La marche génétique de l'évolution^{*}

Laurent Angelo Arnoult

Aix-Marseille Université, CNRS, UMR 7288, Institut de Biologie du Développement de Marseille-Luminy, 13288 Marseille Cedex 9, France

Auteur correspondant : Laurent Angelo Arnoult, arnoult.laurent@gmail.com

Reçu le 5 septembre 2014

Résumé – Les mutations génétiques constituent le moteur principal de l'évolution. À chaque génération, elles produisent de nouvelles variations, qui sont triées par sélection naturelle ou sexuelle. Les mutations sont certes générées au hasard ; mais quelles sont celles que l'évolution retient et pourquoi ? Cette revue a pour objectif de présenter certains résultats récents qui viennent éclairer cette question. Nous nous pencherons d'abord sur les résultats obtenus aux niveaux moléculaire et cellulaire, grâce à des approches synthétiques et des paradigmes d'évolution expérimentale. Puis, nous soulignerons l'apport constitué par les études de l'évolution au niveau pluricellulaire, notamment lorsque l'évolution se répète, c'est-à-dire, lorsqu'elle produit dans des lignées indépendantes des traits similaires. Nous présenterons les résultats dérivés des premières méta-analyses et quantifications dans le domaine. Ces dernières suggèrent que les mutations retenues au cours de l'évolution s'accumulent souvent autour de « points chauds », situés à différents niveaux d'organisation génétique, ce que nous illustrons par un exemple concret. Nous montrerons comment certains travaux suggèrent qu'une multitude de causes, correspondant à une multitude de contextes biologiques différents, peuvent expliquer l'existence de ces points chauds génétiques. Enfin, nous discuterons des limites méthodologiques, des défis empiriques et des horizons de ce pan de la recherche consacrée à la marche génétique de l'évolution.

Mots clés : Génétique / convergence / évolution répétée / méta-analyse / points chauds

Abstract – The genetic walk of evolution.

Genetic mutations are the main fuel of evolution. In each generation, they produce new variations, which may be sorted out by natural or sexual selection. Mutations are generated by chance; yet which are the mutations actually sorted out by evolution, and why? This review presents some recent advances regarding this question. First, we gather results obtained at molecular and cellular levels, through synthetic experiments and under artificial selection paradigms. Next, we highlight studies at the multi-cellular level, especially studies of repeated evolution, whereby independent lineages acquire similar traits. Recent meta-analysis and quantifications are being presented; together they suggest that evolutionary relevant mutations accumulate around hotspots, spanning different levels of genetic organization. Pioneering work suggests that many causes, corresponding to many biological contexts, may explain the existence of these genetic hotspots. We finally discuss methodological limits, empirical challenges and a few future potential directions for this domain of research dedicated to the genetic path of evolution.

Key words: Genetics / convergence / repeated evolution / meta-analysis / hotspots

^{*} Cet article est issu d'une thèse de Doctorat de l'Université Paris-Descartes, intitulée "Bricolage évolutif chez des drosophiles à pois noirs" soutenue le 17 mai 2013. Cette thèse a reçu le Prix 2014 de la Fédération Réaumur des Sciences de la Vie.

Introduction

Le monde biologique se caractérise par une diversité exubérante à tous les niveaux d'organisation du vivant, ainsi que par la troublante et fréquente correspondance entre les formes vivantes et leur milieu de vie. La synthèse moderne de l'évolution propose une explication puissante de ces deux caractéristiques : l'arrivée permanente de mutations aléatoires crée un large spectre de variations, dont certaines favorisent la survie et la reproduction. La sélection et d'autres forces évolutives trient ces germes héréditaires de changement, de génération en génération. Sous certaines conditions, la sélection forge ainsi les adaptations des individus à leur environnement (Lewontin, 1978).

On décrit souvent les mutations aléatoires comme le carburant génétique de l'évolution, et la sélection comme son moteur (López-Fanjul & García-Dorado, 2011); néanmoins la sélection n'opère jamais qu'à travers le filtre du développement, au cours duquel les variations génétiques sont traduites en variations phénotypiques. Le développement, entendu comme la correspondance entre génotype et phénotype, détermine en effet les capacités du vivant à varier, c'est-à-dire détermine de quelle manière chaque mutation affectera le caractère phénotypique soumis à la sélection. Les mutations sont certes générées au hasard, mais le développement fixe pour partie celles que l'évolution retiendra.

C'est pourquoi tant d'efforts sont dévolus à comprendre le développement afin d'éclairer notre compréhension de l'évolution, plus particulièrement afin de résoudre les questions suivantes.

Le développement d'un caractère phénotypique définit-il le jeu des variations possibles ? Les mutations qui pavent les trajets adaptatifs sont-elles en partie déterminées par l'état ancestral du caractère ?

Si une telle relation déterminante entre un caractère et la distribution des mutations qui peuvent le modifier existe, cela rend-il l'évolution plus contingente ou plus déterministe ? Peut-on isoler des caractéristiques communes à ces mutations, parfois appelées « mutations pertinentes au regard de l'évolution » (Stern & Orgogozo, 2008) ? Ces mutations ont-elles un effet de faible ou de grande amplitude sur le phénotype qui évolue ? Peut-on distinguer les mutations responsables de la variation entre individus d'une même espèce (échelle micro-évolutive) *versus* les mutations responsables de la variation entre espèces (échelle macro-évolutive) ?

Dans les années 1970 à 1990, la littérature se réfère à ces questions en parlant de contraintes développementales (Jacob, 1977; Gould & Lewontin, 1979; Smith *et al.*, 1985), dans les années 2000 et 2010, les mêmes questions se regroupent sous la formule

de la prédictibilité génétique de l'évolution (Stern & Orgogozo, 2009).

La résolution de ces questions passe principalement par une approche expérimentale. Il existe bien une littérature théorique, qui étudie d'un point de vue statistique les caractéristiques génétiques de chemins adaptatifs (Orr, 2005a, 2005b), mais il y a très peu d'interactions entre ce corpus théorique et le corpus expérimental auquel nous consacrons cette revue (sauf par exemple dans l'étude de Rokyta *et al.*, 2005).

J'évoque d'abord deux études menées aux niveaux moléculaire et cellulaire, utilisant des approches synthétiques et d'évolution expérimentale, notamment parce qu'elles illustrent le mouvement de la recherche vers des standards de preuves fonctionnelles plus élevés (Laubichler & Maienschein, 2007). Puis je ciblerai mon analyse sur les systèmes multicellulaires, où de nombreux progrès ont été réalisés. Pour reprendre une formulation de François Jacob, quels sont « les outils que le développement de l'embryon fournit au bricolage de l'évolution » (Jacob, 1981) ? Souvent caricaturée, cette question a longtemps été réduite en une opposition entre partisans de l'importance des mutations codantes d'une part (Carroll, 2008) *versus* partisans de l'importance des mutations régulatrices d'autre part (Hoekstra & Coyne, 2007). Nous montrerons comment l'apport de l'étude de l'évolution répétée, les premières tentatives de quantification et l'émergence d'axes d'analyse différents ont déplacé cette opposition et ont soulevé de nouvelles questions.

Enfin nous évoquerons certaines perspectives de ce champ d'investigation, ses limites actuelles et ses promesses de développement futur.

Étudier l'évolution dans un tube

Au niveau moléculaire, des chemins étroits et prédictibles

En se focalisant sur un trait moléculaire simple, la séquence d'une enzyme par exemple, et en reconstituant *in vitro* l'ensemble des mutations intermédiaires possibles entre une situation ancestrale et une situation dérivée correspondant à un maximum de *fitness* (ou valeur adaptative) connu, il est possible de reconstruire, mutation par mutation, l'intégralité des chemins adaptatifs possibles.

Par exemple, Weinreich *et al.* (2006) se sont focalisés sur la β -lactamase, une enzyme bactérienne qui catalyse l'hydrolyse d'antibiotiques comme la pénicilline ou la céfotaxime. Cinq mutations, lorsqu'elles coexistent, augmentent la résistance à la céfotaxime d'un facteur 100 000. Les auteurs de l'étude montrent que cette combinaison allélique constitue un maximum unique de *fitness*. Il y a donc

a priori 5!, soit 120 façons de monter vers cet optimum, mais combien de trajets adaptatifs exactement la nature emprunte-t-elle ?

Les auteurs calculent la résistance à la céfotaxime des 120 allèles intermédiaires possibles ; seuls 18 trajets mènent de la séquence sauvage à la séquence hautement résistante en ne passant que par des intermédiaires de plus en plus résistants. Utilisant deux modélisations indépendantes du régime de sélection, les auteurs calculent la probabilité de mise en œuvre de chacun de ces trajets. L'évolution a ainsi une chance sur deux de passer par moins de 4 trajets.

S'il existe peu de trajets, c'est que l'ordre de fixation des mutations importe. L'effet sur la *fitness* d'une mutation à une position donnée varie selon l'état allélique aux quatre autres positions, un phénomène appelé épistasie. À vrai dire, la plupart des mutations peuvent avoir un effet positif ou négatif sur la *fitness* selon l'état allélique aux 4 autres positions, un phénomène appelé épistasie de signe.

Si l'ordre de fixation des mutations importe, c'est en partie pour des raisons biologiques propres à la fonction de l'enzyme. Par exemple, la mutation G238S dans un contexte sauvage accélère l'hydrolyse de la céfotaxime mais stimule l'agrégation protéique ; inversement, dans le même contexte, M182T a une faible influence cinétique mais réduit l'agrégation protéique. À cause de ce double rôle fonctionnel à l'intérieur de l'enzyme – un cas simple de pléiotropie –, la mutation M182T réduit la résistance à la céfotaxime dans le contexte sauvage, mais l'augmente en combinaison avec la mutation G238S.

Ainsi, quelques études au niveau moléculaire, passées en revue par Poelwijk *et al.* (2007), montrent que des contraintes fonctionnelles dirigent souvent l'évolution sur des trajets adaptatifs peu nombreux. Établir l'architecture fonctionnelle des molécules étudiées permet *a posteriori* d'expliquer le détail des chemins empruntés et pourrait donc permettre *a priori* de les prédire.

Au niveau cellulaire, des trajets morcelés et contingents

Depuis 1988, le laboratoire de R. Lenski sélectionne en parallèle 12 populations d'*E. coli* pour un taux de croissance accéléré sur un milieu limité en glucose. Toutes les 500 générations, des bactéries de toutes les populations sont conservées au congélateur, ménageant la possibilité de les ramener à la vie. Ce dispositif d'évolution expérimentale permet ainsi de reconstituer avec précision le contenu génétique de nombreux chemins adaptatifs indépendants.

Le milieu de culture est tamponné par du citrate, une source métabolique potentielle qu'une *E. coli*

sauvage est incapable d'utiliser en condition aérobie, un caractère dénommé Cit⁻. En effet, dans cette condition, l'expression du transporteur citrate est activement inhibée (Blount *et al.*, 2008).

De façon spectaculaire, après plus de 31 000 générations, une seule population d'*E. coli* a commencé à exploiter le citrate dans les conditions aérobies de l'expérience, un caractère nouveau dénommé Cit⁺. Il existe de nombreuses définitions de la nouveauté évolutive, passées en revue par Pigliucci (2008). Nouveauté évolutive ou pas, le caractère Cit⁺ est bien un trait qualitativement différent qui a ouvert une niche écologique et promu la diversification des bactéries. Quelle est la composition génétique du chemin adaptatif qui a mené à Cit⁺ ?

En 30 000 générations, chaque population a subi plusieurs milliards de mutations communes, un nombre largement supérieur à la taille du génome de la bactérie. En d'autres termes, toutes les mutations communes sont passées isolément plusieurs fois au crible de la sélection naturelle. L'émergence du trait Cit⁺ requiert donc soit un événement génétique extrêmement rare (comme une inversion chromosomique massive), soit une combinaison très particulière de plusieurs mutations.

Pour distinguer ces deux hypothèses, des bactéries Cit⁻ congelées de cette population ont été remises en culture pour environ 3 700 générations. Le trait Cit⁺ a émergé à 18 reprises différentes, mais seulement à partir des populations fondées par un clone provenant d'une génération récente. Il semblerait donc que l'émergence du trait Cit⁺ ne soit possible que dans un fonds génétique précis, qui s'est progressivement construit au cours de générations antérieures. Le séquençage du génome complet de 29 colonies d'âges différents a permis de reconstruire leurs relations phylogénétiques, comme le montre la figure 1 (Blount *et al.*, 2012).

Au moment précis de son émergence, le trait Cit⁺ ne confère à la bactérie qu'une efficacité d'exploitation très limitée du citrate et donc un avantage sélectif très faible. Très rapidement, le trait Cit⁺ s'est optimisé et l'exploitation du citrate est devenu très efficace, expliquant pourquoi les bactéries Cit⁺ sont devenues majoritaires, 2000 générations seulement après leur apparition (Blount *et al.*, 2012).

Le chemin adaptatif semble pouvoir se décomposer en trois phases, indiquées sur la figure 1. Une phase de *potentialisation*, pendant laquelle s'accumulent des mutations sans effets directs sur le métabolisme du citrate, mais qui rendent possible l'émergence du caractère. Une phase d'*actualisation*, correspondant à l'arrivée de la première mutation qui confère à la bactérie la capacité d'exploiter le citrate. Une phase de *raffinement*, où la sélection naturelle améliore rapidement l'efficacité d'exploitation du citrate.

Fig. 1. L'évolution de l'exploitation du citrate en trois étapes génétiques. Différentes lignées ont évolué depuis l'ancêtre fondateur, dont les relations phylogénétiques ont été révélées par séquençage. L'émergence du caractère Cit⁺ semble être un processus en trois étapes. Pendant la potentialisation, des mutations s'accumulent qui augmentent la probabilité d'actualisation du trait. Une fois actualisé, le trait est raffiné par la sélection naturelle. Si l'actualisation du caractère Cit⁺ n'est survenue que dans une lignée précise après environ 31 000 générations, au moins trois autres lignées sont passées par la phase de potentialisation. Adapté de Hendrickson & Rainey (2012).

Ces différentes phases correspondent-elles à des événements génétiques de différentes natures ? De façon remarquable, dans les 19 différentes bactéries nouvellement Cit⁺, la mutation d'actualisation se produit toujours au même *locus* (l'opéron qui contient le transporteur du citrate), provoque toujours le même changement (une expression du transporteur en condition aérobie), mais sa nature moléculaire précise n'est jamais la même (Blount *et al.*, 2012).

À elle seule, cette mutation d'actualisation est incapable d'entraîner l'émergence du trait Cit⁺, ce qui renforce l'idée d'un ensemble de relations épistatiques avec d'autres mutations. Malheureusement, la nature exacte des mutations responsables de cette potentialisation n'est pas encore déterminée. Il n'existe pas davantage de données précises sur les mutations responsables du raffinement du caractère.

Néanmoins, le contraste entre la signature évidente qu'ont laissée dans toutes les populations Cit⁺ les mutations actualisatrices, et la difficulté d'identification des mutations des phases de potentialisation et de raffinement, suggèrent que ces dernières sont sans doute nombreuses et probablement différentes dans chaque trajectoire indépendante qui a vu l'émergence d'individus Cit⁺.

Au niveau morphologique, de nombreuses questions ouvertes

Les traits moléculaires sont les seuls dont la correspondance génotype → phénotype est univoque : il s'agit du code génétique. Par ailleurs, la plupart des contraintes pertinentes s'exercent à l'intérieur de la séquence et sont d'ordre biochimique ou biophysique : elles sont donc relativement faciles à étudier. Au niveau cellulaire, les traits sont polygéniques et leur émergence semble conditionnée à l'arrivée ordonnée d'un jeu de mutations épistatiques (Bridgham *et al.*, 2009 ; Breen *et al.*, 2012).

La puissance des modèles expérimentaux aux niveaux moléculaire et cellulaire permet à l'expérimentateur de conduire des tests fonctionnels à l'intérieur d'un environnement contrôlé, de façon hautement répliquable (Dean & Thornton, 2007) ; identifier et tester la contribution d'une ou d'un groupe de mutations responsables d'un changement évolutif chez des espèces pluricellulaires est un défi majeur, notamment parce que la correspondance entre génotype et phénotype s'établit à travers le filtre complexe du développement embryonnaire.

Un cas de figure particulièrement propice et expérimentalement fécond pour relever ce défi consiste

à étudier l'évolution répétée (Gompel & Prud'homme, 2009) que nous abordons à la section suivante.

Étudier l'évolution quand elle se répète

Un cadre d'étude (presque) idéal

Hors du champ de l'évolution expérimentale, il existe une configuration phylogénétique particulièrement favorable, qui permet de tester le caractère prédictif de l'évolution au niveau génétique chez les organismes pluricellulaires.

En effet, dans des lignées différentes, l'évolution produit souvent des solutions phénotypiques similaires en réponse à des problèmes sélectifs identiques (Gompel & Prud'homme, 2009; Nadeau & Jiggins, 2010; Conte & Arnegard, 2012; Martin & Orgogozo, 2013; Stern, 2013) ainsi par exemple l'invention de l'écholocation qui a eu lieu indépendamment chez les cétacés et chez les chauves-souris (Liu *et al.*, 2010).

Ces cas d'évolution répétée, un terme appelé à remplacer ceux de « convergence » et « parallélisme » (Gompel & Prud'homme, 2009), offrent une opportunité précieuse pour comparer le détail génétique de trajectoires évolutives indépendantes. Si effectivement le développement limite le jeu des trajets génétiques possibles, lorsque l'évolution se répète au niveau phénotypique, la probabilité qu'elle passe par deux chemins génétiques similaires devrait être élevée.

Pour que l'étude de l'évolution répétée serve ce dessein, encore faut-il être certain que l'évolution se répète bien de façon indépendante au niveau génétique, autrement dit, que les mutations responsables du changement répété aient une origine distincte. La tâche n'est pas toujours facile. Il faut d'abord pouvoir polariser la direction du changement évolutif; dans l'exemple de l'écholocation, il faut établir qu'il s'agit de deux apparitions indépendantes, plutôt qu'une apparition chez un ancêtre commun aux cétacés et aux chauves-souris, suivis de multiples retours à l'état ancestral chez leurs cousins mammifères. Cette étape est souvent réalisée en appliquant le principe de parcimonie, ou bien en appliquant des méthodes de probabilité Bayésienne (Prud'homme *et al.*, 2006). Parfois, lorsque l'arbre phylogénétique est mal connu, ou lorsqu'un grand nombre de changements phénotypiques se produisent dans un ensemble d'espèces proches, il n'est pas possible d'établir avec certitude la polarité du changement évolutif (Kopp *et al.*, 2000; Gompel & Carroll, 2003; Jeong *et al.*, 2006, 2008).

Ensuite, il faut s'assurer que l'évolution répétée au niveau morphologique ne s'explique pas par un transfert latéral de gènes; par exemple, le séquençage de trois lignées de papillons arborant un motif de pigmentation rouge similaire a révélé que le transfert de

fragments de chromosome par hybridation est partiellement responsable de ce mimétisme (Consortium, 2012).

Enfin, il faut distinguer les cas où l'évolution répétée s'appuie sur la sélection d'un allèle déjà présent dans la situation ancestrale, à faible fréquence; par exemple, différentes preuves indiquent que, de façon répétée, un même groupe d'allèles du gène EDA, responsable d'une réduction de plaques osseuses, s'est fixé chez différentes populations d'épinoches d'eau douce, par suite de la colonisation répétée à partir d'une même espèce marine parente (Colosimo *et al.*, 2005).

Dans ces deux derniers cas, ce sont les mêmes allèles, apparus seulement une fois au cours de l'évolution, qui se répandent et causent l'évolution de phénotypes similaires dans des lignées évolutives distinctes.

Premières méta-analyses

En excluant ces deux derniers cas, Martin & Orgogozo (2013) ont répertorié plus de 100 cas où des mutations dans les mêmes gènes ont contribué à l'évolution répétée d'un caractère similaire entre deux lignées évolutives indépendantes. En utilisant un jeu de données beaucoup plus partiel, Conte et Arnegard ont estimé la probabilité de réutilisation d'un même gène lorsque l'évolution se répète entre 0,32 et 0,55. Malgré des biais expérimentaux et méthodologiques très nombreux (Rockman, 2011; Stern, 2013), ces méta-analyses suggèrent que de façon générale, l'évolution, lorsqu'elle emprunte des trajectoires phénotypiques similaires, passe par des chemins génétiques similaires plus souvent que si tous les trajets génétiques étaient équiprobables.

Par ailleurs, ces études révèlent que le degré de similarité génétique de ces chemins parallèles s'étale sur un large spectre : à une extrémité de ce spectre, les cas d'évolution répétée qui reposent sur des changements identiques au niveau protéique, comme souvent observé pour les cas de résistance aux insecticides ou aux fongicides; à l'autre extrémité, les cas d'évolution répétée qui reposent sur des changements affectant les mêmes cascades de signalisation ou modules de développement (Martin & Orgogozo, 2013).

Pour apprécier ce point d'importance, détaillons un cas plutôt bien résolu au niveau génétique, chez trois lignées de drosophiles ayant gagné de façon indépendante un motif de tache pigmentaire sur l'aile.

Similitude à plusieurs étages chez des drosophiles à pois noirs

Dans les groupes *melanogaster*, *obscura* et *guttifera*, une tache noire de pigmentation est apparue sur l'aile,

Fig. 2. L'évolution répétée de la pigmentation dans trois groupes de drosophile. Chez *D. biarmipes* (A), *D. tristis* (B) et *D. guttifer* (C) le motif de pigmentation sur l'aile de l'adulte est préfiguré pendant le développement par le motif d'expression de Yellow. Chez *D. biarmipes* et *D. guttifer*, Ebony est exprimé dans un patron d'expression réciproque à Yellow. Chez *D. biarmipes*, *D. tristis* et *D. guttifer*, des changements en cis de *yellow* sont responsables de l'évolution de nouveaux enhancers (D) qui contrôlent l'expression du gène dans l'aile en motif précis (A)b : rapporteur *spot^{bia675}* transformé chez *D. biarmipes* (Arnoult et al., 2013) ; (B)t : rapporteur *intron-spot* transformé chez *D. melanogaster* (Prud'homme et al., 2006) ; (C) *g*¹ : rapporteur *vein spot* transformé chez *D. guttifer* (Werner et al., 2010) ; (C) *g*² : rapporteur *intervein shade* transformé chez *D. guttifer* (Werner et al., 2010). Chez *D. biarmipes* et *D. guttifer*, un régulateur principal (respectivement Dll et Wg) coordonne les patrons d'expression d'effecteurs enzymatiques qui produisent la pigmentation.

de façon indépendante, comme le montre la figure 2. Chez *D. biarmipes* (groupe *melanogaster*), *D. tristis* (groupe *obscura*) et *D. guttifer* (groupe *guttifer*) l'expression de Yellow dans l'aile pupale préfigure le motif de pigmentation adulte (Gompel et al., 2005; Prud'homme et al., 2006; Werner et al., 2010) comme le montrent les immuno-marquages de la figure 2. Chez toutes les drosophiles, cette enzyme joue un rôle clé, mal connu (Drapeau, 2003), dans le dépôt de pigmentation pendant le développement (Walter et al., 1991; Riedel et al., 2011). Par ailleurs, chez *D. biarmipes* et *D. guttifer*, le patron d'expression d'Ebony chez une pupa tardive est approximativement le négatif de la pigmentation adulte (Gompel et al., 2005), comme le montre la figure 2. Ebony fait partie du même réseau métabolique que Yellow et agit comme un répresseur de la pigmentation mélanique sombre (Wittkopp et al., 2002).

Chez *Drosophila biarmipes*, *Drosophila tristis* et *Drosophila guttifer* (Gompel et al., 2005; Prud'homme et al., 2006; Werner et al., 2010) des changements non codants, en cis de *yellow*, ont modifié la régulation spatiale du gène et sont en grande partie responsables du changement du motif d'expression de la protéine.

Pour *Drosophila biarmipes* et *Drosophila tristis*, des mutations à proximité d'un *enhancer* déjà fonctionnel, mais dans deux régions régulatrices distinctes, ont provoqué une modification de la régulation spatiale de *yellow*, comme le montrent les constructions de rapporteurs présentées figure 2 (A)b et (B)t. Chez *Drosophila guttifer*, le motif d'expression complexe de *yellow* est généré par deux *enhancers*, comme le montre la figure 2 (C)g¹ et (C)g².

Enfin, chez *Drosophila biarmipes* et *Drosophila guttifer*, le patron de la pigmentation et de

Fig. 3. Émergence et diversification d'une tache de pigmentation par deux bricolages génétiques distincts. Adapté de Arnoult *et al.*, 2013. L'émergence d'une tache de la pigmentation – transition d'un état ancestral, décrit panneau (A), vers un état dérivé décrit panneau (B) – a reposé sur l'assemblage d'un nouveau module de régulation génétique : différents effecteurs enzymatiques, comme yellow ou ebony, tombant sous la régulation de Distal-less. La diversification du motif de pigmentation entre espèces voisines (C) a par la suite résulté d'une redistribution spatiale de Distal-less. La couleur bleue indique l'évolution de nouveaux liens régulateurs sur le panneau (B) et l'évolution de la distribution spatiale de Distal-less sur le panneau (C).

l'expression de *yellow* est contrôlé par un régulateur positif principal. Chez *Drosophila guttifera*, le morphogène Wingless régule positivement un *enhancer* de *yellow* et contrôle plus généralement le motif de la pigmentation dans l'aile. Le facteur de transcription qui médie la signalisation Wingless au niveau de l'*enhancer* de *yellow* n'a pas été identifié. Chez *D. biarmipes*, le facteur de transcription Distal-less régule directement *yellow* de façon positive, au niveau de son *enhancer* 5'; Distal-less régule également négativement *ebony* et de façon plus générale, contrôle le développement du motif de pigmentation dans l'aile.

Au vu de ces résultats, l'évolution semble contrainte à différents niveaux génétiques lorsqu'elle conduit à l'émergence d'un motif pigmentaire chez les drosophiles. Sur les trois trajets génétiques indépendants, les trois impliquent un changement

d'expression de Yellow, à chaque fois médié par des mutations *cis*-régulatrices, constituant un ou deux nouveaux *enhancers*, deux chemins impliquent la régulation en motif inversé d'Ebony, et à deux reprises un régulateur positif de Yellow et de la pigmentation semble contrôler le motif spatial du caractère.

Toutes ces similarités, à tous ces niveaux de description génétique, ont-elles les mêmes causes ?

Les causes multiples de l'évolution répétée

Les premières tentatives de méta-analyse de la littérature (Arendt & Reznick, 2008; Stern & Orgogozo, 2008; Nadeau & Jiggins, 2010; Martin & Orgogozo, 2013) indiquent que si les mutations sont bien générées au hasard, celles que l'évolution retient finalement sont distribuées de façon non

Boîte 1 : Différentes hypothèses expliquant l'existence des « points chauds » génétiques

Pourquoi certaines mutations s'accumulent-elles dans certaines régions génétiques ? Les nombreuses hypothèses de travail émises dans la littérature (Kopp, 2009 ; Stern & Orgogozo, 2009 ; Martin & Orgogozo, 2013) s'appuient souvent sur un redécoupage de cette question. Il s'agit de comprendre d'une part pourquoi, individuellement, une mutation se fixe ou non dans une population. Et d'autre part, il s'agit de comprendre pourquoi, individuellement ou collectivement, une ou plusieurs mutations se fixe(nt) autour d'un même gène. À ces deux questions, on peut trouver dans la littérature une multitude de réponses théoriques, brièvement exposées ici.

Pourquoi individuellement, une mutation se fixe-t-elle ou non dans une population ?

Le destin d'une mutation dépend de la probabilité d'apparition de cette mutation (ce paramètre est lui-même modulé par différents facteurs, comme le taux de mutation ou la taille effective de la population), de l'effet de la mutation sur la *fitness* (paramètre qui dépend souvent de sa pléiotropie, de l'état allélique à d'autres locus – épistasie ou d'un facteur environnemental – plasticité).

Ce destin est également fonction du régime de sélection (par exemple, un régime de sélection très efficace, comme celui opéré artificiellement par des éleveurs, favorise les mutations de fortes amplitudes phénotypiques), de la variation dans le temps et dans l'espace de ce régime de sélection, ainsi que de la structure démographique de la population (dans une population de petite taille par exemple, la dérive génétique autorise fréquemment la fixation d'une mutation, indépendamment de son effet sur la *fitness*). La taille de la population, mais également sa structuration en dèmes (populations), l'existence et l'intensité de la migration sont autant de facteurs démographiques qui jouent sur la probabilité de fixation d'une mutation.

Pourquoi, individuellement ou collectivement, une ou plusieurs mutations se fixe(nt)-elle(s) autour d'un même gène ?

La probabilité qu'une ou plusieurs mutations s'accumulent de préférence autour d'un même gène au cours de l'évolution est dépendante de plusieurs facteurs. Certains gènes présentent à un instant *T* de l'évolution une plus grande variabilité naturelle (« *standing genetic variation* »), c'est-à-dire que de nombreux allèles coexistent dans une population ; plus cette variabilité est grande, plus grande est la chance que l'évolution filtre un de ces allèles. La taille de la cible susceptible d'être mutée (« *mutational target size* ») est un autre facteur clé : il s'agit du nombre théorique de mutations bénéfiques possibles par gène. Les gènes de grande taille ou bien présents en grand nombre de copies dans le génome, ou ceux qui possèdent une relation structure-fonction particulière, sont une cible de taille plus importante pour l'évolution. Enfin, certains gènes sont situés dans des zones de grande instabilité chromosomique, ce qui augmente le taux de mutation local.

uniforme dans les génomes. Pourquoi les mutations s'accumulent-elles de façon plus probable à certains *locus*, parfois appelés « points chauds » de l'évolution ? Peut-on trouver des qualités communes à tous ces points chauds, et à toutes ces mutations ? On trouve dans la littérature – et sans doute dans la nature ! – une grande diversité de réponses possibles, brièvement évoquées dans la boîte 1.

Comment déterminer quelles sont, parmi ces multiples déterminants, les causes les plus à même d'expliquer l'existence de points chauds génétiques ?

Multiplier les axes d'analyse

La fin des oppositions monolithiques

Pendant longtemps une distinction polémique entre l'importance relative des mutations dans les régions régulatrices en *cis*, *versus* des mutations codantes en *trans* (Lynch & Wagner, 2008) a dominé la discussion

autour du substrat génétique de l'évolution (Hoekstra & Coyne, 2007 ; Carroll, 2008).

L'argument « classique » suit toujours la même ligne de raisonnement (Carroll, 2005 ; Wray, 2007) : soit un gène donné (un facteur de transcription par exemple), qui possède plusieurs rôles différents dans le développement d'un organisme, autrement dit, à forte pléiotropie mosaïque ; une mutation ponctuelle dans sa séquence codante sera très rarement bénéfique dans tous les contextes où ce gène est impliqué durant le développement. En revanche, une mutation ponctuelle dans un de ses éléments *cis*-régulateurs affectera son expression de manière discrète, spécifiquement dans le patron d'expression spatio-temporel contrôlé par cet élément. Il en résulterait que les mutations dans ces régions *cis*-régulatrices sont généralement moins pléiotropes.

Le problème de cet argument est, en plus d'être très qualitatif et d'ignorer la modularité des domaines protéiques (Lynch & Wagner, 2008), d'être trop général, et d'ignorer tous les autres facteurs

biologiques qui peuvent influencer le destin évolutif des mutations. Une étude récente illustre les limites de cette opposition : chez plusieurs insectes herbivores, l'évolution répétée de résistance à une toxine végétale a résulté alternativement ou conjointement de mutations codantes, de régulations et de duplications ; toutes ces modifications ont eu lieu à un même locus, la cible moléculaire de la toxine (Zhen *et al.*, 2012).

Récemment, des axes d'analyses nouveaux ont été proposés, permettant de dépasser cette opposition trop simple et de fragmenter la question du substrat génétique de l'évolution, en fonction d'un aspect particulier du contexte biologique.

Dans les sections suivantes, je résume trois propositions parmi ces axes d'analyse émergents.

Répertoire spécialisé vs. pléiotropie mosaïque

Parmi les cas extrêmes d'évolution répétée qui se reflète au niveau protéique, celui des changements de vision chez différents vertébrés est exemplaire. Les opsines forment une grande famille multigénique ; sous forme protéique elles participent spécifiquement à la formation des molécules photosensibles qui donnent à nos yeux la capacité de percevoir la lumière. Leur propriété d'absorption détermine la gamme de longueurs d'ondes à l'intérieur de laquelle nous pouvons voir. De façon répétée et indépendante, des mutations codantes dans la séquence de certaines opsines, ou des duplications suivies de pareilles mutations, ont altéré leur spectre d'absorption et conduit à un changement correspondant de la vision (Yokoyama, 2002).

Il existe de très nombreux exemples de changements évolutifs associés à des duplications ou des mutations codantes de gènes effecteurs spécialisés, comme les opsines ; c'est le cas dans la respiration, le métabolisme digestif, et la défense immunitaire (Carroll, 2005). De façon générale, lorsqu'un ou plusieurs gènes portent une fonction moléculaire *spécifiquement* dévolue à la fabrication d'un aspect du phénotype, il semble que l'évolution passe dans ces cas par un raffinement de ce répertoire moléculaire spécialisé : soit par une expansion, soit par un léger changement fonctionnel.

À l'inverse, lorsque le développement d'un caractère – par exemple un caractère morphologique chez un organisme pluricellulaire – repose sur l'action de gènes impliqués dans un grand nombre d'autres processus développementaux (on parle de pléiotropie mosaïque du gène), il semble que les changements évolutifs conduisent souvent plutôt à une réutilisation du répertoire génétique, notamment *via* des changements dans les systèmes de régulation génétique. Par exemple, la méta-analyse

de Martin & Orgogozo (2013) révèle que les facteurs de transcription sont environ quatre fois plus représentés dans leur catalogue que leur proportion dans les génomes.

Court-terme vs. long terme

Parmi les axes d'analyse récemment mis en avant, se trouve l'axe du temps. Les mutations pertinentes au regard de l'évolution sont-elles de même nature lorsqu'un caractère varie à l'intérieur d'une espèce ou entre espèces ?

De façon remarquable, Stern & Orgogozo (2008) ont recensé une proportion inversée de mutations nulles responsables de changement aux échelles de temps micro-évolutives *vs.* macro-évolutives : ces dernières représentent plus de 50 % des mutations responsables de changements entre races domestiquées et seulement 5 % des mutations responsables de changements entre espèces. Ce contraste pourrait être expliqué par la force de la sélection artificielle, ou par la petite taille des populations sélectionnées.

Concernant les changements de morphologie, les mutations *cis*-régulatrices sont toujours plus représentées pour les changements évolutifs sur le long terme. D'après les chiffres de la méta-analyse la plus récente (Martin & Orgogozo, 2013), elles ne représentent que 22 % (44 sur 195) des mutations causant un changement de morphologie chez des races domestiquées, 26 % (40 sur 151) des mutations causant un changement de morphologie entre populations d'une même espèce *vs.* 83 % (33 sur 36) des mutations causant un changement entre espèces.

En utilisant un jeu de données beaucoup plus restreint, Conte & Arnegard (2012) calculent que la probabilité que des mutations frappent indépendamment les mêmes gènes, dans deux lignées indépendantes, décroît en fonction du temps passé depuis la divergence de ces deux lignées. Ce résultat, s'il est confirmé par d'autres méta-analyses, suggère que si le développement limite effectivement le jeu des mutations possibles, ce jeu de contraintes diverge avec le temps dans chaque branche du vivant.

Émergence vs. diversification

En observant la nature, il nous semble parfois que l'évolution se montre réellement plus inventive : plutôt que d'optimiser un caractère existant, elle en crée un radicalement nouveau, par exemple la carapace des tortues, ou bien pour reprendre un exemple développé plus haut, la tache noire sur les ailes de *D. biarmipes*. La nouveauté évolutive, entendue comme une propriété qualitativement différente qui ouvre ou crée une niche écologique, s'accompagne souvent d'une diversification du trait chez des espèces cousines

Fig. 4. Des changements génétiques déterminés dans une poignée d'espèces. Le nombre de mutations responsables d'un changement évolutif, recensées par Martin et Orgogozo (2013), est indiqué sur une échelle logarithmique chez différentes espèces étudiées. Leur contribution relative est également précisée en pourcentage.

(Pigliucci, 2008). Par exemple, chez les espèces cousines de *D. biarmipes*, la tache est plus ou moins claire ou foncée, brune ou noire, ramassée ou émietée, allongée ou recourbée, voire absente (Arnoult *et al.*, 2013).

Nous avons déjà mentionné que dans ce groupe de drosophiles, l'émergence du caractère a résulté du recrutement par un même facteur de transcription, Distal-less, d'au moins deux cibles différentes, jouant des rôles enzymatiques réciproques dans la production de la pigmentation. Une fois ce module de régulation génétique assemblé, la figure 4 montre comment la diversification spatiale du motif de pigmentation s'est appuyée sur un changement d'expression de ce facteur de transcription (Arnoult *et al.*, 2013).

Cette étude suggère qu'au cours des radiations évolutives, deux versants de l'évolution, émergence et diversification, correspondent à deux bricolages génétiques distincts : le recrutement de différents effecteurs enzymatiques par régulateur commun pour l'émergence; la redistribution de ce chef d'orchestre moléculaire pour la diversification spatiale du caractère. Ce modèle pourrait s'appliquer à d'autres cas de radiations recensés dans la littérature, où un caractère nouvellement apparu se trouve redistribué dans différents motifs spatiaux, chez différentes espèces cousines (Brakefield *et al.*, 1996; Abzhanov *et al.*, 2004; Reed *et al.*, 2011).

Conclusions et perspectives

Les nombreuses limites empiriques

La recherche et la caractérisation des mutations responsables de changement évolutif s'appuient souvent sur une approche gène-candidat, comme c'est le cas montré dans la figure 2. Cette approche biaise donc les résultats vers l'ensemble limité de gènes dont la fonction durant le développement est connue.

Alternativement, des approches pan-génomiques sont utilisées : les mutations sont détectées par QTL (*Quantitative Trait Loci*) (Barton & Keightley, 2002), par association génétique (Mackay *et al.*, 2012), ou par crible génomique (Hohenlohe *et al.*, 2010). Soit pour des raisons techniques, soit à cause de biais statistiques, l'ensemble des approches pan-génomiques ne permet la détection que d'une dizaine de changements génétiques au maximum (Nadeau & Jiggins, 2010), et biaisent la recherche vers la caractérisation de mutations à forts effets phénotypiques (Rockman, 2011). Mettre en œuvre ces approches n'est par ailleurs possible que dans une poignée d'espèces, comme le montre la figure 4. Des mutations responsables de changements évolutifs n'ont été répertoriées de façon robuste que chez 146 espèces, sur les quelques 8 millions d'espèces estimées (Mora *et al.*, 2011); la figure 4 montre les espèces principales chez lesquelles de telles

mutations ont été déterminées : plus de la moitié des investigations ont été menées sur les 18 mêmes espèces. Le biais concernant les études chez les espèces domestiquées et les humains est frappant : plus de 40 % des études référencées concernent ces espèces.

Enfin, tester le caractère causal d'une ou de plusieurs mutations dans un changement évolutif, nécessite très souvent une manipulation génétique, par transgénèse ou introgression. Ces manipulations génétiques ne sont possibles que dans un nombre très limité d'espèces modèles.

Bon nombre de limites empiriques actuelles sont en passe d'être levées grâce aux nouvelles générations de séquençage, et à l'importation dans toutes sortes d'espèces non modèles d'outils génétiques puissants, comme les CRISPR (*Clustered Regularly Interspaced Short Palindromic Repeats*) (Hsu *et al.*, 2014). Ces nouvelles techniques permettent par exemple d'étudier directement des populations naturelles plutôt que des populations de laboratoire (Elmer & Meyer, 2011). Le séquençage toujours moins cher, toujours plus précis, permet de gagner en pouvoir de discrimination statistique à tous les niveaux d'analyse ; par exemple, il est plus facile de distinguer la contribution d'un allèle unique *vs.* plusieurs allèles liés à un changement évolutif, lorsque la couverture de séquençage est importante. Enfin, le re-séquençage massif de lignées de laboratoire, ayant subi un régime de sélection artificielle, permet d'étudier à l'échelle pan-génomique les ressorts génétiques de l'adaptation, comme dans l'exemple des bactéries *Cit*⁺.

Au-delà de ces limitations techniques, subsiste un biais conceptuel, hérité de notre perception actuelle du développement.

Une vision « GRN » centrée

La compréhension de l'évolution peut être éclairée par une connaissance approfondie du développement. Il en résulte qu'une connaissance partielle et biaisée du développement ne jette qu'une lumière biaisée sur les ressorts de l'évolution. Les approches récentes du développement ont conduit à une vision principalement centrée sur les réseaux de régulation génétique (Davidson & Erwin, 2006 ; Davidson & Levine, 2008) ; très souvent, le développement est réduit à l'établissement coordonné dans le temps et dans l'espace de l'expression de gènes sous l'action de réseaux de régulation génétique (ou *Gene Regulation Network*, GRN). Nulle surprise de voir que certaines avancées récentes établissent donc l'importance évolutive de mutations dans ces GRN. Il est fort à parier que des mutations dans des gènes impliqués à d'autres niveaux de régulation génétique (architecture chromatinienne,

taux de dégradation des ARN, épissage alternatif, modifications post-traductionnelles) seront identifiées dans le futur.

Conclusion : le défi de l'adaptation

Évoquée au tout début de l'introduction, l'adaptation a presque disparu de la suite de cette revue. En effet, hormis dans les systèmes moléculaires ou cellulaires présentés brièvement, il est très difficile de mesurer l'adaptation d'un individu à son environnement ; encore plus difficile d'atomiser l'individu en traits quasi-indépendants susceptibles d'être considérés comme des adaptations ; et encore plus périlleux de reconstituer l'histoire sélective d'un tel trait de son état ancestral à son état dérivé (pour une discussion détaillée de ces difficultés, se référer à Lewontin, 1978 ; Brandon, 1999).

Il en résulte que de nombreuses études centrées sur l'évolution d'un caractère particulier (Sucena *et al.*, 2003 ; Tanaka *et al.*, 2011), limpides d'un point de vue génétique, ne font absolument aucun lien entre le caractère, le changement génétique et une quelconque fonction adaptative.

Une telle tâche n'est néanmoins pas illusoire ; il est parfois possible d'isoler et de tester la contribution d'un allèle à la *fitness* actuelle (se référer à Barrett & Hoekstra (2011) pour une revue détaillée). Une fois un allèle candidat retenu (par une approche gène candidat ou pan-génomique), il est parfois possible de mettre en place des tests expérimentaux dans la nature et de calculer les coefficients de sélection, qui s'appliquent sur les différents génotypes.

Revenons à l'exemple des épinoches. Une cinquantaine d'individus marins rares, hétérozygotes au *locus* EDA - le *locus* responsable de la variation du nombre de plaques osseuses - ont été transférés dans quatre bassins d'eau douce indépendants (Barrett *et al.*, 2008) ; tout au long d'un cycle de reproduction, la fréquence de cet allèle a été suivie dans la population en effectuant un échantillonnage génétique régulier des quatre populations. De façon remarquable, cette expérience a confirmé que l'allèle fixé dans les populations d'eau douce est associé à un taux de croissance plus élevé, une meilleure survie, et une maturité sexuelle plus précoce, mais elle a également révélé un effet délétère de cet allèle (ou d'un allèle physiquement proche sur le chromosome) sur les stades juvéniles, avant le développement des plaques osseuses, sans doute sur un autre trait important d'un point de vue adaptatif.

Cet exemple souligne à la fois la difficulté de l'étude de l'adaptation au niveau génétique, et son intérêt pour lever une partie du voile sur les liens compliqués entre génotype, phénotype et *fitness*.

Références

- Abzhanov A., Protas M., Grant B., Grant P., Tabin C., Bmp4 and morphological variation of beaks in Darwin's finches. *Science*, 2004, 305, 1462–1465.
- Arendt J., Reznick D., Convergence and parallelism reconsidered: what have we learned about the genetics of adaptation? *Trends Ecol Evol*, 2008, 23, 26–32.
- Arnoult L., Su K.F.Y., Manoel D., Minervino C., Magrina J., Gompel N., Prud'homme B., Emergence and Diversification of Fly Pigmentation Through Evolution of a Gene Regulatory Module. *Science*, 2013, 339, 1423–1426.
- Barrett R., Hoekstra H., Molecular spandrels: tests of adaptation at the genetic level. *Nat Rev Genet*, 2011, 12, 767–780.
- Barrett R.D., Rogers S.M., Schluter D., Natural selection on a major armor gene in threespine stickleback. *Science*, 2008, 322, 255–257.
- Barton N., Keightley P., Understanding quantitative genetic variation. *Nat Rev Genet*, 2002, 3, 11–21.
- Blount Z., Borland C., Lenski R., Historical contingency and the evolution of a key innovation in an experimental population of *Escherichia coli*. *Proc Natl Acad Sci USA*, 2008, 105, 7899–7906.
- Blount Z.D., Barrick J.E., Davidson C.J., Lenski R.E., Genomic analysis of a key innovation in an experimental *Escherichia coli* population. *Nature*, 2012, 489, 513–518.
- Brakefield P., Gates J., Keys D., Kesbeke F., Wijngaarden P., Monteiro A., French V., Carroll S., Development, plasticity and evolution of butterfly eyespot patterns. *Nature*, 1996, 384, 236–242.
- Brandon R., The units of selection revisited: the modules of selection. *Biol Philos*, 1999, 14, 167–180.
- Breen M., Kemena C., Vlasov P., Notredame C., Epistasis as the primary factor in molecular evolution. *Nature*, 2012, 490, 535–538.
- Bridgham J.T., Ortlund E.A., Thornton J.W., An epistatic ratchet constrains the direction of glucocorticoid receptor evolution. *Nature*, 2009, 461, 515–519.
- Carroll S., Evolution at two levels: on genes and form. *PLoS Biol*, 2005, 3, e245.
- Carroll S.B., Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell*, 2008, 134, 25–36.
- Colosimo P.F., Hosemann K.E., Balabhadra S., Villarreal G., Dickson M., Grimwood J., Schmutz J., Myers R.M., Schluter D., Kingsley D.M., Widespread parallel evolution in sticklebacks by repeated fixation of Ectodysplasin alleles. *Science*, 2005, 307, 1928–1933.
- Consortium T.H.G., Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature*, 2012, 487, 94–98.
- Conte G., Arnegard M., The probability of genetic parallelism and convergence in natural populations. *Proc Biol Sci*, 2012, 279, 5039–5047.
- Davidson E.H., Erwin D.H., Gene regulatory networks and the evolution of animal body plans. *Science*, 2006, 311, 796–800.
- Davidson E.H., Levine M.S., Properties of developmental gene regulatory networks. *Proc Natl Acad Sci USA*, 2008, 105, 20063–20066.
- Dean A., Thornton J., Mechanistic approaches to the study of evolution: the functional synthesis. *Nat Rev Genet*, 2007, 8, 675–688.
- Drapeau M., A novel hypothesis on the biochemical role of the *Drosophila* Yellow protein. *Biochem Biophys Res Com*, 2003, 311, 1–3.
- Elmer K., Meyer A., Adaptation in the age of ecological genomics: insights from parallelism and convergence. *Trends Ecol Evol*, 2011, 26, 298–306.
- Gompel N., Carroll S., Genetic mechanisms and constraints governing the evolution of correlated traits in drosophilid flies. *Nature*, 2003, 424, 931–935.
- Gompel N., Prud'homme B., Wittkopp P.J., Kassner V.A., Carroll S.B., Chance caught on the wing: cis-regulatory evolution and the origin of pigment patterns in *Drosophila*. *Nature*, 2005, 433, 481–487.
- Gompel N., Prud'homme B., The causes of repeated genetic evolution. *Dev Biol*, 2009, 332, 36–47.
- Gould S., Lewontin R., The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme. *Proc R Soc Lond B Bio Sci*, 1979, 205, 581–598.
- Hendrickson H., Rainey P., Evolution: How the unicorn got its horn. *Nature*, 2012, 489, 504–505.
- Hoekstra H.E., Coyne J.A., The locus of evolution : Evo Devo and the genetic adaptation. *Evolution*, 2007, 61, 995–1016.
- Hohenlohe P., Bassham S., Etter P., Stiffler N., Johnson E., Cresko W., Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genet*, 2010, 6, e1000862.
- Hsu P.D., Lander E.S., Zhang F., Development and applications of CRISPR-Cas9 for genome engineering. *Cell*, 2014, 157, 1262–1278.
- Jacob F., Evolution and tinkering. *Science*, 1977, 196, 1161–1166.
- Jacob F., Le Jeu des possibles: Essai sur la diversité du vivant. 1981, Fayard, Paris.
- Jeong S., Rokas A., Carroll S.B., Regulation of body pigmentation by the Abdominal-B Hox protein and its gain and loss in *Drosophila* evolution. *Cell*, 2006, 125, 1387–1399.
- Jeong S., Rebeiz M., Andolfatto P., Werner T., True J., Carroll S.B., The evolution of gene regulation underlies a morphological difference between two *Drosophila* sister species. *Cell*, 2008, 132, 783–793.
- Kopp A., Duncan I., Godt D., Carroll S.B., Genetic control and evolution of sexually dimorphic characters in *Drosophila*. *Nature*, 2000, 408, 553–559.
- Kopp A., Metamodels and phylogenetic replication: a systematic approach to the evolution of developmental pathways. *Evolution*, 2009, 63, 2771–2789.
- Laubichler M., Maienschein J., From embryology to evo-devo: a history of developmental evolution. 2007, MIT Press, Cambridge.
- Lewontin R.C., Adaptation. *Sci Am*, 1978, 239, 212–218, 220, 222 passim.

- Liu Y, Cotton J.A., Shen B., Han X., Rossiter S.J., Zhang S., Convergent sequence evolution between echolocating bats and dolphins. *Curr Biol*, 2010, 20, R53–54.
- López-Fanjul C., García-Dorado A., The fuel of evolution. *Heredity*, 2011, 106, 535–536.
- Lynch V.J., Wagner G.P., Resurrecting the role of transcription factor change in developmental evolution. *Evolution*, 2008, 62, 2131–2154.
- Mackay T., Richards S., Stone E., Barbadilla A., Ayroles J., Zhu D., Casillas S., Han Y., Magwire M., Cridland J., The *Drosophila melanogaster* genetic reference panel. *Nature*, 2012, 482, 173–178.
- Martin A., Orgogozo V., The loci of repeated evolution: A catalog of genetic hotspots of phenotypic variation. *Evolution*, 2013, 67, 1235–1250.
- Mora C., Tittensor D., Adl S., Simpson A., Worm B., How many species are there on Earth and in the ocean? *PLoS Biol*, 2011, 9, e1001127.
- Nadeau N.J., Jiggins C.D., A golden age for evolutionary genetics? Genomic studies of adaptation in natural populations. *Trends Genet*, 2010, 26, 484–492.
- Orr H., The genetic theory of adaptation: a brief history. *Nat Rev Genet*, 2005a, 6, 119–127.
- Orr H., Theories of adaptation: what they do and don't say. *Genetica*, 2005b, 123, 3–13.
- Pigliucci M., What, if anything, is an evolutionary novelty? *Philosophy of Science*, 2008, 75, 887–898.
- Poelwijk F., Kiviet D., Weinreich D., Tans S., Empirical fitness landscapes reveal accessible evolutionary paths. *Nature*, 2007, 445, 383.
- Prud'homme B., Gompel N., Rokas A., Kassner V.A., Williams T.M., Yeh S.-D., True J.R., Carroll S.B., Repeated morphological evolution through cis-regulatory changes in a pleiotropic gene. *Nature*, 2006, 440, 1050–1053.
- Reed R., Papa R., Martin A., Hines H., Counterman B., Pardo-Diaz C., Jiggins C., Chamberlain N., Kronforst M., Chen R., Optix drives the repeated convergent evolution of butterfly wing pattern mimicry. *Science*, 2011, 333, 1137–1141.
- Riedel F., Vorkel D., Eaton S., Megalin-dependent Yellow endocytosis restricts melanization in the *Drosophila* cuticle. *Development*, 2011, 138, 149–158.
- Rockman M.V., The QTN program and the alleles that matter for evolution: all that's gold does not glitter. *Evolution*, 2011, 66, 1–17.
- Rokyta D., Joyce P., Caudle S., Wichman H., An empirical test of the mutational landscape model of adaptation using a single-stranded DNA virus. *Nat Genet*, 2005, 37, 441–444.
- Smith J., Burian R., Kauffman S., Alberch P., Campbell J., Goodwin B., Lande R., Raup D., Wolpert L., Developmental constraints and evolution: a perspective from the Mountain Lake conference on development and evolution. *Q Rev Biol*, 1985, 60, 265–287.
- Stern D., The genetic causes of convergent evolution. *Nat Rev Genet*, 2013, 14, 751–764.
- Stern D.L., Orgogozo V., The loci of evolution : How predictable is genetic evolution? *Evolution*, 2008, 62, 2155–2177.
- Stern D.L., Orgogozo V., Is genetic evolution predictable? *Science*, 2009, 323, 746–751.
- Sucena E., Delon I., Jones I., Payre F., Stern D.L., Regulatory evolution of shavenbaby/ovo underlies multiple cases of morphological parallelism. *Nature*, 2003, 424, 935–938.
- Tanaka K., Barmina O., Sanders L., Arbeitman M., Kopp A., Evolution of sex-specific traits through changes in HOX-dependent doublesex expression. *PLoS Biol*, 2011, 9, e1001131.
- Walter M.F., Black B.C., Afshar G., Kermabon A.Y., Wright T.R., Biessmann H., Temporal and spatial expression of the yellow gene in correlation with cuticle formation and dopa decarboxylase activity in *Drosophila* development. *Dev Biol*, 1991, 147, 32–45.
- Weinreich D., Delaney N., DePristo M., Hartl D., Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science*, 2006, 312, 111–114.
- Werner T., Koshikawa S., Williams T.M., Carroll S.B., Generation of a novel wing colour pattern by the Wingless morphogen. *Nature*, 2010, 464, 1143–1148.
- Wittkopp P.J., True J.R., Carroll S.B., Reciprocal functions of the *Drosophila* yellow and ebony proteins in the development and evolution of pigment patterns. *Development*, 2002, 129, 1849–1858.
- Wray G., The evolutionary significance of cis-regulatory mutations. *Nat Rev Genet*, 2007, 8, 206–216.
- Yokoyama S., Molecular evolution of color vision in vertebrates. *Gene*, 2002, 300, 69–78.
- Zhen Y., Aardema M., Medina E., Schumer M., Parallel Molecular Evolution in an Herbivore Community. *Science*, 2012, 337, 1634–1637.