Mechanistic approaches to the study of evolution:

the functional synthesis

Antony M. Dean[*] and Joseph W. Thornton[‡]

[*]*140 Gortner Laboratories, University of Minnesota, St Paul, Minnesota 55108, USA*

[‡]*5289 University of Oregon, Eugene, Oregon 97403, USA*

Abstract

An emerging synthesis of evolutionary biology and experimental molecular biology is providing much stronger and deeper inferences about the dynamics and mechanisms of evolution than were possible in the past. The new approach combines statistical analyses of gene sequences with manipulative molecular experiments to reveal how ancient mutations altered biochemical processes and produced novel phenotypes. This functional synthesis has set the stage for major advances in our understanding of fundamental questions in evolutionary biology. Here we describe this emerging approach, highlight important new insights that it has made possible, and suggest future directions for the field.

During the twentieth century, evolutionary and molecular biology diverged in subject matter and scientific culture. Few scientists were trained in both fields, and many biology departments were split up. This is a great pity, for the two modalities are complementary, not dichotomous. Today, a new wave of research is bridging this unnatural schism, bringing the techniques, subject matter and modes of inference of molecular biology to bear on classic and modern questions in evolution and, in the process, shedding new light on the origins and organization of molecular systems.

A great strength of molecular biology is the high standard of evidence-based inference. A causal link between a factor and its effect is established only when that factor has been isolated, with all other variables held constant in controlled experiments. To meet this standard, molecular biologists set aside much of the complexity of nature, including variability across individuals, populations and environments. But ignoring biological variation imposes a cost: although strong inferences can be drawn, the extent to which they can be generalized is sometimes unclear.

Most evolutionary biologists emphasize a less reductionist approach. Variation among individuals, populations and taxa is viewed not as sources of unwanted noise, but as the very phenomenon that requires explanation. Inferences about the historical mechanisms that generate variation are usually drawn from patterns of association that have been detected in surveys of sequences, allele frequencies and phenotypes. The strength of this approach is its focus on real biological systems in their natural and historical contexts; the weakness is that statistical associations are not reliable indicators of causality. Too often, alternative explanations are viable. Despite important experimental work in evolutionary biology (for example, REFS [1-9]), claims that are based solely on associations remain standard in the field.

Correspondence to A.M.D. or J.W.T. e-mails: deanx024@umn.edu;joet@oregon.edu.

As a result, inferences about historical evolution are seldom as decisive as those in molecular biology.

The functional synthesis

Today, the reductionist culture of molecular biology is being fused with the historical realism of evolutionary biology, creating a new functional synthesis. Research in the functional synthesis combines the techniques of evolutionary and phylogenetic analysis with those of molecular biology, biochemistry and structural biology. Phylogenetic, population genetic or quantitative genetic techniques are used to detect mutations associated with putatively adaptive phenotypes. These methods also provide historical insights into the process of evolution by identifying the period and lineage in which a function changed, the mutations that arose in that interval and the sequence sites that retain putative signatures of selection. Protein structures can help to identify historical amino-acid replacements that are likely to be involved in functional evolution.

Molecular biology provides experimental means to test these hypotheses decisively. Gene synthesis allows ancestral sequences, which can be inferred using phylogenetic methods, to be physically 'resurrected', expressed and functionally characterized[10]. Using directed mutagenesis, historical mutations of putative importance are introduced into extant or ancestral sequences. The effects of these mutations are then assessed, singly and in combination, using functional molecular assays. Crystallographic studies of engineered proteins — resurrected and/or mutagenized — allow determination of the the structural mechanisms by which amino-acid replacements produce functional shifts. Transgenic techniques permit the effect of specific mutations on whole-organism phenotypes to be studied experimentally. Finally, competition between genetically engineered organisms in defined environments allows the fitness effects of specific mutations to be assessed and hypotheses about the role of natural selection in molecular evolution to be decisively tested.

This dual strategy can be readily applied to the evolution of any gene or regulatory region with intrinsic functions — such as binding affinity, catalytic performance, transcriptional activity, signal transduction and so on — that can be studied in the laboratory. It can be used to reveal the mechanistic basis for the evolution of virtually any kind of phenotype that is amenable to reductionist analysis, including changes in morphology and development, physiology and behaviour[11-21]. In this sense, the functional synthesis subsumes and extends previous syntheses of evolution with specific biological subfields, such as the evo-devo synthesis[22] and sociobiology[23].

> **Evo-devo synthesis**
>
> The study of the origin and evolution of development, originally restricted to comparative methods, but increasingly using experimental approaches.

Molecular adaptation: the evolution of function

Four decades ago, Hubby and Lewontin introduced enzyme electrophoresis, launching an exciting new era in evolutionary biology. Perceiving what this technique would make possible, they articulated a newly focused research agenda for the field:

'[A] description of the genetic variation in a population is the fundamental datum of evolutionary studies; and it is necessary to explain the origin and maintenance of this variation and to predict its evolutionary consequences.'[24]

Scores of researchers began to catalogue genetic variation within and among populations in the form of allozyme polymorphisms. They attempted to infer the evolutionary processes that produced the observed patterns of variation, but the lack of resolution that is intrinsic to allozyme data severely hindered their progress. Lewontin's paradigm was reinvigorated in the 1980s with the introduction of DNA sequencing[25]. There followed a second flowering, this time of nucleotide polymorphisms, coalescent theory and more powerful statistical analyses of selection, drift and other forces.

Coalescent theory

A mathematical framework, based on the genealogy of alleles, for estimating population genetic alleles.

Box 1

Some long-standing questions

Studies in the functional synthesis are positioned to address key questions that have remained recalcitrant for decades, such as:

- How many mutations are required to produce a new function?

- Does functional or phenotypic evolution proceed by jerks (a few crucial mutations of large effect) or by creeps (many mutations, each of small effect)?

- What role does epistasis have in structuring evolutionary trajectories?

- Are there many evolutionary paths to the same final sequence?

- Could alternative solutions to the same problem have evolved and, if so, how might they differ in sequence, structure and mechanism?

- How does complexity arise in living systems?

Comparative analysis of sequence variation proved enormously productive: it yielded many important insights and continues to dominate studies of molecular evolution today. However, the statistical paradigm that Lewontin advocated is limited in the kinds of knowledge it can generate and the strength of the inferences it provides. By incorporating molecular analyses of gene action, the functional synthesis extends the old statistical approach in three ways to provide deeper, more decisive insights into evolutionary processes.

First, experimental studies in the functional synthesis provide strong independent corroboration of statistical analyses that, in isolation, would remain speculative[26]. For example, statistical 'signatures' of positive selection[27-30] can also be forged by chance, fluctuating population sizes, selection at linked or synonymous sites and other factors that violate test assumptions[31,32]. In the functional synthesis, the phenotypic impact of each putatively important mutation is characterized empirically by introducing it into a reconstructed ancestral gene and evaluating its effects on function and fitness. If the 'evolved' phenotype is reproduced in whole or in part — and if fitness is increased — then the statistical hypothesis is affirmed. Experimental corroboration makes inferences in the functional synthesis far stronger than those made in the old paradigm.

Second, the functional synthesis explicitly connects genotype with phenotype to allow deeper insights into the causes of evolutionary change. The old paradigm focused exclusively on markers of genetic polymorphism and divergence; ignoring phenotype, it could offer no explanation of how genetic differences alter function and fitness. By revealing the mechanisms

by which specific mutations produce new phenotypes, the functional synthesis can determine the causes and biological mechanisms of adaptation.

Third, the functional synthesis can resolve fundamental questions about evolutionary processes that have remained unresolved for decades (BOX 1). Answers to these questions depend on the 'maps' that relate changes in gene sequence to changes in phenotype and fitness. In the functional synthesis, manipulative experiments are used to characterize these maps directly. By empirically determining how mutations — singly or in combination, historical or unnatural — change function and fitness, studies in the functional synthesis analyse evolutionary processes at multiple levels. By ignoring phenotype, the old paradigm was mute on such topics.

Early recognition of the old paradigm's limitations led a few researchers to investigate the phenotypic and functional aspects of allozyme adaptation[4,33-35]. Technical limitations hampered their progress, but these have now been overcome. In the following sections, we discuss several exemplary studies in the functional synthesis that combine evolutionary analyses of history with manipulative analyses of gene function. Space limitations prevent us from covering many kindred studies in protein evolution (REFS 35[-47], to name just a few) and the evolution of development and behaviour (for example, REFS [14-21,48-56]). Instead, we focus on a few projects that illustrate how this new approach deepens our knowledge of evolution, moving us towards a decisive understanding of the molecular mechanisms of adaptation.

Insecticide resistance

In an early exemplar of the functional synthesis, John Oakeshott and colleagues studied the evolution of resistance to diazinon in the sheep blowfly *Lucilia cuprina*[57]. Diazinon is an organophosphate insecticide that is commonly used in Australasia. Like nerve gases, diazinon binds to and inactivates esterases, including acetylcholinesterase, an enzyme that is crucial to neurotransmission (FIGS 1a,b).

Resistance to diazinon had been statistically associated with an allele at another esterase locus, *E3* (REFS 58[-60]). To assess whether this allele confers the ability to detoxify the insecticide and to determine the mechanism for this shift, Newcomb *et al*. compared *E3* sequences from several natural populations and found that alleles from resistant and susceptible flies differed by five amino acids[57]. They then constructed chimeric enzymes from the resistant and sensitive alleles, expressed them in cultured cells and tested their enzymatic activities *in vitro*. Only one of the five replacements — a Gly135Asp switch in the active site of the enzyme — conferred on E3 the novel capacity to hydrolyse organophosphates at rates as high as resistant alleles from the wild. The reverse Asp135Gly replacement, when introduced into a naturally resistant allele, produces the susceptible phenotype that cannot hydrolyse organophosphates[57] (FIG. 1c).

To better understand the mechanism for this functional shift, Newcomb *et al*. then modelled the E3 active site on the crystallographic structure of a closely related esterase[57] (FIG. 1d). In the susceptible allele, as in acetylcholinesterase, diazinon binds to the enzyme but undergoes only the first half-reaction, producing an enzyme-inhibitor complex that resists further hydrolysis. This effectively blocks the natural esterase activities of both E3 and acetylcholinesterase. However, in the Gly135Asp mutant of E3, the side-chain carboxylate of the aspartate activates a water molecule in the active site, leading to hydrolysis of the intermediate complex. The functional E3 enzyme is regenerated, and the pesticide is hydrolysed and rendered non-toxic.

Extending their study to other populations[61], Oakeshott's group identified a second set of active-site mutations in E3, Trp251Leu and Trp251Thr, both of which confer resistance to

another organophosphate, malathion. A survey of DNA from museum specimens showed that these mutants were present in some blowflies that were collected before organophosphates were first used in Australia in 1955. By contrast, the Gly135Asp mutant that confers diazinon resistance was not found in a single specimen. These data suggest that malathion resistance in *L. cuprina* evolved from pre-existing genetic variation, whereas diazinon resistance required a novel mutation. Why the difference? Functional assays reveal that the mutations at site 251 confer malathion hydrolase activity with little or no effect on esterase activity. By contrast, the Gly135Asp mutant gains diazinon hydrolase activity but loses normal esterase activity, causing developmental instability and asymmetry in homozygous flies. In the absence of pesticides, one would expect purifying selection to keep the Gly135Asp mutant at low frequency, whereas the Trp251Leu mutant could segregate neutrally. Intriguingly, both the Trp251Leu and Gly135Asp replacements arose independently in a sister species, *Lucilia sericata* — in regions where they have similar geographic distributions with respect to pesticide use — and also in the distantly related housefly, *Musca domestica*[61,62].

Coat colour in mice

Newcomb *et al*.'s experiments were among the first to supplement sequence comparisons with functional molecular analyses to definitively identify the mutations responsible for an adaptive phenotype. More recently, Hopi Hoekstra and colleagues used a similar approach to identify a single amino-acid replacement that confers an adaptive change in the coat colour of the beach mouse *Peromyscus polionotus*[63].

In response to predation, mice living in beach environments evolved lighter coats than those living in dark rocky environments. Hoekstra *et al*. focused on a candidate gene — the melanocortin-1-receptor gene (*Mc1r*) — that was known to have a crucial role in the vertebrate signal-transduction pathway that controls pigment deposition. Sequencing *Mc1r* alleles from light-coloured and dark-coloured mouse populations revealed a single fixed difference in the peptide sequence that was perfectly associated with coat colour.

To test the functional effect of this substitution, Hoekstra *et al*. expressed both alleles in cultured cells, and examined their ability to function in the signalling cascade. The receptor from the light-coloured beach mice had a much lower affinity for the melanocortin ligand, and had lost the ability to stimulate cyclic AMP (cAMP) production, the second messenger that triggers the pigmentation-related cascade. Crosses between light and dark mice confirmed that *Mc1r* alleles contribute to coat colour, but showed that they explain only 9-28% of the variation in pigmentation that is observed in wild populations. Evidently, the complete adaptive phenotype involves other loci.

Hoekstra's study provides strong evidence that the *Mc1r* amino-acid replacement contributes to light coat colour by interfering with the cellular response to the melanocortin signal. The story, however, is not yet as complete as that of the blowfly esterase, largely because the genetic basis of coat colour is more complex than that of insecticide resistance. With time, the structures of *Mc1r* and its paralogues will be solved to determine how the replacement compromises receptor function, the other loci that contribute to coat colour will be identified, and transgenic experiments will corroborate the effect of these mutations on the coat phenotype. Indeed, subsequent mapping and gene expression studies have already shown that changes in the agouti gene, which encodes an MC1R antagonist, are statistically associated with light pigmentation in these populations[64]. Even at this relatively early stage, Hoekstra's work shows how functional molecular biology can be used to identify definitively the specific mutations that contribute to a quantitative genetic trait of known adaptive significance.

The evolution of colour vision

The blowfly and beach mouse stories describe recent adaptations for which genetic variation still segregates in natural populations. Many ancient adaptations are fixed in present-day taxa, and are therefore not amenable to population and quantitative genetic analyses. In such cases, phylogenetic techniques can be combined with molecular assays to explore how historical mutations have produced new gene functions and phenotypes.

Over the past two decades, Shozo Yokoyama and co-workers have studied how vertebrate visual pigments evolved sensitivity to different wavelengths of light. A visual pigment consists of a G-protein-coupled transmembrane protein, called an opsin, complexed with a chromophore derived from vitamin A. When a visual pigment absorbs a photon, the chromophore isomerizes, inducing a conformational change in the opsin. This in turn triggers a phototransduction cascade that produces a signal via the optic nerve to the brain. Most vertebrates encode several opsins that are tuned to the dominant wavelengths in their habitats. Yokoyama's programme has been to identify the mutations that control the evolution of opsin spectral sensitivity.

The subfamily of rhodopsins is important for dimlight vision. Most rhodopsins are maximally sensitive to green light, but the coelacanth has two rhodopsin genes, both of which are maximally sensitive to the blue light that is prevalent in deep water. To identify the amino-acid replacements that produced blue sensitivity, Yokoyama *et al.* used phylogenetic methods to infer the sequences of ancestral rhodopsins at numerous branch points during vertebrate evolution[65-67]. They identified amino-acid changes that occurred on the two branches leading from the ancestral lobe-finned fish to each modern-day coelacanth rhodopsin; of these, four replacements — two on each branch — were prime candidates for a functional role on the basis of the sequences and functions of other opsins. Yokoyama *et al*. then tested the hypothesis that these two pairs of mutations caused the evolution of blue sensitivity by introducing the ancestral residues into the coelacanth rhodopsins using directed mutagenesis. The extant and 'ancestralized' rhodopsins were then expressed in cultured cells, purified and reconstituted with chromophore, and their ability to absorb light across the spectrum was assessed *in vitro*. Reintroducing each ancestral amino acid shifted sensitivity towards green light, and reversing the paired replacements on either branch completely restored the ancestral phenotype. Interestingly, one of the replacements occurred in parallel in several other blue-shifted rhodopsins, such as those of dolphins and marine eels.

Yokoyama and co-workers have also examined functional evolution in another opsin subfamily that is optimized for short-wavelength light[68,69]. Most fish and reptiles, and some rodents and birds, use ultraviolet (UV) vision when communicating, selecting mates and foraging. Most mammals, some birds and some amphibians cannot see UV light because their opsins absorb in the violet range. To explore how UV and violet sensitivity evolved, shi and Yokoyama applied the same approach: ancestral sequences that were inferred phylogenetically were reconstructed and functionally characterized *in vitro*[69]. All the ancestral sequences turned out to be UV-sensitive, indicating that responses to longer wavelengths were derived independently in the primate, frog and bird lineages (FIG. 2).

Shi and Yokoyama then sought to identify the genetic basis for the evolutionary shift from UV sensitivity in the common ancestor of birds and reptiles to violet sensitivity in the ancestor of birds. Of the nine amino-acid replacements that separate the two ancestral sequences, three were at sites that were previously identified as participating in spectral tuning. However, when all three replacements were introduced into the reconstructed bird-reptile ancestral opsin, the spectral sensitivity shifted only one-third of the way towards the derived violet-sensitive phenotype. Hence, the three replacements were not sufficient to produce the derived phenotype.

Of the six remaining replacements, which were tested subsequently, only one affected spectral tuning, but addition of this fourth replacement — a site of previously unknown functional effect — produced the full bird-ancestor phenotype.

In these and other studies (for example, REFS [70-73]), Yokoyama's group was among the first to use functional assays to definitively identify mutations that contributed to the evolution of new functions from those that arose coincidentally during the same historic interval. Their work is a beautiful exemplar of how evolutionary analyses can be used to identify sites of previously unsuspected functional importance and enrich our understanding of the mechanisms of gene function.

The fitness landscape of antibiotic resistance

Landscapes have long served as metaphors for the process of adaptation, with individuals or populations moving through the 'space' of possible sequence combinations or frequencies, and with altitude representing fitness[8,74-80]. Landscape topographies have important implications for evolutionary processes. Are they smooth, with a single adaptive peak accessible by selection from every point on the landscape? Or are they rugged, with multiple isolated peaks? Are all mutational paths equally likely, or are some more likely than others? Despite much discussion of how evolution might proceed on different kinds of landscapes, there have been few attempts to characterize empirically the terrain on which adaptive processes actually occur.

Dan Weinreich and colleagues addressed these issues experimentally with TEM β-lactamase, a bacterial enzyme that confers resistance to penicillins and their congeners, such as cefotaxime[81]. Clinical studies had identified five amino-acid replacements that are associated with increased resistance to cefotaxime. To characterize the landscape on which resistance evolves, *TEM* alleles containing all $2^5 = 32$ possible combinations of residues at the five sites were engineered. These alleles were transformed into *Escherichia coli*, and the resistance of each strain to cefotaxime assayed. The quintuple mutant, *TEM**, conferred the highest resistance. For every other combination there was at least one mutation that increased resistance. This landscape has one peak.

But not all adaptive walks to the optimal genotype are plausible. Weinreich *et al*. calculated which of the $5! = 120$ possible paths from the wild-type susceptible allele *TEM^{wt}* to *TEM** were most likely to be taken under the strong selection-weak mutation model of population genetics[77,82]. Weinreich *et al*. found that most paths include one or more neutral or deleterious steps. Only 18 paths involve increased resistance at each step (FIG. 3).

> **Strong selection-weak mutation model**
>
> A population genetic model in which beneficial mutations are fixed sequentially in the population through a series of selective sweeps, and in which neutral and deleterious mutations can be ignored as having low probabilities of fixation.

The number of strictly uphill paths is limited because some replacements are beneficial only in the presence of others, a phenomenon called sign epistasis[83]. For example, the Met182Thr replacement reduces resistance in the *TEM^{wt}* background, but increases it once the Gly238Ser replacement has occurred. Paths that fix Gly238Ser before Met182Thr are plausible, but those in which the temporal order is reversed are not. Weinreich *et al*. pointed out that previous biochemical studies suggested a mechanism for this epistasis. The Met182Thr replacement inhibits protein aggregation and increases stability, but reduces the rate of cefotaxime hydrolysis. In the wild-type background, in which protein aggregation is insignificant and the rate of cefotaxime hydrolysis is low, the cost outweighs the benefit. The Gly238Ser

replacement enhances the hydrolysis rate but reduces stability and increases the tendency to aggregate. In a Gly238Ser background, the benefit of Met182Thr in reducing aggregation and enhancing stability is greater than the cost of a slight reduction in the rate of hydrolysis.

Weinreich *et al*. then calculated the probability that each of the 18 plausible adaptive walks would have been followed, on the basis of the number of available beneficial replacements and the strength of selection at each step. Only a few were found to be likely: depending on the model, just two or four paths accounted for 50% of the total probability of all plausible walks to the peak.

Weinreich *et al*. 's conclusion is surprising. Only a limited number of adaptive walks are plausible, even in a landscape with a single adaptive peak. Although their strategy was not exhaustive — mutations other than the five replacements that they studied might open up new adaptive trajectories through sequence space, and close off others — the general conclusion is well supported: epistasis so severely limits possible histories that the order in which mutations were accrued in a given period of time might be deduced, even when the total number of possible orders is large.

Landscapes, constraints and adaptation

A full mechanistic description of an adaptive landscape requires the maps that describe the transformations of genotype into phenotype and of phenotype into fitness. The genotype-fitness map is the sum of these two component maps.

Dean and colleagues have characterized all three maps controlling coenzyme use in isopropylmalate dehydrogenase (IMDH), an ancient enzyme with a key function in leucine biosynthesis[84]. All IMDHs use the coenzyme NAD^+ as an electron acceptor (FIG. 4a). By contrast, a highly divergent paralogous enzyme — isocitrate dehydrogenase (IDH), which catalyses a step in the Krebs' cycle — uses NAD^+ as a coenzyme in some species and $NADP^+$ in others[85]. To understand why all IMDHs use NAD^+, Dean and colleagues characterized the adaptive landscape on which coenzyme preference evolved.

The first step was to identify the residues controlling coenzyme use. Guided by crystal structures[86,87], directed mutagenesis was used to replace six residues in the coenzyme-binding pocket of *E. coli* IMDH, which uses NAD^+, with the homologus residues from $NADP^+$ using IDHs. Enzyme activities determined *in vitro* revealed a dramatic switch in preference, from 200-fold in favour of NAD^+ to 200-fold in favour of $NADP^+$ (REF. 84).

The second step was to characterize the adaptive terrain. Ninety mutants carrying various combinations of residues at the six key sites were constructed, their kinetic activities assayed *in vitro* and their fitnesses determined using chemostat competition assays[84]. These data were then used to construct the three maps of the adaptive landscape. The genotype-phenotype map of IMDH, which describes the effects that amino-acid replacements have on enzyme activities, shows no epistasis. A simple linear additive model accurately predicts enzyme performance from genotype: adding more residues from $NADP^+$-dependent IDH always increased the preference for $NADP^+$. By contrast, the genotype-fitness map of IMDH shows extensive epistasis, with only a few adaptive walks that lead continuously upwards to the adaptive peak.

Chemostat competition assay

A precise assay of the relative growth rates (fitnesses) of competing strains can be obtained in the chemostat, a continuous culture device that is used to impose starvation for a specific resource in a constant environment.

Why the difference in epistasis between the two maps? The answer lies in the concave relationship between phenotype and fitness (Fig. 4b). As NAD$^+$ performance increases along the right axis, fitness rises steeply, but soon reaches a plateau where further increases in enzyme activity confer little or no benefit. A similar relationship between NADP$^+$ performance and fitness along the left axis leads to a lower fitness plateau. Epistasis occurs because a mutation that improves the performance of an inefficient enzyme increases fitness substantially, but the same improvement in an already-efficient enzyme produces a negligible increase in fitness.

Why is NADP$^+$ use selected against? The metabolic model underlying the IMDH landscape indicates that the product NADPH, which is abundant in cells, strongly inhibits NADP$^+$-using IMDHs, reducing their *in vivo* activity sufficiently to lower their fitness plateau[88]. The NAD$^+$-using wild type retains high fitness because NADH is scarce in cells, and therefore product inhibition is weak. A directed evolution experiment showed that mutations that weaken NADP$^+$ binding also weaken NADPH binding, relieving the inhibition and increasing fitness[88].

Directed evolution

A library of random mutants that have been generated by PCR amplification of a gene is ligated into a plasmid, transformed into a strain and screened for a desired function.

These findings raised another question: if NADP$^+$ use is deleterious in IMDHs, why do some IDHs use it as a coenzyme? Structural differences in the Michaelis complex of IDH reduce affinity for NADPH, allowing the enzyme to use NADP$^+$ efficiently without intense product inhibition (FIG. 4C). But what forces drove some IDH lineages to evolve NADP$^+$ use from an NAD$^+$-using ancestor 3 billion years ago? Cells use NADH to generate energy in the form of ATP and NADPH to provide reducing power for biosynthesis. Dean and colleagues suggested that the switch in coenzyme might have been an adaptation to growth on oxidized carbon sources such as acetate[89]. Consistent with this hypothesis, a genomic association study showed that all 97 sequenced genomes encoding isocitrate lyase (an enzyme that is essential for growth on acetate) have an NADP$^+$-using IDH, whereas all 32 encoding only NAD$^+$-using IDHs lack the lyase[85]. Competition experiments between *E. coli* strains carrying the NADP$^+$-using wild-type IDH and an engineered NAD$^+$-using IDH[90] showed that NADP$^+$ use is indeed advantageous during competition for limiting acetate, but not during competition for limiting glucose[89]. Although these experiments confirm that NADP$^+$ use by IDH is advantageous during growth on acetate, they do not resolve whether its historical emergence was a direct adaptation to growth on acetate. Prokaryotes use — and probably used during their early evolution — various oxidized carbon sources.

Michaelis complex

A complex of substrate bound to enzyme just before catalysis.

Almost three decades ago, Gould and Lewontin drove home the point that the presence of a phenotype does not mean that it was selected for; each hypothesis of adaptation must be tested by proposing a specific mechanism by which the phenotype increases fitness and then determining whether it actually does so through that mechanism[91]. Dean and colleagues' work shows how manipulative techniques can be used to carry out this programme at the molecular level, definitively determining whether specific aspects of a protein's phenotype are indeed adaptations — and if so, why.

Evolution of complexity

A fundamental evolutionary question concerns how biological complexity evolves. Virtually everything that living cells do is regulated by specific interactions between molecules — enzymes and substrates, ligands and receptors, transcription factors and DNA-binding sites. How these kinds of tight partnerships evolve is a crucial question for both molecular and evolutionary biology.

Thornton and colleagues have addressed this question using the specific partnerships between steroid hormones and their receptors as a model. In humans and other tetrapods, the hormone aldosterone activates the mineralocorticoid receptor (MR), an intracellular transcription factor that controls electrolyte homeostasis. MR is related by gene duplication to the glucocorticoid receptor (GR), which is activated by cortisol to regulate the stress response. Aldosterone is known to be a recent novelty that is specific to tetrapods, which raises the question: how did the new MR-aldosterone partnership emerge?

Bridgham *et al.* phylogenetically inferred the sequence of the ancestral corticoid receptor (AncCR), the ~450 million year old protein from which GR and MR evolved by gene duplication[92]. A cDNA coding for the ancient receptor was then synthesized *de novo* and expressed in cultured cells. Using a reporter-gene transcription assay, they found that tiny doses of aldosterone activated the AncCR, as did cortisol and deoxycorticosterone (DOC), the ligand for MRs in extant fish, which is thought to be the ancestral ligand (FIG. 5a). The MRs of all vertebrates studied, including jawless, cartilaginous and teleost fish, were also sensitive to aldosterone.

Why was the ancestral receptor sensitive to aldosterone tens of millions of years before the hormone itself evolved? Thornton's group proposed that AncCR's aldosterone sensitivity was a structural by-product of its sensitivity to DOC. To test this idea, Ortlund *et al.* determined the crystal structures of AncCR in complex with DOC and with aldosterone[93]. As predicted, the ancestral receptor binds aldosterone in precisely the same way that it binds DOC. These results indicate that the ancestral DOC receptor was structurally preadapted for aldosterone activation. When the hormone emerged in tetrapods because of mutations in a steroidogenic enzyme, it recruited the receptor into a new functional partnership. Similarly, Thornton's group has found that several other hormone-receptor pairs also evolved when new molecules co-opted old molecules that once had different functions[94,95].

If MR's phenotype is ancestral, then GR's specificity is due to evolutionary loss of sensitivity to MR's ligands. To understand the mechanistic basis for this shift, Thornton's group first determined the historical interval during which sensitivity to aldosterone and DOC was lost by resurrecting successive ancestral receptors on the gene-family tree (FIG. 5a). Of the 37 changes in protein sequence on this branch, phylogenetic patterns of conservation pointed to five as prime candidates for producing GR's cortisol specificity. Thornton's group introduced these into the ancestral background singly and in combination to determine their effects on receptor function. Two substitutions — Ser106Pro and Leu111Gln — reduced the AncCR's sensitivity to aldosterone by three orders of magnitude while maintaining a GR-like response to cortisol. Strong sign epistasis between these replacements was apparent: alone, Ser106Pro severely reduced sensitivity to all ligands that were tested, whereas Leu111Gln had a negligible effect on function. The Leu111Gln mutant therefore represents a selectively accessible stepping-stone through sequence space to the new function[92,93].

To understand the mechanism for this shift in function, Ortlund *et al.* compared the crystal structure of AncCR to those of the extant and modelled ancestral GRs (FIG. 5b). Leu111Gln has a minimal effect on the receptor-ligand interactions. By contrast, Ser106Pro introduces a

sharp kink into the protein main-chain, repositioning a helix that borders the ligand-binding pocket, destabilizing the structure and compromising activation by all ligands. However, the movement of the helix also repositions site 111, bringing it close to the one hydroxyl group on cortisol that is lacking in DOC and aldosterone. Thus repositioned, the Leu111Gln replacement forms a new hydrogen bond to cortisol's hydroxyl, specifically stabilizing this complex and increasing the sensitivity of the receptor to cortisol. The structure shows how GR's new function evolved, and reveals the mechanism for the epistatic interaction between two key sequence changes.

These two replacements were necessary but not sufficient to yield the full GR phenotype: unlike extant GRs, the Ser106Pro-Leu111Gln mutant ancestor retains a weak response to aldosterone and DOC. Surprisingly, adding the other three strictly conserved sequence changes did not complete the GR phenotype but instead caused a complete loss of receptor function. The crystal structures indicated that, although these three conserved changes destabilized crucial parts of the receptor, 2 of the 32 other historical replacements might epistatically stabilize these same regions. When introduced into the ancestral background, the two additional replacements had negligible effects on function. However, when added to the five strictly conserved changes, these two 'permissive' changes completely restored the receptor's response to cortisol and conferred the full aldosterone-insensitive and DOC-insensitive phenotype (FIG. 5c). In this way, Thornton's group identified seven substitutions that, if introduced in the appropriate order, allow the evolution of the complete GR phenotype.

The Reverend William Paley famously argued that, just as the intricate complexity of a watch implies a design by a watchmaker, so complexity in Nature implies design by God[96]. Evolutionary biologists have typically responded to this challenge by sketching scenarios by which complex biological systems might have evolved through a series of functional intermediates. Thornton and co-workers have gone much further: they have pried open the historical and molecular 'black box' to reconstruct in detail — and with strong empirical support — the history by which a tightly integrated system evolved at the levels of sequence, structure and function.

Box 2

Lessons learnt

Studies in the functional synthesis already suggest answers to evolutionary questions that have been debated for decades.

Convergence

Do convergent phenotypes evolve through identical or different mutations in independent lineages? Apparently both. Studies in the functional synthesis have documented three mechanisms by which convergent phenotypes have evolved:

- Beneficial alleles that are present in an ancestral population can be driven to fixation in different lineages by similar selective pressures. Examples include malathion resistance alleles in blowflies[61] and the evolution of reduced external armour in numerous freshwater species of stickleback fish[14].

- The same beneficial mutations can arise *de novo* and be selected for in separate lineages. Examples include the blue-sensitive rhodopsins of the coelacanth and dolphin[66], insecticide resistance in the E3 esterases of blowflies and houseflies[61] and adaptation to foregut fermentation in the RNases of Colobine monkeys[38] and the lysozymes of artiodactyls, Colobine monkeys and the hoatzin (an unusual leaf-eating bird) [111].

- The same beneficial phenotype can be produced by different mutations in different lineages. Examples include the parallel evolution of light coat colours in beach mice and other mouse species in Arizona and New Mexico[63,112], USA, parallel loss of pigmentation in Mexican cave fish populations[55], parallel changes in gene expression in the wings of *Drosophila* species[16] and increased oxygen affinity in the haemoglobins of two species of goose that are adapted to high-altitude flight[42,113].

Which mechanism pertains in a given case presumably depends on the number of possible genetic ways to produce an adaptive phenotype, the timescale on which it has evolved and whether genetic and ecological contexts preserve ancestral genetic variation on which parallel selection in new environments can act.

Evolutionary step size

Does adaptation proceed by a few mutations of large effect[114-116] or the gradual accumulation of many mutations of small effect[79]? Apparently both. Studies in the functional synthesis demonstrate a key role for mutations of large effect. Single amino-acid replacements caused the evolution of insecticide resistance in blowflies[57], new substrate specificity in lactate dehydrogenase[117], and biofilm formation in *Pseudomonas fluorescens*[21]. Just a few substitutions are needed for spectral tuning in opsins[67,73], antibiotic resistance in TEM lactamase[81] and resistance to newt venom in garter snake sodium channels[37,47]. But steroid receptors evolved by a combination of large-effect mutations that yield new functions, small-effect mutations that optimize them, and permissive mutations of little or no immediate effect that allow the protein to tolerate the others[93].

The preponderance of large-effect mutations in recent studies probably reflects ascertainment bias, at least in part. Detecting mutations with large phenotypic effects (for example, loss of traits[14,17,48,51,55,63,118]) is far easier than detecting those with small effects. The fact that mutations of large effect dominate the literature concerning historical adaptations does not mean that they dominate all evolutionary processes; indeed, *in vitro* protein evolution experiments show that some changes in enzyme function require dozens of mutations, all of small effect[119-122].

One should bear in mind that the terms 'large effect' and 'small effect' are meaningful only at a specified level of analysis. In isopropylmalate dehydrogenase (IMDH), mutations with large effects on enzyme activity elicit small changes in flux and fitness[84]. Similarly, in beach mice, a near loss-of-function mutation in the melanocortin-1 receptor gene (*Mc1r*) contributes at most 28% of the variation in coat colour[63]. Whether these mutations should be regarded as having large or small effects depends on the trait under discussion.

Epistasis

Does epistasis determine the trajectory and the outcome of evolutionary processes? The answer is yes for the trajectory, and maybe for the outcome. Sign epistasis produces 'rugged' landscapes in which multiple patches of sequence space can be isolated by genetic intermediates of low fitness[83]. On such landscapes, escape from local optima to ascend high fitness peaks depends on chance events that render evolutionary trajectories less predictable, more historically contingent[74]. How important sign epistasis actually is in structuring adaptive landscapes and influencing evolutionary processes is a debate that has smouldered for almost a century[74,79,123-125].

The functional synthesis enables the impact of epistasis to be evaluated directly by characterizing the phenotypic and fitness effects of mutations singly and in combination. Results to date indicate that sign epistasis is widespread and influential. In the evolution of hormone receptors[92], antibiotic resistance[81] and coenzyme use[84], sign epistasis blocks

many adaptive walks through sequence space, so that mutations must be acquired in a particular order to avoid deleterious intermediates. Still, there is no evidence that sign epistasis renders some adaptive peaks isolated, inaccessible by mutation and selection alone. However, no study has addressed whether more distant adaptive peaks might exist, let alone whether they are accessible.

Future directions

The functional synthesis is just beginning to take hold. Even so, the handful of studies reviewed here, along with others that take a similar tack[36-47], have already provided insights into fundamental evolutionary questions that had remained intractable for decades (BOX 2). As more examples are accumulated, it will be possible to move from dissecting specific systems to making broad characterizations.

Functional tests should become routine in studies of molecular evolution. Statistical inferences from sequence data will remain important, but they should be treated as a starting point, not the centrepiece or end of analysis as in the old paradigm. In our opinion, it is now incumbent on evolutionary biologists to experimentally test their statistically generated hypotheses before making strong claims about selection or other evolutionary forces. With the advent of new capacities, the standards of evidence in the field must change accordingly. To meet this standard, evolutionary biologists will need to be trained in molecular biology and be prepared to establish relevant collaborations across disciplines.

The functional synthesis should be broadened to address other important issues. For example, the origin and diversification of gene families is a key question. What are the roles of subfunctionalization[97], neofunctionalization[98] and the optimization of 'promiscuous' ancestral functions[99-102] in gene-family evolution? Another issue is the evolution of gene regulation. How important are changes in *cis*-regulatory regions versus those in protein-coding sequences for the evolution of adaptive phenotypes[50,103,104]? The functional synthesis has the capacity to resolve these questions by identifying the specific genetic changes that have produced evolutionary shifts in gene function and expression.

A key challenge for the functional synthesis is to thoroughly connect changes in molecular function to organismal phenotype and fitness. Ideally, results obtained *in vitro* should be verified *in vivo*. Transgenic evolutionary studies identifying the functional impact of historical mutations have been conducted in microbes and a few model plant and animal species[14-20], but an expanded repertoire of models will be required to reach this goal for other taxa. By integrating the functional synthesis with advances in developmental genetics and neurobiology, this approach has the potential to yield important insights into the evolution of development, behaviour and physiology. Experimental studies of natural selection in the laboratory can also be enriched by functional approaches to characterize the specific genetic changes that underlie the evolution of adaptive phenotypes[21,105-108].

Finally, the functional synthesis should move beyond studies of single genes to analyse the evolution of pathways and networks that are made up of multiple genes. Most cellular processes and organismal phenotypes are determined by metabolic pathways and regulatory cascades; however, we have little empirical knowledge on the evolutionary history of such networks[109]. By studying the mechanistic history of the members of an interacting gene set, it should be possible to reconstruct how metabolic and regulatory gene networks emerged and functionally diversified over time.

Conclusion

Research paradigms shift when opportunistic scientists adopt new approaches that allow them to ask new questions and produce new knowledge[110]. By merging the cultures of evolution and molecular biology, the functional synthesis refocuses research on a new agenda — the empirical evaluation of the mechanisms by which genetic change has produced adaptive phenotypes — using new tools and revised standards of inference. With methods and insights gathered across subdisciplines, this programme provides ample opportunities for the reunification of biology. As its strategies mature and its use expands, the functional synthesis will provide an ever richer view of the processes of evolution and the history of molecular systems.

DATABASES

UniProtKB: http://ca.expasy.org/sprot

GR | IDH | IMDH | MR | TEM[wt]

FURTHER INFORMATION

Anthony M. Dean's homepage: http://www.cbs.umn.edu/labs/deanlab/

Joseph W. Thornton's homepage: http://www.uoregon.edu/~joet

References

1. Dykhuizen DE, Hartl DL. Selection in chemostats. Microbiol. Rev 1983;47:150–168. [PubMed: 6308409]

2. Elena SF, Lenski RE. Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. Nature Rev. Genet 2003;4:457–469. [PubMed: 12776215]

3. Losos JB, Jackman TR, Larson A, Queiroz K, Rodriguez-Schettino L. Contingency and determinism in replicated adaptive radiations of island lizards. Science 1998;279:2115–2118. [PubMed: 9516114]

4. Powers DA, Lauerman T, Crawford D, DiMichele L. Genetic mechanisms for adapting to a changing environment. Annu. Rev. Genet 1991;25:629–659. [PubMed: 1812817]

5. Wourms MK, Wasserman FE. Butterfly wing markings are more advantageous during handling than during the initial strike of an avian predator. Evolution 1985;39:845–851.

6. Moller AP. Female choice selects for male sexual tail ornaments in the monogamous swallow. Nature 1988;332:640–642.

7. Rainey PB, Rainey K. Evolution of cooperation and conflict in experimental bacterial populations. Nature 2003;425:72–74. [PubMed: 12955142]

8. Denver DR, et al. The transcriptional consequences of mutation and natural selection in *Caenorhabditis elegans*. Nature Genet 2005;37:544–548. [PubMed: 15852004]

9. Endler JA. Natural selection on color patterns in *Poecilia reticulata*. Evolution 1980;34:76–91.

10. Thornton JW. Resurrecting ancient genes: experimental analysis of extinct molecules. Nature Rev. Genet 2004;5:366–375. [PubMed: 15143319]An introduction to gene 'resurrection' — that is, phylogenetic reconstruction, biochemical synthesis and functional characterization of ancient sequences — as a strategy for testing evolutionary hypotheses.

11. Fitzpatrick MJ, Feder E, Rowe L, Sokolowski MB. Maintaining a behaviour polymorphism by frequency-dependent selection on a single gene. Nature 2007;447:210–212. [PubMed: 17495926]

Polymorphism in activity of a cGMP-dependent protein kinase that elicits different foraging behaviours in *Drosophila melanogaster* larva is maintained by negative frequency-dependent selection during starvation.

12. Douglas SJ, Dawson-Scully K, Sokolowski MB. The neurogenetics and evolution of food-related behaviour. Trends Neurosci 2005;28:644–652. [PubMed: 16203044]

13. Toth AL, Robinson GE. Evo-devo and the evolution of social behavior. Trends Genet 2007;23:334–341. [PubMed: 17509723]

14. Colosimo PF, et al. Widespread parallel evolution in sticklebacks by repeated fixation of Ectodysplasin alleles. Science 2005;307:1928–1933. [PubMed: 15790847]Using transgenic, phylogenetic and quantitative genetic analysis, the authors identify an allele that is involved in the loss of external armour during the evolution of numerous freshwater stickleback populations.

15. Gompel N, Prud'homme B, Wittkopp PJ, Kassner VA, Carroll SB. Chance caught on the wing: *cis*-regulatory evolution and the origin of pigment patterns in *Drosophila*. Nature 2005;433:481–487. [PubMed: 15690032]This pioneering study in the evolution of development (along with reference 16) used transgenic techniques to identify decisively specific regulatory elements that underlie evolutionary differences in the expression of genes that drive pigmentation patterns between fruitfly species.

16. Prud'homme B, et al. Repeated morphological evolution through *cis*-regulatory changes in a pleiotropic gene. Nature 2006;440:1050–1053. [PubMed: 16625197]

17. Konishi S, et al. An SNP caused loss of seed shattering during rice domestication. Science 2006;312:1392–1396. [PubMed: 16614172]A beautiful example of the functional synthesis to study the evolution of development: fine quantitative genetic mapping identified a single substitution associated with the loss of seed shattering that occurs during rice domestication; transgenic and functional analysis established the specific effects of the historical mutation on gene expression, seed development and the shattering phenotype.

18. Wang H, et al. The origin of the naked grains of maize. Nature 2005;436:714–719. [PubMed: 16079849]

19. Daborn PJ, et al. A single p450 allele associated with insecticide resistance in *Drosophila*. Science 2002;297:2253–2256. [PubMed: 12351787]

20. Chung H, et al. *Cis*-regulatory elements in the accord retrotransposon result in tissue-specific expression of the *Drosophila melanogaster* insecticide resistance gene *Cyp6g1*. Genetics 2007;175:1071–1077. [PubMed: 17179088]Using genetic manipulation in fruitflies, the authors show decisively that a transposon insertion in the regulatory region of a gene that metabolizes insecticides is sufficient to recapitulate the evolution of DDT (dichloro-diphenyl-trichloroethane) resistance.

21. Bantinaki E, et al. Adaptive divergence in experimental populations of *Pseudomonas fluorescens*. III. Mutational origins of wrinkly spreader diversity. Genetics 2007;176:441–453. [PubMed: 17339222]

22. Gilbert SF, Opitz JM, Raff RA. Resynthesizing evolutionary and developmental biology. Dev. Biol 1996;173:357–372. [PubMed: 8605997]

23. Wilson, EO. Sociobiology: The New Synthesis. Belknap; Cambridge: 1975.

24. Hubby JL, Lewontin RC. A molecular approach to the study of genic heterozygosity in natural populations. I. The number of alleles at different loci in *Drosophila pseudoobscura*. Genetics 1966;54:577–594. [PubMed: 5968642]

25. Kreitman M. Nucleotide polymorphism at the alcohol dehydrogenase locus of *Drosophila melanogaster*. Nature 1983;304:412–417. [PubMed: 6410283]

26. Golding GB, Dean AM. The structural basis of molecular adaptation. Mol. Biol. Evol 1998;15:355–369. [PubMed: 9549087]A review summarizing classic early work in the functional synthesis.

27. McDonald JH, Kreitman M. Adaptive protein evolution at the *adh* locus in *Drosophila*. Nature 1991;351:652–654. [PubMed: 1904993]

28. Yang Z, Nielsen R. Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. Mol. Biol. Evol 2002;19:908–917. [PubMed: 12032247]

29. Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 1989;123:585–596. [PubMed: 2513255]

30. Nielsen R. Molecular signatures of natural selection. Annu. Rev. Genet 2005;39:197–218. [PubMed: 16285858]

31. Nielsen R. Statistical tests of selective neutrality in the age of genomics. Heredity 2001;86:641–647. [PubMed: 11595044]

32. Eyre-Walker A. Changing effective population size and the McDonald-Kreitman test. Genetics 2002;62:2017–2024. [PubMed: 12524367]

33. Koehn RK, Hilbish TJ. The adaptive importance of genetic variation. Am. Sci 1987;75:134–141.

34. Watt, WB. The Evolution of Population Biology. Singh, RS.; Uyenoyama, MK., editors. Cambridge Univ. Press; Cambridge: 2004.

35. Wheat CW, Watt WB, Pollock DD, Schulte PM. From DNA to fitness differences: sequences and structures of adaptive variants of Colias phosphoglucose isomerase (PGI). Mol. Biol. Evol 2006;23:499–512. [PubMed: 16292000]

36. Brideau NJ, et al. Two Dobzhansky-Muller genes interact to cause hybrid lethality in *Drosophila*. Science 2006;314:1292–1295. [PubMed: 17124320]

37. Geffeney SL, Fujimoto E, Brodie ED 3rd, Brodie ED Jr, Ruben PC. Evolutionary diversification of TTX-resistant sodium channels in a predator-prey interaction. Nature 2005;434:759–763. [PubMed: 15815629]An elegant mechanistic-structural explanation of the repeated evolution of resistance to the tetrodotoxin of toxic newt prey, caused by amino-acid replacements in the voltage-gated sodium channels of garter snake muscles.

38. Zhang J. Parallel adaptive origins of digestive RNases in Asian and African leaf monkeys. Nature Genet 2006;38:819–823. [PubMed: 16767103]Evolution of foregut fermentation in Asian and African leaf-eating monkeys is characterized by parallel amino-acid replacements that produce similar functional shifts in digestive RNases.

39. Zhang J, et al. The crystal structure of a high oxygen affinity species of haemoglobin (bar-headed goose haemoglobin in the oxy form). J. Mol. Biol 1996;255:484–493. [PubMed: 8568892]

40. Zhang J, Rosenberg HF. Complementary advantageous substitutions in the evolution of an antiviral RNase of higher primates. Proc. Natl Acad. Sci. USA 2002;99:5486–5491. [PubMed: 11917138]

41. Jermann TM, Opitz JG, Stackhouse J, Benner SA. Reconstructing the evolutionary history of the artiodactyl ribonuclease superfamily. Nature 1995;374:57–59. [PubMed: 7532788]

42. Jessen TH, Weber RE, Fermi G, Tame J, Braunitzer G. Adaptation of bird hemoglobins to high altitudes: demonstration of molecular mechanism by protein engineering. Proc. Natl Acad. Sci. USA 1991;88:6519–6522. [PubMed: 1862080]

43. Gaucher EA, Thomson JM, Burgan MF, Benner SA. Inferring the palaeoenvironment of ancient bacteria on the basis of resurrected proteins. Nature 2003;425:285–288. [PubMed: 13679914]

44. Chang BS, Jonsson K, Kazmi MA, Donoghue MJ, Sakmar TP. Recreating a functional ancestral archosaur visual pigment. Mol. Biol. Evol 2002;19:1483–1489. [PubMed: 12200476]

45. Thomson JM, et al. Resurrecting ancestral alcohol dehydrogenases from yeast. Nature Genet 2005;37:630–635. [PubMed: 15864308]Ancestral yeast alcohol dehydrogenase (ADH) was reconstructed, expressed and shown to have the functional characteristics that are typical of extant ADH1, which is involved in ethanol production, rather than ADH2, which is involved in ethanol consumption.

46. Ugalde JA, Chang BS, Matz MV. Evolution of coral pigments recreated. Science 2004;305:1433. [PubMed: 15353795]

47. Soong TW, Venkatesh B. Adaptive evolution of tetrodotoxin resistance in animals. Trends Genet 2006;22:621–626. [PubMed: 16959367]

48. Doebley J, Stec A, Hubbard L. The evolution of apical dominance in maize. Nature 1997;386:485–488. [PubMed: 9087405]

49. Shapiro MD, et al. Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. Nature 2004;428:717–723. [PubMed: 15085123]

50. Stern DL. Evolutionary developmental biology and the problem of variation. Evolution 2000;54:1079–1091. [PubMed: 11005278]

51. Sucena E, Delon I, Jones I, Payre F, Stern DL. Regulatory evolution of *shavenbaby/ovo* underlies multiple cases of morphological parallelism. Nature 2003;424:935–938. [PubMed: 12931187]

52. Shimizu KK, et al. Darwinian selection on a selfing locus. Science 2004;306:2081–2084. [PubMed: 15604405]

53. de Meaux J, Pop A, Mitchell-Olds T. *cis*-regulatory evolution of chalcone-synthase expression in the genus *Arabidopsis*. Genetics 2006;174:2181–2202. [PubMed: 17028316]

54. Benderoth M, et al. Positive selection driving diversification in plant secondary metabolism. Proc. Natl Acad. Sci. USA 2006;103:9118–9123. [PubMed: 16754868]

55. Protas ME, et al. Genetic analysis of cavefish reveals molecular convergence in the evolution of albinism. Nature Genet 2006;38:107–111. [PubMed: 16341223]

56. Osborne KA, et al. Natural behavior polymorphism due to a cGMP-dependent protein kinase of *Drosophila*. Science 1997;277:834–836. [PubMed: 9242616]

57. Newcomb RD, et al. A single amino acid substitution converts a carboxylesterase to an organophosphorus hydrolase and confers insecticide resistance on a blowfly. Proc. Natl Acad. Sci. USA 1997;94:7464–7468. [PubMed: 9207114]

58. Newcomb RD, East PD, Russell RJ, Oakeshott JG. Isolation of α cluster esterase genes associated with organophosphate resistance in Lucilia cuprina. Insect Mol. Biol 1996;5:211–216. [PubMed: 8799740]

59. Parker AG, Campbell PM, Spackman ME, Russell RJ, Oakeshott JG. Comparison of an esterase associated with organophosphate resistance in *Lucilia cuprina* with an orthologue not associated with resistance in *Drosophila melanogaster*. Pestic. Biochem. Physiol 1996;55:85–99. [PubMed: 8980033]

60. Newcomb RD, Campbell PM, Russell RJ, Oakeshott JG. cDNA cloning, baculovirus-expression and kinetic properties of the esterase, E3, involved in organophosphorus resistance in *Lucilia cuprina*. Insect Biochem. Mol. Biol 1997;27:15–25. [PubMed: 9061925]

61. Hartley CJ, et al. Amplification of DNA from preserved specimens shows blowflies were preadapted for the rapid evolution of insecticide resistance. Proc. Natl Acad. Sci. USA 2006;103:8757–8762. [PubMed: 16723400]

62. Claudianos C, Russell RJ, Oakeshott JG. The same amino acid substitution in orthologous esterases confers organophosphate resistance on the house fly and a blowfly. Insect Biochem. Mol. Biol 1999;29:675–686. [PubMed: 10451921]

63. Hoekstra HE, Hirschmann RJ, Bundey RA, Insel PA, Crossland JP. A single amino acid mutation contributes to adaptive beach mouse color pattern. Science 2006;313:101–104. [PubMed: 16825572]

64. Steiner CC, Weber JN, Hoekstra HE. Adaptive variation in beach mice produced by interacting pigmentation genes. PLoS Biol 2007;5:e219. [PubMed: 17696646]

65. Yokoyama S, Zhang H, Radlwimmer FB, Blow NS. Adaptive evolution of color vision of the Comoran coelacanth (*Latimeria chalumnae*). Proc. Natl Acad. Sci. USA 1999;96:6279–6284. [PubMed: 10339578]

66. Yokoyama S, Tada T. Adaptive evolution of the African and Indonesian coelacanths to deep-sea environments. Gene 2000;261:35–42. [PubMed: 11164035]

67. Yokoyama S. Color vision of the coelacanth (*Latimeria chalumnae*) and adaptive evolution of rhodopsin (*RH1*) and rhodopsin-like (*RH2*) pigments. J. Hered 2000;91:215–220. [PubMed: 10833047]

68. Shi Y, Radlwimmer FB, Yokoyama S. Molecular genetics and the evolution of ultraviolet vision in vertebrates. Proc. Natl Acad. Sci. USA 2001;98:11731–11736. [PubMed: 11573008]

69. Shi Y, Yokoyama S. Molecular analysis of the evolutionary significance of ultraviolet vision in vertebrates. Proc. Natl Acad. Sci. USA 2003;100:8308–8313. [PubMed: 12824471]

70. Yokoyama S, Starmer WT, Takahashi Y, Tada T. Tertiary structure and spectral tuning of UV and violet pigments in vertebrates. Gene 2006;365:95–103. [PubMed: 16343816]

71. Yokoyama S, Radlwimmer FB. The molecular genetics and evolution of red and green color vision in vertebrates. Genetics 2001;158:1697–1710. [PubMed: 11545071]

72. Yokoyama S, Radlwimmer FB. The molecular genetics of red and green color vision in mammals. Genetics 1999;153:919–932. [PubMed: 10511567]

73. Yokoyama S, Radlwimmer FB. The 'five-sites' rule and the evolution of red and green color vision in mammals. Mol. Biol. Evol 1998;15:560–567. [PubMed: 9580985]

74. Wright S. The roles of mutation, inbreeding, crossbreeding, and selection in evolution. Proc. 6th Int. Cong. Genet 1932;1:356–366.

75. Phillips PC, Arnold SJ. Visualizing multivariate selection. Evolution 1989;43:1209–1222.

76. Gavrilets S. A dynamical theory of speciation on holey adaptive landscapes. Am. Nat 1999;154:1–22.

77. Gillespie JH. Molecular evolution over the mutational landscape. Evolution 1984;38:1116–1129.

78. Kauffman, SA. The Origins of Order: Self-organization and Selection in Evolution. Oxford Univ. Press; Oxford: 1993.

79. Fisher, RA. The Genetical Theory of Natural Selection. Oxford Univ. Press; Oxford: 1930.

80. Maynard Smith J. Natural selection and the concept of a protein space. Nature 1970;225:563–564. [PubMed: 5411867]

81. Weinreich DM, Delaney NF, Depristo MA, Hartl DL. Darwinian evolution can follow only very few mutational paths to fitter proteins. Science 2006;312:111–114. [PubMed: 16601193]

82. Gillespie JH. A simple stochastic gene substitution model. Theor. Popul. Biol 1983;23:2020–2015.

83. Weinreich DM, Watson RA, Chao L. Perspective: sign epistasis and genetic constraint on evolutionary trajectories. Evolution 2005;59:1165–1174. [PubMed: 16050094]

84. Lunzer M, Miller SP, Felsheim R, Dean AM. The biochemical architecture of an ancient adaptive landscape. Science 2005;310:499–501. [PubMed: 16239478]

85. Zhu G, Golding GB, Dean AM. The selective cause of an ancient adaptation. Science 2005;307:1279–1282. [PubMed: 15653464]

86. Hurley JH, Dean AM. Structure of 3-isopropylmalate dehydrogenase in complex with NAD^+: ligand-induced loop closing and mechanism for cofactor specificity. Structure 1994;2:1007–1016. [PubMed: 7881901]

87. Hurley JH, Dean AM, Koshland DE Jr, Stroud RM. Catalytic mechanism of $NADP^+$-dependent isocitrate dehydrogenase: implications from the structures of magnesium-isocitrate and $NADP^+$ complexes. Biochemistry 1991;30:8671–8678. [PubMed: 1888729]

88. Miller SP, Lunzer M, Dean AM. Direct demonstration of an adaptive constraint. Science 2006;314:458–461. [PubMed: 17053145]

89. Dean AM, Golding GB. Protein engineering reveals ancient adaptive replacements in isocitrate dehydrogenase. Proc. Natl Acad. Sci. USA 1997;94:3104–3109. [PubMed: 9096353]

90. Chen R, Greer A, Dean AM. A highly active decarboxylating dehydrogenase with rationally inverted coenzyme specificity. Proc. Natl Acad. Sci. USA 1995;92:11666–11670. [PubMed: 8524825]

91. Gould SJ, Lewontin RC. The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. Proc. R. Soc. Lond. B Biol. Sci 1979;205:581–598. [PubMed: 42062]

92. Bridgham JT, Carroll SM, Thornton JW. Evolution of hormone-receptor complexity by molecular exploitation. Science 2006;312:97–101. [PubMed: 16601189]

93. Ortlund EA, Bridgham JT, Redinbo MR, Thornton JW. Crystal structure of an ancient protein: evolution of a new function by conformational epistasis. Science. August 16;2007 doi:101126/science.1142819

94. Thornton JW, Need E, Crews D. Resurrecting the ancestral steroid receptor: ancient origin of estrogen signaling. Science 2003;301:1714–1717. [PubMed: 14500980]

95. Thornton JW. Evolution of vertebrate steroid receptors from an ancestral estrogen receptor by ligand exploitation and serial genome expansions. Proc. Natl Acad. Sci. USA 2001;98:5671–5676. [PubMed: 11331759]

96. Paley, W. Natural Theology: or, Evidences of the Existence and Attributes of the Deity, Collected from the Appearances of Nature. E. Paulder; London: 1802.

97. Force A, et al. Preservation of duplicate genes by complementary, degenerate mutations. Genetics 1999;151:1531–1545. [PubMed: 10101175]

98. Ohno, S. Evolution by Gene Duplication. Springer; New York: 1970.

99. Copley SD. Enzymes with extra talents: moonlighting functions and catalytic promiscuity. Curr. Opin. Chem. Biol 2003;7:265–272. [PubMed: 12714060]

100. O'Brien PJ, Herschlag D. Catalytic promiscuity and the evolution of new enzymatic activities. Chem. Biol. (London) 1999;6:R91–R105.

101. Khersonsky O, Roodveldt C, Tawfik DS. Enzyme promiscuity: evolutionary and mechanistic aspects. Curr. Opin. Chem. Biol 2006;10:498–508. [PubMed: 16939713]

102. Gerlt JA, Babbitt PC. Divergent evolution of enzymatic function: mechanistically diverse superfamilies and functionally distinct suprafamilies. Annu. Rev. Biochem 2001;70:209–246. [PubMed: 11395407]

103. Carroll SB. Endless forms: the evolution of gene regulation and morphological diversity. Cell 2000;101:577–580. [PubMed: 10892643]

104. Hoekstra HE, Coyne JA. The locus of evolution: evo devo and the genetics of adaptation. Evolution 2007;61:995–1016. [PubMed: 17492956]

105. Leu JY, Murray AW. Experimental evolution of mating discrimination in budding yeast. Curr. Biol 2006;16:280–286. [PubMed: 16461281]

106. Velicer GJ, et al. Comprehensive mutation identification in an evolved bacterial cooperator and its cheating ancestor. Proc. Natl Acad. Sci. USA 2006;103:8107–8112. [PubMed: 16707573]

107. Woods R, Schneider D, Winkworth CL, Riley MA, Lenski RE. Tests of parallel molecular evolution in a long-term experiment with *Escherichia coli*. Proc. Natl Acad. Sci. USA 2006;103:9107–9112. [PubMed: 16751270]

108. Zhong S, Khodursky A, Dykhuizen DE, Dean AM. Evolutionary genomics of ecological specialization. Proc. Natl Acad. Sci. USA 2004;101:11719–11724. [PubMed: 15289609]

109. Davidson, EH. The Regulatory Genome: Gene Regulatory Networks in Development and Evolution. Academic; Burlington: 2006.

110. Kuhn, T. The Structure of Scientific Revolutions. Univ. Chicago Press; Chicago: 1996.

111. Kornegay JR, Schilling JW, Wilson AC. Molecular adaptation of a leaf-eating bird: stomach lysozyme of the hoatzin. Mol. Biol. Evol 1994;11:921–928. [PubMed: 7815930]

112. Nachman MW, Hoekstra HE, D'Agostino SL. The genetic basis of adaptive melanism in pocket mice. Proc. Natl Acad. Sci. USA 2003;100:5268–5273. [PubMed: 12704245]

113. Hiebl I, Braunitzer G, Schneeganss D. The primary structures of the major and minor hemoglobin-components of adult Andean goose (*Chloephaga melanoptera*, Anatidae): the mutation Leu > Ser in position 55 of the β-chains. Biol. Chem. Hoppe-Seyler 1987;368:1559–1569. [PubMed: 3442599]

114. Bateson, W. Preface from Mendel's Principles of Heredity: A Defense. Cambridge Univ. Press; Cambridge: 1902.

115. Goldschmidt, R. The material basis of evolution. Yale; New Haven: 1940.

116. Bateson, W. Materials for the Study of Variation Treated with Especial Regard to Discontinuity in the Origin of Species. Macmillan; London: 1894.

117. Wilks HM, et al. A specific, highly active malate dehydrogenase by redesign of a lactate dehydrogenase framework. Science 1988;242:1541–1544. [PubMed: 3201242]

118. Cresko WA, et al. Parallel genetic basis for repeated evolution of armor loss in Alaskan threespine stickleback populations. Proc. Natl Acad. Sci. USA 2004;101:6050–6055. [PubMed: 15069186]

119. Spiller B, Gershenson A, Arnold FH, Stevens RC. A structural view of evolutionary divergence. Proc. Natl Acad. Sci. USA 1999;96:12305–12310. [PubMed: 10535917]

120. Rothman SC, Voorhies M, Kirsch JF. Directed evolution relieves product inhibition and confers *in vivo* function to a rationally designed tyrosine aminotransferase. Protein Sci 2004;13:763–772. [PubMed: 14767072]

121. Oue S, Okamoto A, Yano T, Kagamiyama H. Redesigning the substrate specificity of an enzyme by cumulative effects of the mutations of non-active site residues. J. Biol. Chem 1999;274:2344–2349. [PubMed: 9891001]

122. Hsu CC, Hong Z, Wada M, Franke D, Wong CH. Directed evolution of D-sialic acid aldolase to L-3-deoxy-manno-2-octulosonic acid (L-KDO) aldolase. Proc. Natl Acad. Sci. USA 2005;102:9122–9126. [PubMed: 15967977]

123. Whitlock MC, Phillips PC, Moore FBG, Tonsor SJ. Multiple fitness peaks and epistasis. Annu. Rev. Ecol. Syst 1995;26:601–629.

124. Provine, WB. The origins of theoretical population genetics. Univ. Chicago Press; Chicago: 1971.

125. Gavrilets S. Evolution and speciation on holey adaptive landscapes. Trends Ecol. Evol 1997;12:307–312.

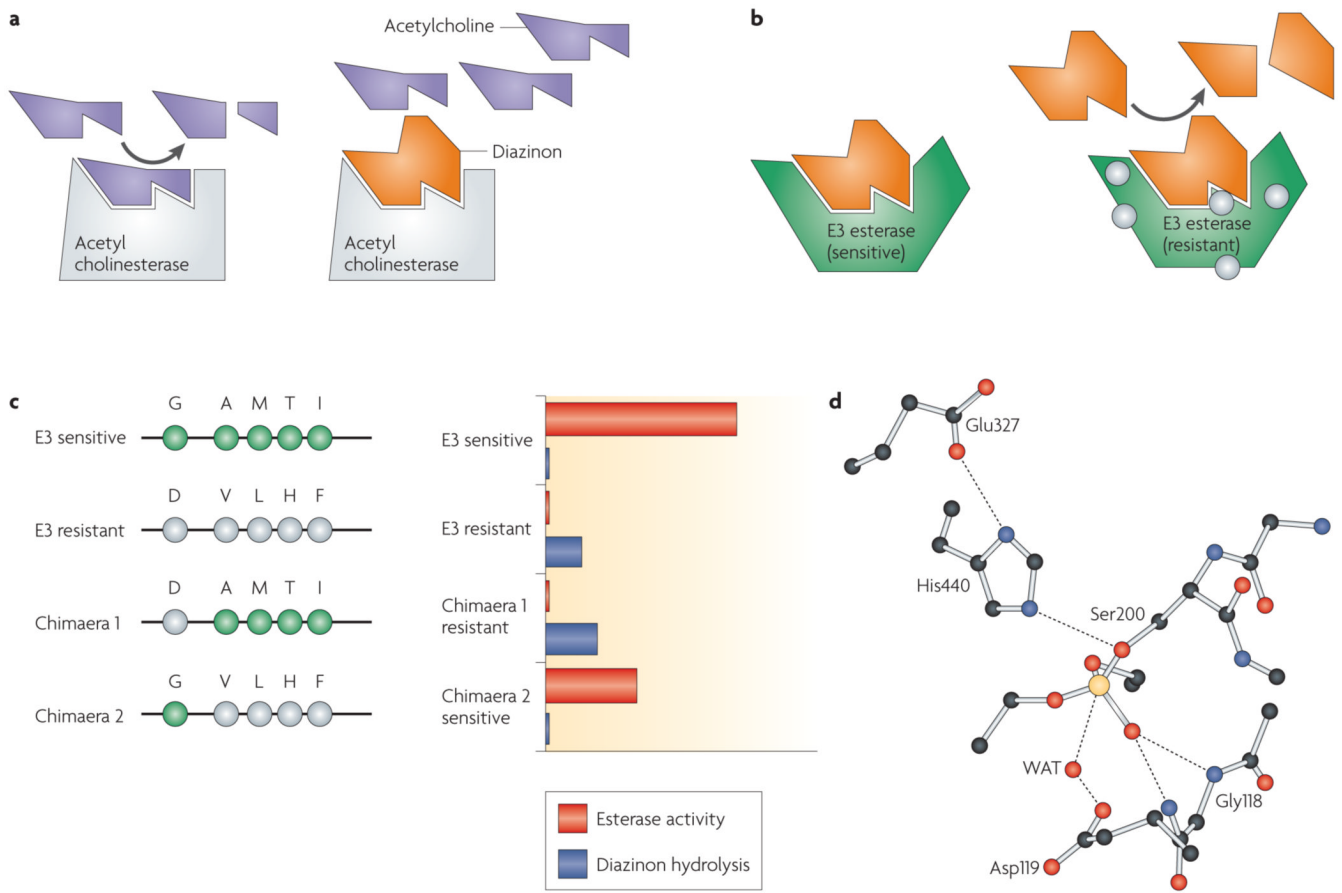126. Yokoyama S. Molecular evolution of color vision in vertebrates. Gene 2002;300:69–78. [PubMed: 12468088]

Figure 1. Evolution of insecticide resistance by a single amino-acid change
a | Diazinon, like other organophosphates, binds to acetylcholinesterase (AChE) and blocks neurotransmitter turnover, causing acetylcholine accumulation and neurotoxicity. **b** | Diazinon also binds to and blocks E3, another esterase. Newcomb *et al*. found that, in populations of diazinon-resistant blowflies, the esterase isozyme E3 hydrolyses diazinon, rendering it non-toxic[57]. Compared with sensitive blowfly populations, E3 from resistant populations contains five amino-acid replacements (represented as grey circles). **c** | Newcomb *et al*. prepared chimeric E3s to test the role of Gly137Asp replacement, the only replacement that is located in the enzyme's active site. Chimaera 1 contains all the residues from sensitive E3 (represented as green circles) except Asp137 (from the resistant allele, represented as a grey circle). Chimaera 2 contains all the residues from the resistant allele except Gly137. The graph shows the loss of carboxyesterase activity and the gain of diazinon-hydrolysis activity in the resistant allele and Chimaera 1, but not Chimaera 2. **d** | A model of the active-site region of a resistant esterase with the Gly137Asp replacement (Asp119 in the model). Asp119 is positioned to act as a general base, activating a water molecule (WAT) for nucleophylic attack on the insecticide's phosphorous atom. The P-O bond to Ser200 breaks, diethyl phosphate is released and the active E3 enzyme is restored. Panel **d** reproduced with permission from REF. [57] © (1997) National Academy of Sciences (USA).

Figure 2. Evolution of spectral sensitivity in vertebrate opsins

a | To understand the evolution of new spectral sensitivities in the short-wave opsins (SWO), Yokoyama and colleagues inferred their phylogeny. The wavelength to which each pigment is most sensitive is shown in parentheses; UV-sensitive pigments are shown in purple and blue-sensitive pigments in blue. Sequences of ancestral SWOs at each node were inferred, and the optimal wavelength for each ancestral pigment is shown in circles. The shift from the UV-sensitive pigment in the bird-reptile ancestor (represented as a purple circle) to the blue-sensitive pigment in the ancestor of birds (represented as a blue circle) is indicated by an arrow. **b** | The inferred sequences of ancenstral SWOs at each node were 'resurrected' by mutagenesis, expressed in mammalian cells (COS1), reconstituted with the chromophore 11-*cis*-retinal and

tested *in vitro*. **c** | Introducing four historical substitutions from this branch into the bird-reptile ancestral SWO is sufficient to recapitulate the evolution of the bird ancestor's new spectral phenotype. Parts **a** and **c** reproduced with permission from REF. [69] © (2003) National Academy of Sciences (USA). Part **b** reproduced with permission from REF. [126] © (2002) Elsevier Sciences.
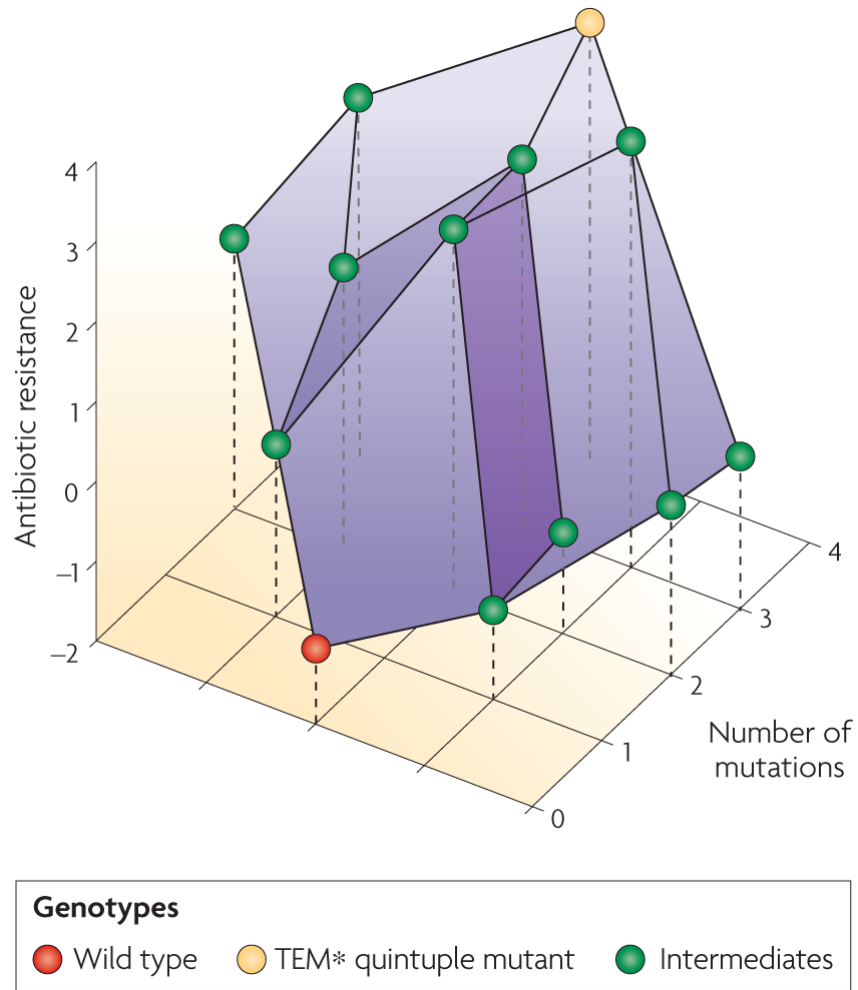
Figure 3. Uphill adaptive walks among amino-acid replacements in the adaptive landscape of TEM β-lactamase

Nodes represent genotypes and edges represent mutations. Only pathways that increase the antibiotic resistance at each step are shown. Parallelograms formed by four connected genotypes indicate additivity among mutations. Asymmetrical quadrilaterals indicate the presence of epistasis.
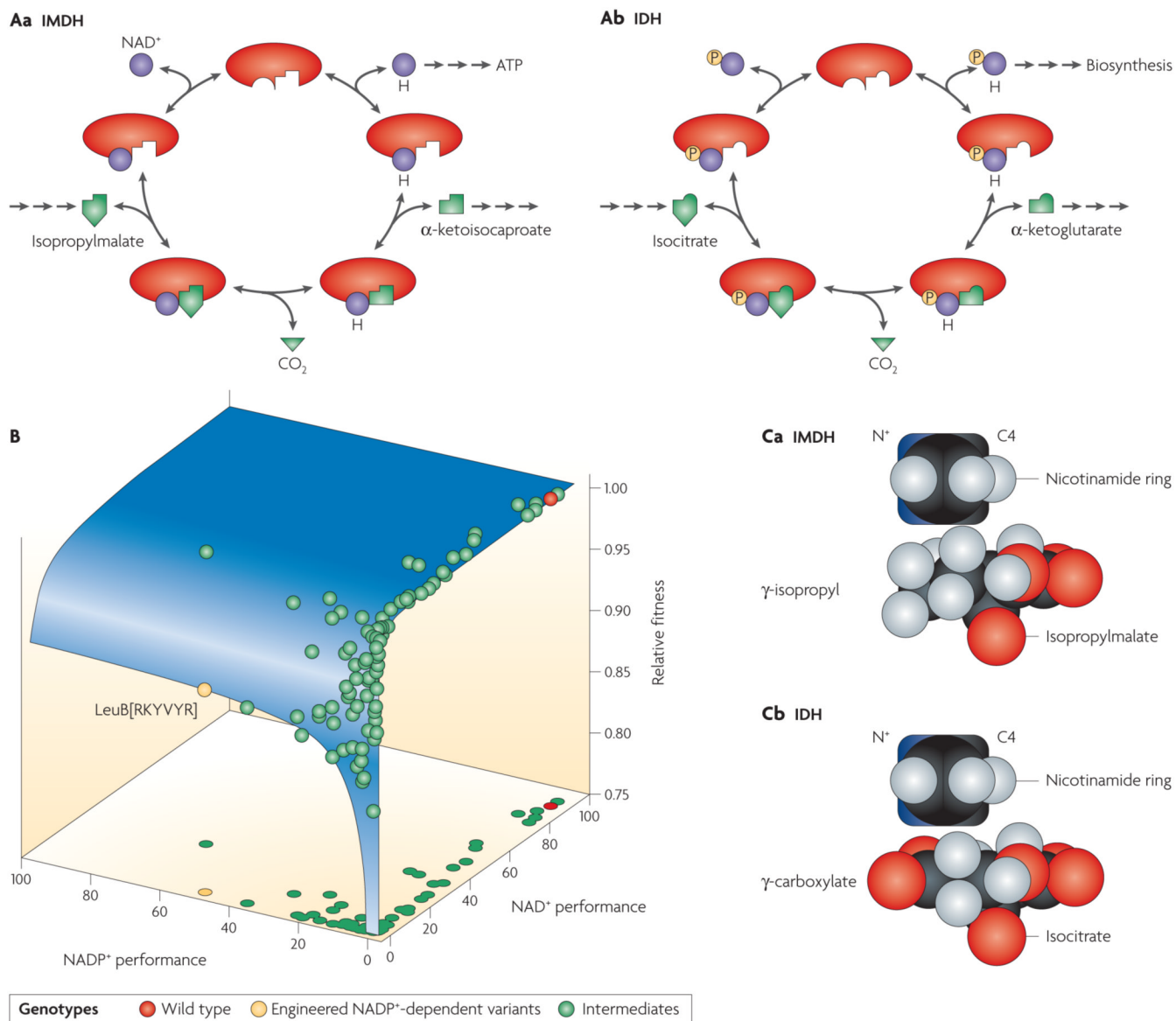
Figure 4. Constraint and opportunity in the evolution of coenzyme use by paralogous dehydrogenases

Aa | The catalytic cycle of isopropylmalate dehydrogenase (IMDH) to produce ATP in the Krebs cycle; ancestral use of NAD+ as a coenzyme has been conserved in all known IMDHs. **Ab** | The catalytic cycle of isocitrate dehydrogenase (IDH) to produce NADPH for biosynthesis; use of NADP+ as a coenzyme has evolved in some but not all IDH lineages. Binding of the reduced coenzyme — NADH to IMDH or NADPH to IDH — inhibits catalysis by tying up the enzyme in abortive complexes. Cellular concentrations of NADH are low, so inhibition of IMDH is insignificant. Although abundant, NADPH has a relatively low affinity for IDH. **B** | The adaptive landscape controlling the evolution of coenzyme use by IMDH. For 90 IMDH mutants, fitness is plotted against coenzyme performances. The fitted phenotype-fitness map is shown in blue. **C** | Structural basis for reduced NADPH inhibition in IDH. In both IDH and IMDH, the nicotinamide ring of the coenzyme lies atop the γ-moiety of the bound substrate (isopropylmalate or isocitrate). During catalysis, hydrid transfer to the C4 neutralizes the positive charge on the ring. In IMDH, the resulting hydrophobic attraction between the

reduced ring and the γ-isopropyl increases affinity. In IDH, the loss of the charge on the ring breaks the salt bridge to the γ-carboxylate and lowers affinity. Hence, cellular NADPH is a potent inhibitor of IMDH but a relatively weak inhibitor of IDH, allowing IDHs to evolve NADP$^+$ use. Parts **B** and **C** modified with permission from REF. [88] © (2006) American Association for the Advancement of Science.
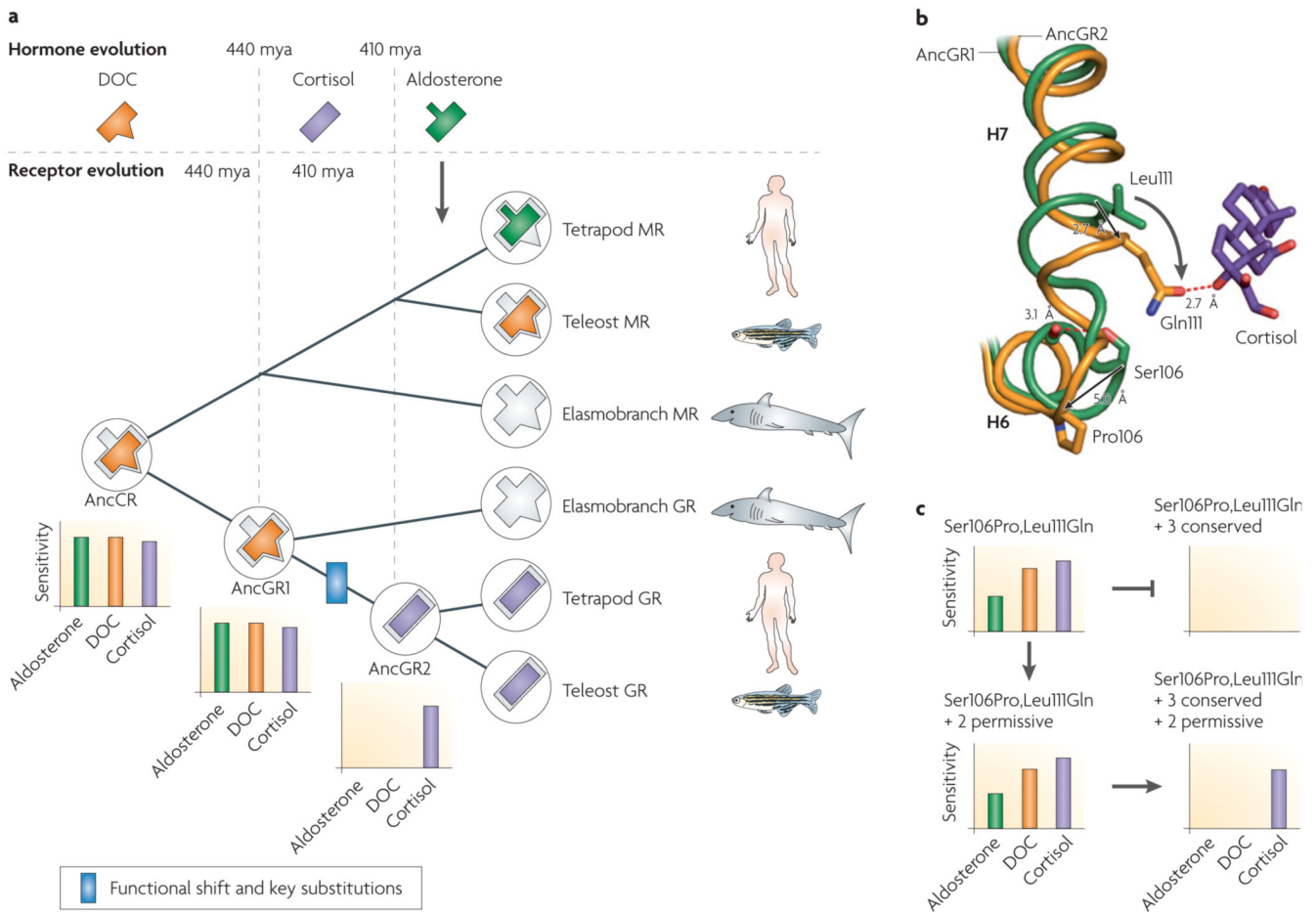
Figure 5. Evolution of corticosteroid-receptor specificity

a | Timeline for evolution of receptors for three structurally similar steroid hormones. Deoxycorticosterone (DOC) is an ancient vertebrate hormone, whereas aldosterone evolved much later in the tetrapod lineage (as indicated by a black arrow). Modern mineralocorticoid receptors (MRs) can be activated by aldosterone, DOC, and to a lesser extent cortisol. The glucocorticoid receptor (GR) is activated only by cortisol in bony vertebrates. The resurrected ancestral corticoid receptor (AncCR) has MR-like sensitivity to all three hormones. Resurrection of GRs from the ancestral jawed vertebrate (AncGR1) and the ancestral bony vertebrate (AncGR2) show that GR's cortisol specificity evolved in the interval between AncGR1 and AncGR2 (represented as a blue box). Dates from the fossil record are indicated in million of years ago (mya). **b** | Evolution of GR's cortisol specificity. When two historical replacements from the AncGR1-AncGR2 interval were introduced into the ancestral background, they switched receptor preference from aldosterone to cortisol. Structures of the ancestral proteins show that replacement Ser106Pro causes a kink in the protein backbone, destabilizing the ligand-receptor complex and reducing activation by all ligands; Ser106Pro also repositions site 111, so that replacement Leu111Gln forms hydrogen bonds with the C17-hydroxyl that is unique to cortisol. **c** | Optimization of the GR phenotype. Three other strict replacements from the same interval abolish function when introduced into AncGR1 with Ser106Pro and Leu111Gln — an unlikely evolutionary path under selection (as indicated by the inhibitory arrow). The ancestral structures identified two other replacements from this interval (out of 37 total) that in isolation have little effect on function, but when combined with the conserved substitutions yield a complete GR-like phenotype (as indicated by the black

arrow). Part **b** reproduced with permission from REF. [93] © (2007) American Association for the Advancment of Science.