuvent avoir des conséquences sur des échelles spatiales bien plus vastes. Un exemple important est celui de la répartition des espèces dans la nature : de façon quasi générale, les individus d'une espèce donnée ne sont pas répartis de façon homogène dans l'espace, mais présentent systématiquement des distributions agglomérées. Ce phénomène avait d'abord été remarqué dans le cas des planctons à la surface des océans ; depuis,



### 5. Évolution de la distribution spatiale à grande échelle de micro-organismes (l'amibe *D. discoideum*).

(a) Distribution spatiale après 9 générations dans un écosystème homogène (12 x 25 mm), mesurée par une combinaison de techniques de microscopie et d'analyse d'image. Chaque point représente un micro-organisme.

(b-d). Évolution de la fonction d'autocorrélation spatiale (histogramme des distances entre tous les individus) mesurée (noire), et sa comparaison à la théorie purement stochastique (vert) à 1 (b), 4 (c) et 9 (d) générations (d'après [5]). La théorie stochastique décrit l'évolution de l'histogramme des distances entre paires d'individus : chaque naissance enrichit les très courtes distances dans l'histogramme, puisque chaque individu apparaît proche de son parent ; le mouvement brownien tend ensuite à augmenter cette distance initiale, c'est-à-dire réduire l'histogramme pour les faibles distances et l'enrichir pour les grandes distances ; les morts, enfin, réduisent l'histogramme uniformément.

sur les centaines d'espèces de divers règnes étudiés à différentes échelles, seules quelques-unes suivent une répartition purement « poissonnienne ». Les théories classiques de l'écologie cherchent l'origine de ces agglomérats dans l'hétérogénéité géographique (lacs, montagnes, gradients de dénivelés...) et l'adaptation des espèces aux niches. Cependant, ces fluctuations peuvent simplement provenir des phénomènes aléatoires asymétriques que sont les morts et les naissances : on naît proche de ses parents, mais on peut mourir partout, favorisant ainsi les courtes distances entre individus. La diffusion et les mouvements aléatoires ne sont pas suffisants pour lisser les hétérogénéités créées par ces phénomènes stochastiques.

Nous venons de démontrer, dans un écosystème expérimental contrôlé et parfaitement homogène, que la distribution des individus devient, au bout de quelques générations, extrêmement agglomérée et exactement prédictible par des théories probabilistes ne tenant compte que des naissances, morts et migrations [5]. Observer des hétérogénéités spatiales de distribution des espèces ne doit donc pas être surprenant *a priori* (fig. 5).

## Conclusion

Les fluctuations stochastiques sont devenues un thème central de la biologie, grâce à une collaboration active entre, d'une part, les physiciens contribuant par les concepts et les outils de la physique statistique et des processus stochastiques et, d'autre part, les biologistes apportant leur compréhension de la complexité du monde vivant et la maîtrise des outils modernes de la biologie moléculaire. Les scientifiques ont ainsi compris que pour contrôler les fluctuations, là où la précision est vitale, les cellules dédient une part importante de leurs ressources à des « circuits » correcteurs d'erreur d'un grand raffinement ; mais aussi, que là où la précision n'est pas vitale, une certaine variabilité est *maintenue* pour conférer à l'organisme un avantage dans un environnement extérieur fluctuant.

Baucoup de chemin reste à parcourir pour avoir une vue générale des fluctuations biologiques, semblable à celle que la physique statistique nous confère pour les fluctuations thermiques. Nous avons besoin de plus de données expérimentales pour étoffer nos connaissances, ainsi que d'outils théoriques pour analyser les circuits génétiques / métaboliques complexes que le monde vivant utilise. Nous ne sommes qu'au début de l'exploration de ce domaine. ■

## Références

- 1 • A. Raj, A. van Oudenaarden, "Nature, nurture, or chance: stochastic gene expression and its consequences", *Cell*, **135** (2008) 216-226.
- 2 • B. Houchmandzadeh, E. Wieschaus, S. Leibler, "Establishment of developmental precision and proportions in the early *Drosophila* embryo", *Nature*, **415** (2002) 798-802.
- 3 • I. Mihalcescu, W. Hsing, S. Leibler, "Resilient circadian oscillator revealed in individual cyanobacteria", *Nature*, **430** (2004) 81-85.
- 4 • M. Amdaoud, M. Vallade, C. Weiss-Schaber, I. Mihalcescu, "Cyanobacterial clock, a stable phase oscillator with negligible intercellular coupling", *Proc Natl Acad Sci U S A*, **104** (2007) 7051-7056.
- 5 • B. Houchmandzadeh, "Neutral clustering in a simple experimental ecological community", *Phys. Rev. Lett.*, **101** (2008) 078103.

(1) En théorie de contrôle, la régulation négative, ou boucle de rétroaction négative, permet qu'une partie du signal de sortie d'un dispositif soit réinjectée à l'entrée avec un signe inversé, stabilisant ainsi la réponse du dispositif.

(2) L'atmosphère terrestre ne contenait pas d'oxygène jusqu'à il y a environ 2 milliards d'années. Le succès évolutif des organismes photosynthétiques (essentiellement des cyanobactéries) produisant l'oxygène comme un déchet de la photosynthèse, a globalement modifié la composition de l'atmosphère qui actuellement contient environ 20% d'oxygène.