

> Les mutations bénéfiques à forts effets sont rares et les mutations délétères sont éliminées par la sélection naturelle. La majorité des mutations qui s'accumulent dans les génomes ont donc des effets sélectifs très faibles, voire nuls ; elles sont alors appelées mutations neutres. Au cours des deux dernières décennies, il a été montré que les mutations, même en l'absence d'effet sur la valeur sélective des organismes, affectent leur évolvabilité¹, en donnant accès à de nouveaux phénotypes par le biais de mutations apparaissant ultérieurement, et qui n'auraient pas été disponibles autrement. En plus de cet effet, de nombreuses mutations neutres – indépendamment de leurs effets sélectifs – peuvent affecter la mutabilité de séquences d'ADN voisines, et moduler l'efficacité de la recombinaison homologue. De telles mutations ne modifient pas le spectre des phénotypes accessibles, mais plutôt la vitesse à laquelle de nouveaux phénotypes seront produits, un processus qui a des conséquences à long terme mais aussi potentiellement à court terme, en lien avec l'émergence de cancers. <

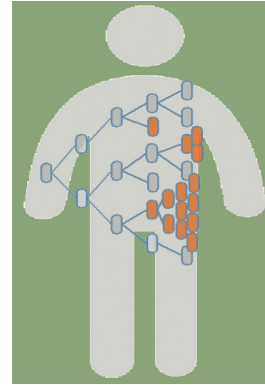
Les mutations sont des modifications permanentes du matériel génétique et représentent la principale source de diversité phénotypique héritable, qui est le substrat de la sélection naturelle. Dans un environnement donné, les mutations peuvent être classées selon leur impact sur la valeur sélective : elles peuvent être létales, délétères, neutres ou bénéfiques. Un formalisme simple consiste à attribuer aux mutations un coefficient de sélection « s », de sorte que la valeur sélective d'un mutant par rapport à son parent est de $1 + s$. Les valeurs de « s » des mutations bénéfiques, délétères, létales et neutres sont ainsi respectivement supérieures à 0, inférieures à 0, égales à -1 et égales à 0. Ces catégories de mutations diffèrent par

Vignette (© Olivier Tenaillon).

¹ L'évolvabilité désigne la capacité d'un organisme à générer de la variation phénotypique héritable favorisant l'adaptation.

L'impact des mutations neutres sur l'évolvabilité et l'évolution des génomes

Olivier Tenaillon¹, Ivan Matic²



¹Université Paris Cité et université Sorbonne Paris Nord, Inserm UMR 1137, Infection, Antimicrobiens, Modélisation, Évolution (IAME), F-75018 Paris, France.

²Université Paris Cité, Inserm U1016, CNRS UMR8104, Institut Cochin, F-75014 Paris, France. olivier.tenaillon@inserm.fr

la fréquence à laquelle elles se produisent, mais aussi par l'importance qu'elles revêtent pour les chercheurs. Les mutations très bénéfiques sont rares, mais elles sont fortement recherchées par les scientifiques pour comprendre l'adaptation. Les mutations délétères et létales, relativement abondantes, sont éliminées par la sélection naturelle, mais elles sont précieuses pour comprendre les fonctions des gènes ou les déterminants moléculaires des maladies. Les mutations très avantageuses étant rares, et les mutations délétères étant éliminées, les mutations qui s'accumulent dans les génomes au cours du temps (on dit qu'elles se fixent et on parle alors de substitutions) sont, en conséquence, pour la plupart, des mutations dont les effets sélectifs sont faibles, voire neutres. Elles sont pourtant très utiles pour construire des généalogies ou des phylogénies. En effet, une mutation d'effet sélectif faible a un destin évolutif qui est largement régi par le hasard. Ce processus aboutit à une fixation quasi constante de substitutions qui dépend principalement du taux de mutation de l'organisme. En supposant que ces taux de mutation sont assez constants entre les taxons, le nombre de substitutions entre deux génomes peut donc être utilisé pour déduire le temps qui s'est écoulé depuis l'existence de leur dernier ancêtre commun. Cependant, si ces nombreuses substitutions n'ont que des effets sélectifs nuls ou très faibles, elles peuvent, comme nous allons le discuter dans cette revue, affecter les trajectoires évolutives, indépendamment de leurs effets sélectifs propres [1].

Mutations neutres, variation cryptique et évolvabilité

Les domaines de la génétique des populations et, plus tard, de l'évolution moléculaire, en se concentrant sur le sort d'un allèle donné, ont



principalement utilisé la fixation des mutations neutres comme marqueurs du temps. Pourtant, l'intégration des mutations neutres dans la métaphore du paysage adaptatif, a permis de relier les mutations neutres à l'évolvabilité, c'est-à-dire à la capacité des organismes à générer une diversité adaptative [2-4]. Un paysage adaptatif est une représentation intégrée des connexions entre les génotypes possibles et leur succès reproductif associé, ce que l'on nomme valeur sélective. Dans cette perspective, l'accent n'est pas mis sur le sort d'une mutation donnée, mais plutôt sur l'ensemble des mutations qui se propageront dans une population compte tenu de toutes les mutations possibles et de leurs interactions. L'impact d'une mutation ne se limite donc pas à son impact sur la valeur sélective mais affecte aussi le panel de génotypes et de phénotypes auquel elle donne accès. L'évolution est alors vue comme une marche qui conduit une population d'une position dans le paysage adaptatif à une autre, chaque étape de la marche donnant accès à un ensemble légèrement différent de mutations et, par conséquent, d'effets sélectifs. La sélection qui favorise la fixation de mutations bénéfiques est une force motrice majeure du mouvement dans le paysage adaptatif. Pourtant, dans l'hypothèse où les mutations bénéfiques sont rares, l'évolution par fixation de mutations neutres est apparue comme un moyen efficace de se déplacer dans le paysage, les mutations délétères étant, elles, éliminées par la sélection naturelle. Pour des raisons mathématiques et conceptuelles, il a souvent été proposé, par la suite, de limiter les paysages adaptatifs à des réseaux neutres dans lesquels le paysage adaptatif est uniquement composé de mutations neutres ou très délétères [5, 6]. Le paysage adaptatif est alors composé de réseaux de génotypes de valeurs sélectives égales, interconnectés par des mutations. Dans ces modèles conceptuels, la connectivité de chaque génotype, ou la taille du réseau neutre accessible par des mutations neutres, a été reliée à l'évolvabilité [2-4]. L'idée sous-jacente simple était que sur un grand réseau neutre, les populations peuvent accumuler des niveaux élevés de diversité génétique, permettant ainsi d'explorer ce grand réseau. Elles peuvent ainsi avoir accès à un plus grand nombre de génotypes par le biais de mutations, dont certaines pourront être bénéfiques dans de nouvelles conditions. En utilisant une approche de modélisation, il a par ailleurs été montré que le temps nécessaire pour trouver un nouveau phénotype spécifique était lié à la taille de ce réseau de mutations neutres.

Quels types de mécanismes moléculaires permettent-ils aux mutations neutres de modifier l'éventail des phénotypes accessibles ?

Les modèles de repliement de l'ARN ont joué un rôle essentiel dans notre compréhension de la biologie de ces processus. Avec l'essor des programmes informatiques prédisant le repliement des petites séquences d'ARN (voir par exemple [7]), il est devenu possible d'étudier *in silico* l'évolution des séquences d'ARN sous une contrainte sélective agissant sur le repliement que prendra la séquence. Pour un repliement d'ARN donné, on peut attribuer à toute séquence une valeur sélective de 1, si la séquence est prédite comme se repliant préférentiellement dans un format particulier, et de 0 dans le cas contraire. Les séquences qui adoptent le même repliement peuvent différer par

une succession de mutations neutres : des mutations qui n'ont pas modifié le repliement. Si ces séquences partagent le même repliement, il n'en est pas nécessairement de même pour les séquences avoisinantes. L'accessibilité d'un nouveau repliement d'ARN, à partir d'un repliement préexistant, a été étudiée à partir de ces réseaux neutres [5]. Des approches similaires ont été utilisées avec des modèles de réseaux de protéines, qui simulent le repliement, dans des réseaux bidimensionnels ou tridimensionnels, de courtes séquences de protéines, mais aussi des réseaux de gènes régulateurs, qui génèrent un modèle d'expression génétique, et des analyses de flux, qui prédisent la capacité de croissance d'un réseau métabolique en fonction de différents substrats [8]. Avec le développement de la mutagenèse et du séquençage à haut débit, certaines approches expérimentales ont même validé le principe selon lequel une variation génétique neutre pourrait favoriser l'adaptation. Par exemple, des populations de bactéries *Escherichia coli*, qui avaient accumulé une certaine variation neutre dans une protéine fluorescente émettant initialement dans le jaune, se sont adaptées plus rapidement que des populations n'ayant pas cette diversité préexistante, lors d'une sélection pour obtenir un changement du spectre d'émission de la protéine vers le vert [9].

Dans ces études des paysages adaptatifs, les mutations neutres sont essentiellement une source de variation phénotypique cryptique². Dans la plupart des cas, c'est la dégénérescence de la carte de valeur sélective-génotype-phénotype qui permet à certaines mutations de modifier le génotype, mais pas la valeur sélective. L'idée sous-jacente est que ces mutations neutres modifient avant tout le génotype (et, éventuellement, certains phénotypes non sélectionnés), mais que les nouvelles séquences auxquelles ce changement de génotype donne accès, peuvent, elles, avoir des phénotypes et des valeurs sélectives différents. Un exemple simple de dégénérescence génotype-phénotype se trouve dans le code génétique, dans lequel certaines mutations génèrent des codons synonymes, c'est-à-dire qui codent le même acide aminé. Bien qu'elles ne modifient pas la séquence protéique, les mutations synonymes peuvent donner accès à de nouveaux phénotypes, en changeant le type d'acides aminés accessibles par des mutations ultérieures. Cette différence dans les ensembles d'acides aminés accessibles est appelée *volatilité des codons*. Par exemple, une mutation du codon CGG, codant l'arginine, en son synonyme AGG, change la gamme d'acides aminés accessibles à ce site

² Sans que ce phénotype diffère selon la variation génotypique.

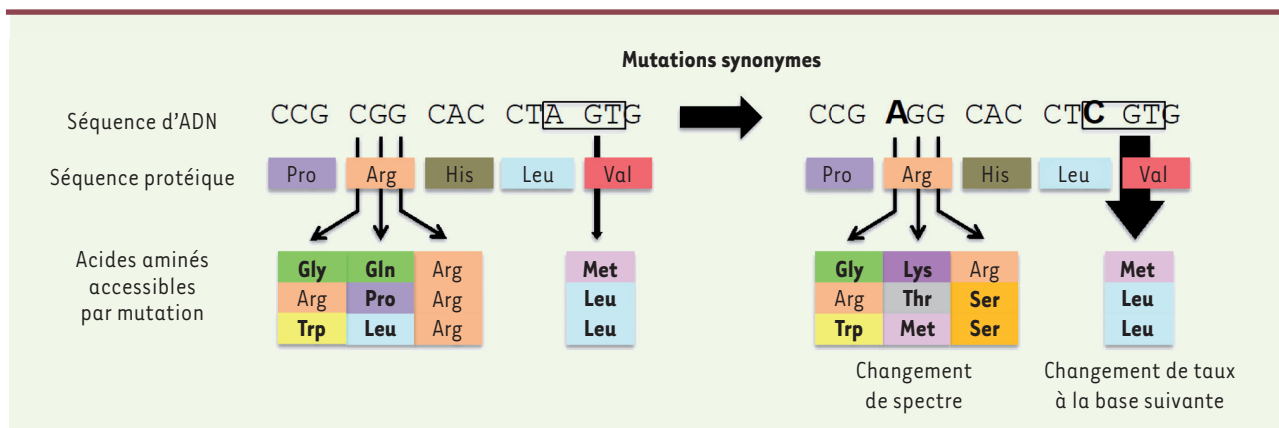


Figure 1. Impact des mutations synonymes sur l'évolvabilité. Une séquence d'ADN codante est mutée avec deux mutations synonymes, indiquées en gras. Bien que ces mutations n'affectent pas la séquence protéique, la première (CGG > AGG) affecte les acides aminés qui sont maintenant accessibles par mutations ultérieures (les acides aminés spécifiques des différents codons synonymes sont en gras) et la seconde (CTA > CTC) affecte le taux de mutation à la base suivante, en raison d'un changement dans le schéma de séquence local d'AGT à CGT, qui est beaucoup plus mutagène.

en cas de mutations ultérieures : de la glycine, du tryptophane, de la glutamine, de la leucine et de la proline, à la glycine, au tryptophane, à la lysine, à la thréonine, à la méthionine et à la sérine (Figure 1). La variabilité de la volatilité des codons a été utilisée pour améliorer la variabilité génétique accessible par mutagenèse aléatoire dans les protéines [10]. Ainsi, chez la bactérie, des mutants synonymes d'un gène de résistance aux antibiotiques, lorsqu'ils ont été soumis à une mutagenèse aléatoire par PCR (*polymerase chain reaction*)³, ont conduit à l'apparition de mutations produisant différentes améliorations de la fonction de la protéine responsable de la résistance.

De nombreux mécanismes alternatifs peuvent conduire à une certaine dégénérescence de la correspondance entre génotype et phénotype et servir de sources de variation cryptique. Par exemple, le flux dans une voie métabolique avec des enzymes fortement exprimées n'est pas affecté par des changements modérés de l'expression des enzymes dans des conditions non limitantes de substrat [11, 12]. De même, une protéine très stable n'est quasiment pas affectée par des mutations qui modifient sa stabilité. Certaines mutations, apparemment neutres, peuvent ainsi moduler l'expression des gènes ou la stabilité des protéines, mais aussi peuvent permettre de tolérer d'autres mutations dont l'effet sera plus important, y compris celles qui apportent de nouvelles fonctions [13]. Par exemple, dans la bêta-lactamase TEM-1 de la bactérie *E. coli*, la mutation stabilisatrice M182T (la méthionine en position 182 substituée par une thréonine) a un effet marginal sur la valeur sélective, mais elle permet l'accumulation ultérieure de nombreuses mutations déstabilisantes [14], dont certaines étendent l'activité de TEM-1 à un spectre plus large de molécules antibiotiques [15]. Dans tous les cas, dans la métaphore du paysage adaptatif, les mutations neutres sont liées à l'évolvabilité par le fait qu'elles modifient qualitativement le panel de mutations et de phénotypes acces-

sibles en raison d'une certaine dégénérescence dans la carte génotype-phénotype.

Cet impact des mutations neutres sur l'évolvabilité est maintenant bien documenté [16]. Dans la suite de cette revue, nous nous concentrerons sur un autre impact des mutations neutres qui est souvent négligé, à savoir, l'impact qu'elles ont sur les taux de mutation et de recombinaison, ce qui est important pour l'évolution. Contrairement aux cas que nous avons évoqués ici, les mutations ne changent pas le spectre des phénotypes accessibles, mais plutôt le degré auquel ces phénotypes mutants seront produits. En effet, indépendamment de leurs impacts directs ou indirects sur les phénotypes sélectionnés, les mutations, en affectant la machinerie de réplication, de réparation et de recombinaison de l'ADN, peuvent affecter le destin évolutif à long terme des espèces.

Impact du polymorphisme local de la séquence d'ADN sur la mutabilité

La séquence nucléotidique locale a un impact sur la fidélité de l'ADN polymérase, l'efficacité de la réparation de l'ADN et sur la mutabilité par différents agents mutagènes. Il n'est donc pas surprenant que l'analyse des polymorphismes naturels dans les génomes microbiens et eucaryotes, les génomes de cellules cancéreuses, montrent que le contexte de la séquence d'ADN a un impact significatif sur la probabilité de génération de mutations. Des études réalisées sur le génome de bactéries *E. coli* déficientes dans des systèmes de réparation de l'ADN, ont révélé que l'ADN polymérase introduit des mutations préférentiellement au niveau de certaines séquences spécifiques [17]. On

³ La technique de mutagenèse aléatoire par PCR repose sur l'utilisation d'une polymérase peu fidèle qui a de forte chance d'introduire des mutations lors des multiples cycles d'élongation réalisés lors de l'amplification par PCR.

a ainsi constaté que les mutations affectant une guanine (G) dans une séquence CGT étaient environ 10 fois plus probables que les mutations affectant une guanine dans une séquence AGT [18]. Similairement, des biais d'ampleur variable sont observés chez les primates dans les taux de substitution d'une même paire de bases dans différents environnements de séquence locale [19]. Dans la lignée germinale humaine, la composition en bases, jusqu'à deux nucléotides de distance d'une base focale, influence significativement le taux de substitution [20]. Une mutation, même sans effet sur le phénotype ou la valeur sélective, peut donc influencer le taux d'erreur de la polymérase sur les bases voisines.

La séquence d'ADN locale a également un impact sur l'efficacité de la réparation de l'ADN, et donc, sur la mutabilité de cette séquence. Par exemple, l'efficacité de la réparation des mésappariements de bases chez *E. coli* dépend non seulement de la nature du mésappariement mais aussi des séquences nucléotidiques qui lui sont proches. Ainsi, l'efficacité de la réparation augmente avec le contenu en G:C de la séquence d'ADN qui flanque le mésappariement [21]. Chez les eucaryotes, la séquence de nucléotides locale a également un impact sur l'efficacité de la réparation par excision de bases et sur la fidélité de l'ADN polymérase [22, 23]. Par exemple, la thymine glycosylase humaine élimine plus efficacement la thymidine des mésappariements G:T (normalement G:C), qui résultent de la désamination de la 5-méthylcytosine, lorsqu'elle se trouve dans un contexte de séquence CpG, par rapport aux contextes ApG, CpC et TpC.

Les propriétés thermodynamiques et conformationnelles intrinsèques de différentes séquences d'ADN dictent ainsi la fidélité de la réplification de l'ADN et l'efficacité de sa réparation. Mais les mutations peuvent également altérer la stabilité des structures secondaires locales de l'ADN formées pendant la transcription. Cela affecte le temps passé par certaines bases dans l'état ADN simple brin, ce qui les rend chimiquement moins stables que les bases plus protégées dans l'état double brin de l'ADN et modifie donc leur mutabilité [24, 25]. Il est important de noter que ces effets mécanistiques sur le taux de mutation contribuent à l'accumulation de diversité à long terme, ce que nous montre la génomique comparative. Par exemple, l'impact d'une base voisine sur la fidélité de la polymérase, révélé dans des études faites au laboratoire, est fortement corrélé aux biais observés dans les mutations accumulées entre deux souches d'*E. coli* distantes, qui divergent d'environ 3 % [17]. Ces biais mécanistiques dans les taux de mutation ont donc des signatures durables.

Impact du polymorphisme de la séquence d'ADN sur l'efficacité de la recombinaison

La recombinaison homologe maintient la stabilité du génome en réparant les cassures de l'ADN et en sauvant les fourches de réplification bloquées ; mais elle génère également de nouvelles combinaisons de séquences d'ADN, par transfert horizontal de gènes⁴ chez les bactéries,

et lors de la méiose chez les eucaryotes. Il existe de nombreuses preuves que le polymorphisme des séquences d'ADN, indépendamment de son impact sur la valeur sélective, a un impact majeur sur l'efficacité de la recombinaison homologe. Le processus de recombinaison homologe est initié par l'action combinée d'une variété de nucléases et d'hélicases qui génèrent de l'ADN simple brin, qui est le substrat d'une famille d'ADN recombinases conservées au cours de l'évolution, dont RecA chez les bactéries, et Rad51 chez les eucaryotes [26]. Ces protéines se lient d'abord à l'ADN simple brin pour former des filaments nucléoprotéiques. Elles s'engagent ensuite dans une recherche de séquences d'ADN cibles. La recherche des sites cibles dans l'ADN est très efficace et très précise, même lorsque ces sites sont extrêmement rares dans le génome. Le seuil minimum pour la reconnaissance d'une séquence d'ADN par ces protéines est de huit nucléotides consécutifs parfaitement appariés. Ce seuil est conservé entre la RecA bactérienne et la Rad51 de levure et humaine [27]. Lorsqu'une longueur donnée d'identité de séquence partagée et ininterrompue est trouvée, les ADN recombinases catalysent l'appariement de l'ADN des filaments de nucléoprotéines avec le brin complémentaire de l'ADN double brin, générant ainsi de l'ADN hétéroduplex (formé de brins d'origines différentes) (Figure 2A). Cette réaction, appelée *invasion des brins*, nécessite au moins 26 nucléotides chez les bactéries, et 250 chez les eucaryotes [28]. Cette longueur minimale d'identité de séquence d'ADN requise pour la recombinaison homologe est appelée *segment minimal de traitement efficace*. L'appariement de séquences d'ADN non identiques produit des molécules hétéroduplex qui sont mal appariées (Figure 2B). Les recombinases d'ADN RecA et Rad51 peuvent détecter des mésappariements simples aux étapes initiales de l'échange de brins d'ADN. Les mésappariements près de l'extrémité 3' du brin entrant empêchent fortement l'échange de brin. Les divergences de séquence d'ADN entre les molécules d'ADN qui recombinent réduisent ainsi le nombre de segment minimal de traitement efficace et donc l'efficacité de la recombinaison homologe [29, 30].

Une fois établi, l'ADN hétéroduplex est étendu par une étape de migration unidirectionnelle des branches. À cette étape de la recombinaison, les ADN recombinases tolèrent la divergence de séquence nucléotidique entre les molécules recombinantes. Cependant, la non identité entre les molécules d'ADN qui recombinent, en aval du segment minimal de traitement efficace, entraîne la génération d'hétéroduplex où les deux brins peuvent différer (Figure 2B). Il en résulte la création d'une

⁴ Processus dans lequel un organisme intègre du matériel génétique provenant d'un autre organisme sans en être le descendant.

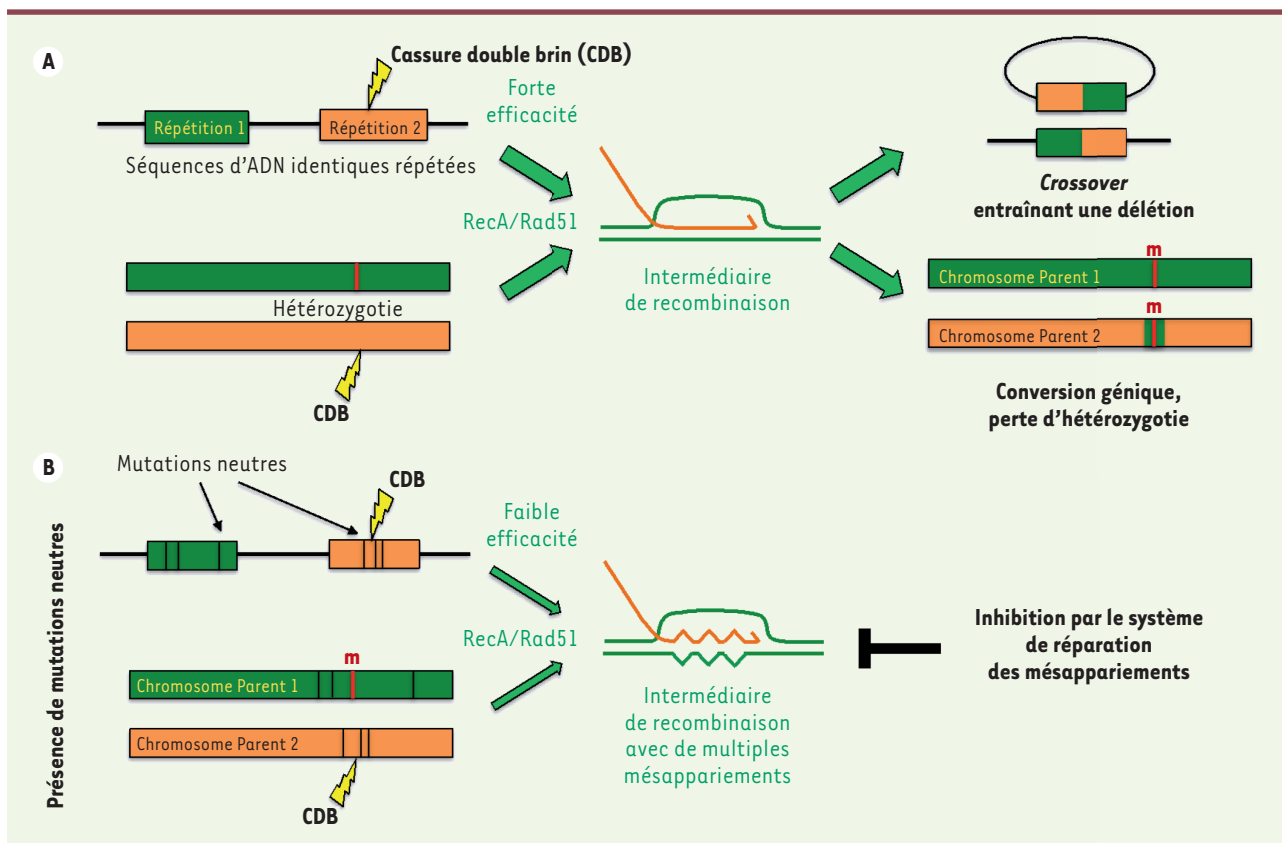


Figure 2. Impact des mutations synonymes sur l'efficacité de la réparation des cassures double brin par recombinaison homologue. **A.** La réparation d'une cassure double brin (CDB) par recombinaison homologue peut entraîner des réarrangements génomiques potentiellement délétères, comme des délétions et une perte d'hétérozygotie. **B.** Des mutations neutres présentes dans la région d'une cassure double-brin peuvent diminuer l'efficacité de la réparation double-brin, car elles affectent l'activité des enzymes de recombinaison et déclenchent l'activité anti-recombinogène des enzymes de réparation des mésappariements.

molécule d'ADN portant des paires de bases non complémentaires qui sont reconnues et détruites efficacement par le système de réparation des mésappariements. Le système de réparation des mésappariements agit comme un élément de fidélité, qui détermine la longueur du segment minimal de traitement efficace en imposant une rigueur accrue pendant la recherche d'homologie catalysée par les filaments nucléoprotéiques [29, 30]. Il suffit d'un seul mésappariement pour que le système de réparation des mésappariements inhibe la recombinaison [31], et des mésappariements supplémentaires ont un effet négatif cumulatif sur l'efficacité de la recombinaison [30]. Le degré et la distribution de la divergence des séquences d'ADN entre les molécules qui recombinent sont donc des paramètres structurels qui influencent l'efficacité de la recombinaison, car ils affectent l'activité des enzymes de recombinaison et déterminent si l'activité anti-recombinogène du système de réparation des mésappariements doit être déclenchée. Les études des effets de la divergence des séquences d'ADN sur l'efficacité de la recombinaison homologue révèlent l'existence d'une relation entre la diminution exponentielle de l'efficacité de la recombinaison et l'augmentation linéaire de la divergence de l'ADN entre molécules recombinantes. Cet effet peut donc avoir un large éventail d'impacts

sur la dynamique du génome à des échelles de temps courtes ou longues.

Effet à court terme des mutations neutres sur les taux de mutation et de perte d'hétérozygotie

La recombinaison ectopique, intra- et inter-chromosomique, entre des séquences d'ADN répétées situées à des emplacements différents, peut entraîner des altérations de la structure du génome chez les procaryotes et chez les eucaryotes. La recombinaison entre chromosomes homologues peut, quant à elle, entraîner la perte d'hétérozygotie chez les eucaryotes. Ces deux types de réarrangements chromosomiques doivent être évités autant que possible car ils affectent la structure même du chromosome et peuvent générer des cellules homozygotes portant deux allèles délétères. L'expression d'une mutation hétérozygote récessive chez des souris consanguines, donc hautement homozygotes, se produit à une fréquence élevée en raison de la perte

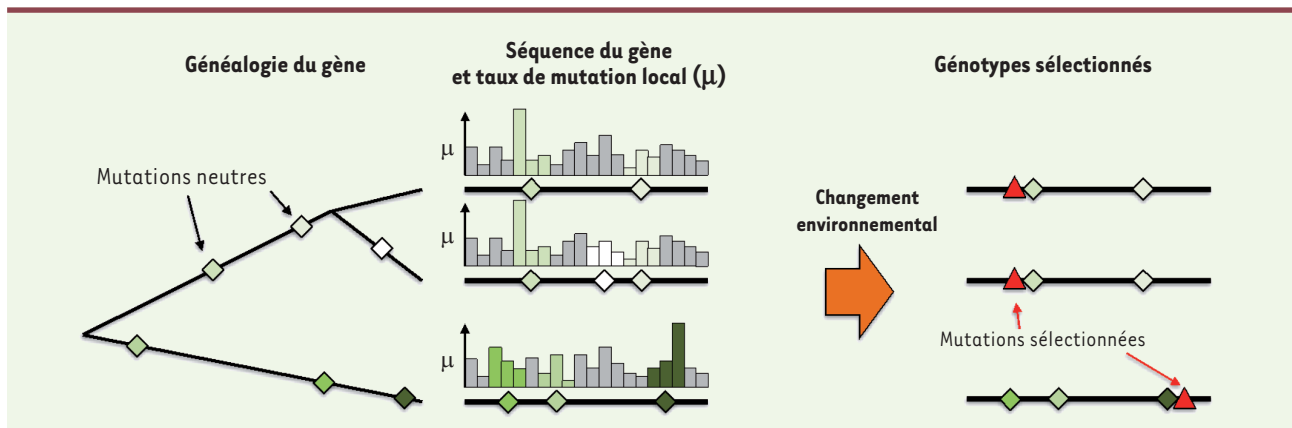


Figure 3. L'accumulation de polymorphismes dans les génomes peut affecter le taux de mutation local dans une fraction non négligeable du génome et conduire au recrutement de différentes mutations face à une nouvelle pression sélective. Nous montrons ici la généalogie de trois séquences dans lesquelles des mutations neutres (diamants colorés) se sont accumulées. Les séquences résultantes diffèrent par leur taux de mutation local, comme le montre la hauteur des histogrammes, qui sont colorés en fonction de la couleur de la mutation entraînant les changements du taux de mutation local. Lorsqu'un changement environnemental est appliqué indépendamment à ces différents clones, ces changements de taux locaux peuvent conduire à la sélection de différentes mutations bénéfiques illustrées par des triangles rouges et entraîner une différenciation supplémentaire du génome.

d'hétérozygotie par recombinaison mitotique au sein des cellules somatiques [32] (Figure 2). Dans les hybrides issus de souches de souris de parentés distantes, cette fréquence de perte d'hétérozygotie est beaucoup plus faible, en raison de la divergence de la séquence d'ADN entre les allèles maternels et paternels, qui affecte l'efficacité de la recombinaison mitotique. Une perte d'hétérozygotie peut entraîner de forts effets sélectifs, car elle permet aux allèles de devenir homozygotes, et donc d'avoir un impact sur le phénotype, s'ils sont récessifs. Cela peut s'avérer bénéfique, en facilitant la sélection de mutations récessives bénéfiques [33]. Cependant, le même processus peut s'avérer très délétère, par exemple dans le cas d'une perte d'hétérozygotie au niveau des gènes suppresseurs de tumeurs, car le gain de valeur sélective au niveau cellulaire peut entraîner le développement de cancers [34].

Les variations des taux de mutation ou de perte d'hétérozygotie dues à la présence de mutations neutres sont substantielles : jusqu'à trois fois plus de perte d'hétérozygotie entre des souris consanguines et des souris non consanguines [32], ou jusqu'à une augmentation d'un facteur dix du taux de mutation sur des sites voisins [18]. Pourtant, une question reste ouverte : dans quelles conditions évolutives ces changements de taux ont-ils un impact sur l'évolution ?

La fixation de mutations neutres est un processus lent. Mais alors, dans quelles conditions ces mutations peuvent-elles avoir un impact ? Il semble que les conditions adéquates soient les mêmes que celles qui sont nécessaires pour que la variation génétique cryptique ait un impact sur l'évolution : des populations qui présentent une grande diversité, en raison de leur taille importante, d'un taux de mutation élevé, ou des deux. Dans ces conditions, la grande diversité génétique au sein de la population entraînera une variabilité entre les différents sites polymorphes des taux de mutation locaux, ce qui contribuera

à l'augmentation de la diversité. Considérons, par exemple, une espèce bactérienne dont la divergence génomique entre les souches peut atteindre 3 % dans la partie conservée de leur génome. En supposant qu'une mutation affecte le taux de mutation sur la base focale et sur les deux bases voisines, les taux différeront entre deux souches échantillonnées au hasard sur 9 % de leur génome. Lorsque ces souches divergentes seront confrontées à un nouveau défi, tel qu'un traitement antibiotique, le panel de mutations quantitativement accessibles sera plus large. Appliquée indépendamment à différentes populations divergentes, une telle sélection peut conduire au recrutement de différentes mutations bénéfiques pour l'adaptation et contribuer ensuite à une plus grande diversification des espèces (Figure 3). Une illustration frappante de cet effet a été observée chez la bactérie *Pseudomonas fluorescens* [35]. Dans un mutant non motile de la souche SBW25 de cette bactérie, la restauration de la mobilité se produit à plus de 95 % grâce à la même mutation (A289C : l'adénine 289 remplacée par une cytosine) dans le gène *ntrB*, qui code une histidine kinase. Cette réponse ciblée est spécifique de SBW25 et n'est pas observée dans des souches qui diffèrent de SBW25 par quelques mutations synonymes dans le gène *ntrB*. L'introduction dans SBW25 de six mutations synonymes, trouvées dans d'autres souches, dans le voisinage proche de la position 289 du gène, abolit ce point chaud de mutation, et la mutation A289C n'est alors plus détectée parmi les mutations compensatrices. Ces effets pourraient être encore plus

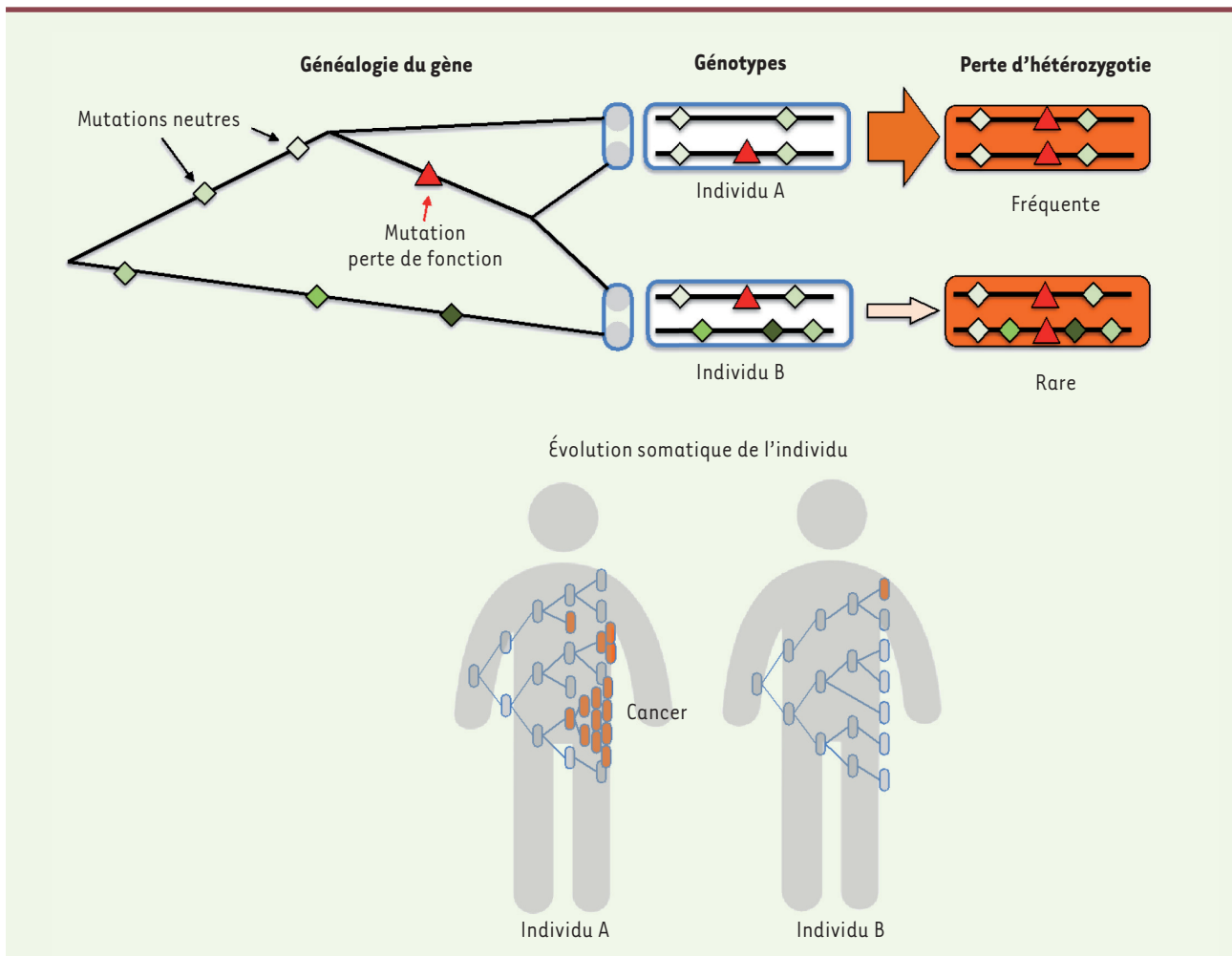


Figure 4. La diversité des séquences parentales présentes chez un individu peut affecter le taux de perte d'hétérozygotie et par conséquent le taux de cancers. Nous montrons ici la généalogie des deux copies d'un gène chez deux individus diploïdes. Les deux copies de l'individu A ont un ancêtre commun récent et ne diffèrent que par une seule mutation, le triangle rouge, alors que les deux copies de l'individu B diffèrent par de nombreuses mutations car leur ancêtre commun est beaucoup plus ancien. La différence entre ces temps de coalescence peut être due au hasard ou au fait que l'individu B est issu d'un mélange de deux populations précédemment isolées. La différence de similarité entre les deux copies entraîne une différence importante (jusqu'à trois fois) dans la vitesse à laquelle se produit la perte d'hétérozygotie. Si la mutation du triangle rouge favorise l'apparition d'un cancer lorsqu'elle est présente en deux copies, l'individu A est beaucoup plus susceptible que l'individu B de développer un cancer, même si les deux sont hétérozygotes à l'identique au niveau de la mutation causale.

importants pour les populations virales qui présentent à la fois un taux de mutation et une taille de population élevés.

Pour étudier l'impact des mutations sur la variation du taux de perte d'hétérozygotie, nous pouvons nous concentrer sur l'évolution dans la lignée somatique d'un organisme. Considérons un allèle sur un site qui peut faciliter l'évolution vers le cancer lorsqu'il est homozygote. Si l'ancêtre commun le plus récent du fragment chromosomique portant cet allèle est récent, les deux copies seront similaires sur une longue distance (sauf pour la mutation d'intérêt) et la perte d'hétérozygotie sera élevée, entraînant un risque élevé de cancer. En revanche, si l'ancêtre commun le plus récent est ancien, par hasard ou si l'individu résulte d'un mélange entre des populations divergentes comme on

l'observe fréquemment chez l'homme [36], les mutations accumulées entre les deux chromosomes réduiront la perte d'hétérozygotie et le risque de cancer sera alors faible (Figure 4).

Effet à long terme des mutations sur le transfert horizontal de gènes et la spéciation

L'impact des mutations neutres peut également affecter quantitativement d'autres processus importants de la recombinaison, impliquant cette fois la recombinaison de séquences divergentes provenant d'un

autre organisme. Chez les bactéries, par exemple, la divergence des séquences d'ADN participe au contrôle du transfert horizontal de gènes. La relation entre l'efficacité de la recombinaison et la divergence des séquences d'ADN a été observée pour le transfert par conjugaison de gènes⁵ chez les bactéries Gram négatives [37] et pour la transformation⁶ des espèces Gram positives *Bacillus* et *Streptococcus* [38, 39]. Ces barrières peuvent jouer un rôle important en bloquant la propagation d'allèles adaptatifs entre différentes souches d'une même espèce, et entre différentes espèces.

Outre le fait qu'ils constituent une barrière à l'intégration d'un fragment d'ADN étranger, ces mécanismes peuvent également contribuer au processus de spéciation à long terme, qui est défini comme la perte du flux génétique entre lignées. Il a été proposé, par exemple, que l'effet anti-recombinaison du polymorphisme neutre était suffisant pour entraîner la spéciation sympatrique⁷ chez les bactéries, à condition que le taux de recombinaison soit suffisamment faible [40, 41]. La spéciation nécessite l'accumulation d'incompatibilités fonctionnelles entre les lignées pour finalement arrêter le flux de gènes. Il est intéressant de noter que ce modèle de spéciation ne nécessite pas l'accumulation de différences phénotypiques entre les populations. L'accumulation d'une diversité neutre peut suffire à créer des barrières durables aux échanges génétiques et à la poursuite de l'accumulation de la diversité génétique. Un exemple concret de la façon dont la divergence des séquences d'ADN pourrait participer à l'établissement et au maintien de barrières génétiques est fourni par l'impact de la divergence des séquences pendant la méiose chez la levure. Il a été démontré que les croisements entre deux espèces de levures étroitement apparentées, *Saccharomyces cerevisiae* et *S. paradoxus*, ne produisent que 1 % de spores viables, dont la plupart donnent naissance à des cellules présentant des fréquences élevées de réarrangements chromosomiques qui produisent des colonies de petite taille et à croissance lente [42]. Cette stérilité, observée dans de nombreux croisements interspécies, est une base de la définition de l'espèce biologique dans la taxonomie des levures. Cependant, le croisement de *S. cerevisiae* et de *S. paradoxus* dépourvus de systèmes fonctionnels de réparation des mésappariements, entraîne une augmentation de la recombinaison méiotique, une diminution des réarrangements chromosomiques et une amélioration de la viabilité des spores. Ces observations suggèrent que la divergence de l'ADN, associée à une activité anti-recombinaison du système de réparation des mésappariements pendant la méiose, peut contribuer à l'établissement de barrières post-zygotiques entre les espèces eucaryotes.

Conclusions

Nous avons présenté dans cette revue les preuves suggérant que les mutations neutres contribuent à l'évolution de deux manières non

exclusives. Premièrement, certaines mutations, sans effet détectable sur la valeur sélective en raison de la dégénérescence de la carte génotype-phénotype, peuvent modifier les phénotypes accessibles par des mutations ultérieures. Deuxièmement, certaines mutations - indépendamment de leur impact sur les phénotypes sélectionnés - ont un impact sur la machinerie moléculaire de la réplication, de la réparation et de la recombinaison de l'ADN. Cela peut entraîner des changements importants dans les taux locaux de mutation et de recombinaison. Alors que le premier point a fait l'objet de recherches intenses de la part des biologistes de l'évolution au cours des deux dernières décennies, le second a été en grande partie le domaine privé des généticiens moléculaires. L'impact des changements de taux de mutation sur l'évolution est une question qui doit maintenant être mieux intégrée, tant au niveau théorique qu'expérimental, afin de mieux quantifier son importance sur les trajectoires évolutives. ♦

SUMMARY

The impact of neutral mutations on genome evolvability

Beneficial mutations with strong effects are rare and deleterious mutations are purged by natural selection. Therefore, the majority of mutations that accumulate in genomes have very weak or no selective effects, being then called neutral mutations. Over the last two decades, it has been shown that mutations, even when they are neutral, affect evolvability by providing access to new phenotypes through later-occurring mutations that would not have been available otherwise. We propose here that in addition to this effect, many mutations -independent of their selective effects- can affect the mutability of neighboring DNA sequences and modulate the efficiency of homologous recombination. Such mutations do not alter the spectrum of accessible phenotypes, but rather the rate at which new phenotypes will be produced, a process that has long-term but also potentially short-term consequences for cancer emergence. ♦

REMERCIEMENTS

IM est soutenu par les financements : FRM DBF20160635736, Labex « Who am I? » Idex ANR-11-IDEX-0005-01 / ANR-11-LABX-0071-Idex-Sorbonne Paris, ANR-18-CE35-007-02 et ANR-21-CE12-0006. OT est soutenu par les financements : ANR Gewiep : ANR-18-CE35-0005-01 et par la Fondation pour la Recherche Médicale : FRM EQU201903007848.

LIENS D'INTÉRÊT

Les auteurs déclarent n'avoir aucun lien d'intérêt concernant les données publiées dans cet article.

⁵ Transmission d'ADN plasmidique ou chromosomique d'une bactérie donneuse (porteuse d'un plasmide) à une bactérie receveuse.

⁶ Intégration d'un fragment d'ADN étranger dans une cellule.

⁷ La spéciation sympatrique est un processus de spéciation par lequel de nouvelles espèces très proches émergent d'un ancêtre commun, alors qu'elles habitent la même région géographique.

RÉFÉRENCES

1. Tenaillon O, Matic I. The impact of neutral mutations on genome evolvability. *Curr Biol* 2020 ; 30 : R527-4.
2. Wagner A. Robustness and evolvability: A paradox resolved. *Proc R Soc B Biol Sci* 2008 ; 275 : 91-100.
3. Lenski RE, Barrick JE, Ofria C. Balancing robustness and evolvability. *PLoS Biol* 2006 ; 4 : e428.
4. Visser JA de, Hermissin J, Wagner GP, et al. Perspective: Evolution and detection of genetic robustness. *Evol Int J Org Evol* 2003 ; 57 : 1959-72.
5. Huynen MA, Stadler PF, Fontana W. Smoothness within ruggedness: the role of neutrality in adaptation. *Proc Natl Acad Sci U S A* 1996 ; 93 : 397-401.
6. Wilke CO. Adaptive evolution on neutral networks. *Bull Math Biol* 2001 ; 63 : 715-30.
7. Hofacker IL, Fontana W, Stadler PF, et al. Fast folding and comparison of RNA secondary structures. *Monatshfte Für Chem Chem Mon* 1994 ; 125 : 167-88.
8. Wagner A. The molecular origins of evolutionary innovations. *Trends Genet* 2011 ; 27 : 397-410.
9. Zheng J, Payne JL, Wagner A. Cryptic genetic variation accelerates evolution by opening access to diverse adaptive peaks. *Science* 2019 ; 365 : 347-53.
10. Cambrey G, Mazel D. Synonymous genes explore different evolutionary landscapes. *PLoS Genet* 2008 ; 4 : e1000256.
11. Dykhuizen DE, Dean AM, Hartl DL. Metabolic flux and fitness. *Genetics* 1987 ; 115 : 25-31.
12. Kemble H, Eisenhauer C, Couce A, et al. Flux, toxicity, and expression costs generate complex genetic interactions in a metabolic pathway. *Sci Adv* 2020 ; 6 : eabb2236.
13. Tokuriki N, Stricher F, Serrano L, et al. How protein stability and new functions trade off. *PLoS Comput Biol* 2008 ; 4 : e1000002.
14. Jacquier H, Birgy A, Nagard HL, et al. Capturing the mutational landscape of the beta-lactamase TEM-1. *Proc Natl Acad Sci U S A* 2013 ; 110 : 13067-72.
15. Weinreich DM, Delaney NF, Depristo MA, et al. Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science* 2006 ; 312 : 111-4.
16. Payne JL, Wagner A. The causes of evolvability and their evolution. *Nat Rev Genet* 2019 ; 20 : 24-38.
17. Couce A, Caudwell LV, Feinauer C, et al. Mutator genomes decay, despite sustained fitness gains, in a long-term experiment with bacteria. *Proc Natl Acad Sci U S A* 2017 ; 114 : E9026-35.
18. Sung W, Ackerman MS, Gout J-F, et al. Asymmetric context-dependent mutation patterns revealed through mutation-accumulation experiments. *Mol Biol Evol* 2015 ; 32 : 1672-83.
19. Blake RD, Hess ST, Nicholson-Tuell J. The influence of nearest neighbors on the rate and pattern of spontaneous point mutations. *J Mol Evol* 1992 ; 34 : 189-200.
20. Krawczak M, Ball EV, Cooper DN. Neighboring-nucleotide effects on the rates of germ-line single-base-pair substitution in human genes. *Am J Hum Genet* 1998 ; 63 : 474-88.
21. Radman M, Wagner R. Mismatch repair in Escherichia Coli. *Annu Rev Genet* 1986 ; 20 : 523-38.
22. Donigan KA, Sweasy JB. Sequence context-specific mutagenesis and base excision repair. *Mol Carcinog* 2009 ; 48 : 362-8.
23. Mendelman LV, Boosalis MS, Petruska J, et al. Nearest neighbor influences on DNA polymerase insertion fidelity. *J Biol Chem* 1989 ; 264 : 14415-23.
24. Hoede C, Denamur E, Tenaillon O. Selection acts on DNA secondary structures to decrease transcriptional mutagenesis. *PLoS Genet* 2006 ; 2.
25. Wright BE, Reschke DK, Schmidt KH, et al. Predicting mutation frequencies in stem-loop structures of derepressed genes: implications for evolution. *Mol Microbiol* 2003 ; 48 : 429-41.
26. Kowalczykowski SC. An overview of the molecular mechanisms of recombinational DNA repair. *Cold Spring Harb. Perspect Biol* 2015 ; 7 : a016410.
27. Lee JY, Terakawa T, Qi Z, et al. Base triplet stepping by the Rad51/RecA family of recombinases. *Science* 2015 ; 349 : 977-81.
28. Rubnitz J, Subramani S. The minimum amount of homology required for homologous recombination in mammalian cells. *Mol Cell Biol* 1984 ; 4 : 2253-8.
29. Sagi D, Tlusty T, Stavans J. High fidelity of RecA-catalyzed recombination: a watchdog of genetic diversity. *Nucleic Acids Res* 2006 ; 34 : 5021.
30. Vulic M, Dionisio F, Taddei F, et al. Molecular keys to speciation: DNA polymorphism and the control of genetic exchange in enterobacteria. *Proc Natl Acad Sci U S A* 1997 ; 94 : 9763-7.
31. Shen P, Huang HV. Effect of base pair mismatches on recombination via the RecBCD pathway. *Mol Gen Genet* 1989 ; 218 : 358-60.
32. Shao C, Stambrook PJ, Tischfield JA. Mitotic recombination is suppressed by chromosomal divergence in hybrids of distantly related mouse strains. *Nat Genet* 2001 ; 28 : 169-72.
33. Mandegar Mohammad A, Otto Sarah P. Mitotic recombination counteracts the benefits of genetic segregation. *Proc R Soc B Biol Sci* 2007 ; 274 : 1301-7.
34. Ryland GL, Doyle MA, Goode D, et al. Loss of heterozygosity: What is it good for? *BMC Med Genomics* 2015 ; 8 : 45.
35. Horton JS, Flanagan LM, Jackson RW, et al. A mutational hotspot that determines highly repeatable evolution can be built and broken by silent genetic changes. *Nat Commun* 2021 ; 12 : 6092.
36. Reich D. *Who We Are and How We Got Here: Ancient DNA and the new science of the human past*. Oxford : Oxford University Press, 2018 : 368 p.
37. Delmas S, Matic I. Cellular response to horizontally transferred DNA in Escherichia coli is tuned by DNA repair systems. *DNA Repair* 2005 ; 4 : 221-9.
38. Zawadzki P, Roberts MS, Cohan FM. The log-linear relationship between sexual isolation and sequence divergence in Bacillus transformation is robust. *Genetics* 1995 ; 140 : 917-32.
39. Majewski J, Zawadzki P, Pickerill P, et al. Barriers to genetic exchange between bacterial species: Streptococcus pneumoniae transformation. *J Bacteriol* 2000 ; 182 : 1016-23.
40. Dixit PD, Pang TY, Maslov S. Recombination-driven genome evolution and stability of bacterial species. *Genetics* 2017 ; 207 : 281-95.
41. Falush Daniel, Torpdahl Mia, Didelot Xavier, et al. Mismatch induced speciation in Salmonella: model and data. *Philos Trans R Soc B Biol Sci* 2006 ; 361 : 2045-53.
42. Hunter N, Chambers SR, Louis EJ, et al. The mismatch repair system contributes to meiotic sterility in an interspecific yeast hybrid. *EMBO J* 1996 ; 15 : 1726-33.

TIRÉS À PART
O. Tenaillon



Avec **m/s**, vivez en direct les progrès et débats de la biologie et de la médecine

CHAQUE MOIS / AVEC LES ARTICLES DE RÉFÉRENCE DE M/S
CHAQUE JOUR / SUR WWW.MEDECINESCIENCES.ORG

Abonnez-vous sur
www.medecinesciences.org