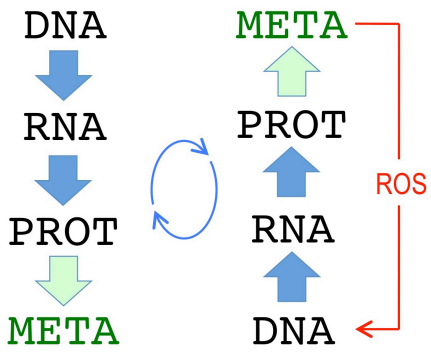
From the *selfish gene* to *selfish metabolism*: revisiting the central dogma

by

Víctor de Lorenzo

Systems & Synthetic Biology Program, Centro Nacional de Biotecnología CSIC Cantoblanco-Madrid, 28049 Spain



Running Title: Metabolism-driven adaptive evolution

Keywords: Central dogma, biodegradation, adaptive evolution, pathogenesis, dinitrotoluene, anti-fragility, metabolism-first

Abbreviations: CD: Central dogma; IVET: *In vivo* expression technology; ROS: reactive oxygen species; CF: cystic fibrosis

* Correspondence to: Víctor de Lorenzo
Centro Nacional de Biotecnología-CSIC
Campus de Cantoblanco, Madrid 28049, Spain
Tel.: 34- 91 585 45 36; Fax: 34-91 585 45 06
E-mail: vdlorenzo@cnb.csic.es

Abstract

The standard representation of the Central Dogma (CD) of Molecular Biology conspicuously ignores metabolism. However, both the metabolites and the biochemical fluxes behind any biological phenomenon are encrypted in the DNA sequence. Metabolism constrains and even changes the information flow when the DNA-encoded instructions conflict with the homeostasis of the biochemical network. Inspection of adaptive virulence programs and emergence of xenobiotic-biodegradation pathways in environmental bacteria suggest that their main evolutionary drive is the expansion of their metabolic networks towards new chemical landscapes rather than perpetuation and spreading of their DNA sequences. Faulty enzymatic reactions on suboptimal substrates produce reactive oxygen species (ROS), which fosters DNA diversification and eventually couples catabolism of the new chemicals to growth. All this calls for a revision of the CD in which metabolism (rather than DNA) has the leading role.

Introduction

How do biological systems adapt to changing environments and evolve to thrive in new physicochemical / nutritional niches? The textbook answer to this question is a combination of classical Darwinian theory and modern Molecular Biology [1]: DNA mutates randomly to bring about a phenotypic diversification of the biological object (from proteins to entire organisms) that explores the solution space until one variant gains a competitive advantage that enables it to reproduce better, thereby reaching a new fitness peak [2]. The implicit framework of the process is epitomized in the so-called central dogma (CD) of Molecular Biology (CD, Fig. 1a). The tenet *DNA makes RNA makes protein* [3,4] accounts for the flow of information in every biological system. Note that the term *information* applied to biological systems is often loosely defined, as it may refer either to that encoded in a sequence or that borne by a given context e.g. compositional or architectural information (see [5-8] for pro and con views). In any case, diversification of DNA at the top of the schema is considered the key occurrence for implementing an evolutionary solution-seeking drive, a job mostly done by the downstream proteins. Whether one considers the more general or the special version (i.e., taking into account RNA self-replication and retro-transcription back to DNA), the CD is based on two essential pillars: [i] that DNA, and RNA are the sole holders and transmitters of residue-by-residue information to proteins and [ii] that such information is basically unidirectional, with the occasional exception of back-copying RNA → DNA which, of itself,

makes no provision for information transfer from proteins back to nucleic acids. This state of affairs places DNA atop the hierarchy of actors that rule the functioning of living systems and orchestrate their adaptation to novel niches. Further inspection of the CD on the background of recent results on the way in which pathogens interact with their hosts and environmental bacteria biodegrade xenobiotic chemicals reveal, however, some inconsistencies in the scheme of Fig. 1a that are worth considering as detailed below.

The canonical CD ignores metabolism

The information stream of the canonical CD ends, astonishingly, in proteins. But … is that the actual finale of the flow? Is there nothing downstream of proteins? Where is metabolism? It is surprising that such a conspicuous absence has essentially been overlooked for so long. Many proteins are enzymes that transform chemicals into nutrients and metabolize such nutrients into all types of building blocks for living systems. The cognate catalytic properties of proteins are certainly encrypted in their primary amino acid sequence, itself encoded in DNA. However, the residue-by-residue transfer of information does not apply in this case, as one cannot predict chemical structures of metabolites or their fluxes directly from the genes involved. But metabolites, biochemical reactions and ensuing fluxes are, indeed, encoded – albeit in a more intricate way, in the DNA sequence. Computational de-convolution of such non-trivial information has allowed, for instance, the development of metabolic models from the sole source of genomic sequences [9-12]. On this basis, a necessary qualification of the CD involves, at the least, the explicit addition of one more layer of information transfer: protein → metabolism (Fig. 1b).

A separate issue regards the connectivity of the CD flowchart with the environment. The naked scheme of Fig. 1a would suggest that DNA, RNA and proteins are self-perpetuating objects with no links to their surroundings. Once more metabolism comes into play: it not only provides the building blocks for the other components of the CD, but also becomes the boundary between the biological and non-biological realms. This should also become clear in a streamlined representation of the dogma: the operation of the information flow depends on the availability of suitable metabolic precursors, cofactors and signals for synthesizing the canonical information-bearing polymers. The step from DNA to RNA (transcription) is checked by the archetypal metabolic “alarmone” ppGpp (an intracellular signalling molecule that is produced in response to harsh environmental conditions [13] and strongly influenced by somewhat

inconspicuous molecules such as α-ketoglutarate [14]. Starvation-associated stress is known to cause errors in DNA [15,16] and to alter translation – besides leading to oxidative stress due to deficiency in NADPH production by glucose-6-P dehydrogenase and other enzymes of the pentoses cycle [17]. Even the rather inexplicable distribution of CG content in bacterial genomes can be traced to the metabolic bottleneck that creates the transition G/C → A/T [18]. And so the examples continue.

A third problem of the official CD is the explicit hierarchy of the actors: DNA gives instructions to RNA, which gives instructions to proteins which (at least in the expanded plot of Fig. 1b) rule metabolism. This scheme (the DNA on top, the other performers below) embodies the view that DNA is the master component of any biological system, and that the others constitute a cohort of servants that execute the orders emanating from the top. One of the extreme interpretations that springs from the CD as a *chain of command* is the metaphor of the *selfish gene*. The concept, brilliantly elaborated by Richard Dawkins in his famous 1976 book, was opportunely published first at the onset of the recombinant DNA era [19]. The main claim of the title is that the sole agenda of genes (specifically, DNA sequences) is to attend to their only evolutionary interest: to continue being replicated -not necessarily for the benefit of the organism they are in. This means that each organism's body just provides the survival and dissemination machinery for the genes it contains. To be sure, inspection of many DNA segments - especially those encoding mobile elements (insertion sequences, transposons, plasmids etc.) - would suggest that they have self-propagation as their only biological program [20]. Typically, IS elements encode an enzyme, the only known function of which is to make its own DNA sequence hop between different locations in the genome [21]. Examination of the environmental “mobilome” has also revealed a high incidence of plasmids consisting of a circular DNA sequence encoding only a replication protein and an origin of replication for the corresponding sequences [22,23]. Inteins are one case of extreme and sophisticated DNA selfishness: intein-mediated protein splicing occurs after the intein-encoding sequence has been transcribed into mRNA and then translated into a polypeptide, thereby avoiding counterselection due to loss of protein function [24]. Evolutionary biologists make a subtle difference between *selfish genetic elements* or *parasitic genetic elements*, depending on whether they contribute more or less to the fitness of the host organism. The ultimate role of such jumping genes is intriguing, as game theory suggest that they go beyond apparent self-interest and may become ultimately advantageous to the population [25]. But can one take for granted that all DNA is (to various degrees) *selfish* just because some sequences look like it?

One cannot underestimate the influence that Dawkins' metaphor has had in subsequent generations of Life Sciences researchers, as it marked the commencement of what one could call the *DNA-centric* or *gene-centric* era of Molecular Biology, not to mention an equally gene-centric view of evolution that influences us to this day. And this is not just an academic enterprise: the *DNA-is-king* paradigm has direct biotechnological implications: control the DNA and you dominate the biological system it encodes. The dividends of such a notion are evident. To name but a few examples, moving genes around with the tools of genetic engineering (and more recently with straight DNA synthesis) has allowed *Escherichia coli* to produce antimalarial drug precursors [26] and even gasoline [27], yeasts to produce insulin [28], plants to produce plastics [29] and cows to produce therapeutic agents [30]. But the *DNA-rules* allegory implicit in the CD has spinoffs in other research endeavours that have been less rosy and which, in fact, challenge the tenet as a whole.

Changing DNA is insufficient for (re)programming biological systems

Let us take, for instance, metabolic engineering of environmental bacteria as agents for bioremediation of xenobiotic pollutants released by the chemical industry [31,32]. The *zeitgeist* of the mid-80s, at the peak of the recombinant DNA hype, was that one could gather genes from various origins (whether naturally-occurring or artificially improved e.g. in substrate specificity), assemble them in a rational way for a given biotechnological purpose, express them in a heterologous host and many (if not all) of mankind's problems would be gone [33]! But subsequent research soon watered down this somewhat naïve perspective. Genetic engineering from the late 80s to the early 2000s did produce a number of bacterial strains with enhanced biodegradative capacities. But the recombinant strains maintained their non-natural phenotypes only under the very controlled selective conditions of the laboratory [34]. Release of such Genetically Engineered Microorganisms (GEMs) in actual environmental niches on which the user had no or little control typically resulted in loss of the knocked-in phenotypes, and takeover by resident bacteria. Very few cases are reported in the literature in which GEMs proved to be better than naturally occurring microbes as bioremediation agents. The problem was not due to a dearth of genetic tools: a large number of genome editing strategies allows insertion of heterologous genes in the chromosome of bacterial hosts with a degree of stability that is virtually identical to that of natural counterparts [35,36]. Something similar happens with recombinant strains engineered for running

whole-cell catalysis in a bioreactor. Even if the recombinant strain is subject to a strong selective pressure (e.g. antibiotics added to the medium) for maintaining and expressing the genes of interest, the activity decreases over time in a sort of “catalytic fatigue” –a phenomenon well known to industrial biotechnologists [37].

More recently, synthetic biologists have created a large number of genetic circuits in which transcription factors and promoters are rationally re-connected following a man-made blueprint aimed at programming new-to-nature properties e. g. oscillators, toggle switches, light-sensing and many others [38,39]. But, again, it is now common knowledge that such devices operate for a limited period of time, after which they succumb to noise and mutations. One remarkable case is that of the T7 phage genome that was entirely re-factored to make it amenable to forward engineering by decompressing regulatory regions and arraying the various functional sequences in an ordered and standardized fashion [40]. Although the resulting virus was certainly infective, its subsequent evolution *in vivo* whilst progressing towards recovering the fitness level of the wild-type phage erased ~ 40% of man-made modifications [41]. In contrast, naturally occurring regulatory circuits are quite robust, and maintain their performance across time and space.

Taken together, all of the above suggest that changing the instructions (the DNA) is not sufficient for ensuring that the whole biological system duly obeys the orders coming from the top of the CD. Instead, it seems that the physiology of the host, of which metabolism is the key component, has a say in whether the directions from DNA are to be implemented or not. This becomes yet more manifest when the DNA-encoded instructions are aimed at changing the biochemical regime of cells: metabolism itself is most difficult to manipulate. A few examples illustrate this circumstance. Knocking-in phosphofructokinase of *E. coli* into *Pseudomonas putida* was unable to generate a glycolytic regime in this bacterium, which naturally consumes sugars through the Entner-Doudoroff pathway [42]. Attempts to alter genetically the biosynthetic flux of the pyrimidine pathway in *E. coli* could not defeat the homeostasis accomplished by the indigenous directed overflow metabolism [43]. Systematic study of the metabolic regimes of large collections of *E. coli* mutants [44,45] revealed only small variations in metabolite levels, indicating a rerouting of fluxes in the metabolic network that otherwise tolerates a large number of genetic changes. This reveals the extraordinary robustness of the cell biochemistry against perturbations coming from DNA, and somehow challenges its position at the summit of the CD

hierarchy. At the very least, one should concede that the operativity of the information/command chain from DNA to the downstream actors is constrained by metabolism, to the point that physiological imbalances may ruin it altogether (Fig. 1c). But there are more grounds yet for questioning the position of DNA as the leading character of the biological play.

Pathogenesis is not only about virulence determinants

Bacterial pathogenicity is a complex phenomenon in which given germs endowed with a number of advantageous traits penetrate animal or plant tissues to the detriment of the whole afflicted organism. The standard view is that virtually all bacteria with ecological compatibility with the host (for instance able to grow at body temperature) have the potential to become pathogens were it not for the permanent surveillance that the immune system imposes on any microbe that the body interacts with [46]. This explains how generally innocuous microorganisms that reside in diverse body habitats (gut, skin etc.) may become pathogenic in immune-compromised individuals. The boundary between pathogenic and non-pathogenic microbes is therefore a bit fuzzy, although some examples clearly map to the extremes. For instance, *Lactobacillus acidophilus* or *Bifidobacterium* are generally considered human-friendly, whereas species such as *Salmonella typhimurium, Shigella flexneri, Listeria monocytogenes* -not to mention enteroinvasive, enteropathogenic or enterohemorrhagic *E. coli* count among the most dreaded species [46]. What, then, makes a virulent strain virulent? Under the CD paradigm, the way to go for answering this question is the identification of genes (i.e. DNA) that determine such a quality.

In a series of papers from the mid-90s, J. Mekalanos [47] and D. Holden [48] proposed the first genetic tools for what would later be called *in vivo expression technology* (IVET [49,50]). The objective of that seminal work was to identify which genes of *Salmonella* were specifically turned on and off during experimental infection of laboratory mice. The procedures yielded lists of genes that were candidates for a role in pathogenesis. It is interesting that metabolic genes were largely ignored: only those that looked like accredited virulence or colonization factors were followed up. Increasingly sophisticated IVET methods have been used to search for such factors in a broad diversity of pathogens (see e.g. [51] and many others). Systematically, the resulting lists contain a large share of metabolic and transport genes and only a handful of specific-looking virulence factors (e.g. type-III / IV secretable effectors, adhesins,

haemolysin etc). The same is true for other pathogenesis-related phenomena e.g. formation of biofilms. Every survey has revealed most genetic actors of biofilm development to be metabolic: only a few others encode "attention-grabbing" functions [52]. But then comes the power of the *geno-centric* paradigm: metabolic functions recorded in every index of virulence culprits are largely ignored and (more or less deliberately) removed from view, as if metabolism had no business in the standard CD. Is it because metabolism is seen as boring or unimportant [53]? One way or the other, virtually all IVET projects end up studying the non-metabolic genes that participate in pathogenesis, while the biochemical players are often consigned to oblivion. But is this ultimately correct? In every conflict (and virulence is a clear case) the most ostentatious actors may not be the ones that are ultimately decisive. In fact, focusing on the most eye-catching components of a virulence program could be misleading if one disregards the metabolic and physico-chemical context in which the phenomenon occurs.

Virulence accompanies the conquest of new chemical landscapes

Metabolism is the economy of living systems. Very few things can happen if instructions coming from one or other gene antagonize the kinetic and thermodynamic optima of the cognate biochemical network of the living carrier. Quite on the contrary, the quest for such metabolic optima is the code that is recorded in the virulence program of pathogens and not *vice versa* (i.e. not DNA instructing the downstream components of the CD for its own sake). This is a serious argument for up-ending the chain of command of the CD as shown in Fig. 1c (not yet the handover of information, see below). In this way metabolism becomes the chief player, on top, and DNA is the clerk that duly archives instructions, based on past experiences, keeping the master satisfied. Following on with the same analogy, one may recall that the alphabet and the first written records were invented for the sake of keeping track of economic transactions and fixing rules for sorting conflicts out. Along the same lines, one could argue that DNA was *invented* as a repository of information to keep a record of already existing metabolic phenomena as well as a sort of memory for re-enacting biochemical reactions when needed. This is the stance of the *metabolism-first* view of the Origin of Life, which rejects the idea of naked self-replicating molecules (e.g. RNA), and suggests instead a seminal metabolism centred on primitive and growingly complex metabolic reactions [54-58] and cycles [59-61]. These could result in generating amino acids and nucleotides, which ultimately led to the emergence of RNA replication. Whether this theory is entirely correct or not, it seems clear that a complex network of (bio)chemical reactions was at some

point possible without genes, but genes or other replicative molecules are not possible without metabolism [62].

BOX 1. **Could metabolic networks be *selfish* evolutionary entities** able to autonomously accept and reject incoming traits? A precedent of this view is the *metabolon* or *structural-metabolic cellular complex*, as conceived by Srere [109]. In this case, reception or dismissal of new activities by such metabolic systems is exerted through e.g. adoption or silencing of particular genes. This means that metabolism can modify the genome, shaping its blueprint and granting inheritance of particular sets of encoded functions -and thus reversing the conventional directionality of the CD flowchart. According to S. J. Gould [1] there is a set of minimal criteria for defining an evolutionary object (*reproduction*, *inheritance*, *variation*, *interaction*). While the *selfish metabolism* proposed in this article has attributes of a true evolutionary entity, it is intertwined with the genome as the storehouse of information.

Considering biochemical networks (not DNA) as the master of a vast variety of biological occurrences could bring both conceptual and methodological changes in the way many phenomena observed in living systems are addressed - including pathogenesis. The *selfish gene* paradigm encourages us to focus on the specific DNA sequences that determine the most immediate virulence phenotypes. The alternative chemo-centric view implicates the drive of the pathogen's metabolic network in the *conquest of new plots of the chemical landscape* of the host for its own sake (Fig. 2). The *selfish gene* metaphor can thus justifiably be replaced by the alternative analogy of *selfish metabolism*. In the first case, the evolutionary agenda of pathogens is the preservation and spreading of the DNA born by the bacteria in question. In the other case, the driving force of virulence is the thermodynamic pull to integrate available substrates into one's own biochemical network. While the two agendas are perfectly compatible, there should be no doubt that chemistry and metabolism prevail over genetics.

Metabolism enables virulence

On the basis of the foregoing, what could, then, determine bacterial pathogenesis –if not the end-of-the-pipe: those conspicuous virulence factors? In some cases, the difference can be made by the possession of biochemical pathways that exploit as nutrients some of the proprietary metabolites of the host. Various examples can be found in the literature on pathogens that do exactly that. For instance, *Mycobacterium tuberculosis,* the cause of two million deaths per year, is able to survive for long periods of time within the host because virulent strains can also utilize the cholesterol of the target cells as a growth substrate [63,64]. As a consequence, deletion of the genes for metabolism of this steroid

decreases *M. tuberculosis* virulence (it should be kept in mind for the later discussion, that Mycobacteria count among the bacterial species that more frequently develop new metabolic routes for environmental xenobiotic and recalcitrant chemicals [65]). Along the same lines, it is known that scarcity of nicotinic acid, a precursor of the central cofactor NAD^+, elicits a large number of virulence traits in various pathogens. In a remarkable case, the lack of NAD in *Shigella* spp. due, for example, to mutation of the *nadB* gene (encoding L-aspartate oxidase) - which disrupts synthesis of quinolinate - is required for virulence. It is thus the lack of a metabolic precursor that triggers the pathogenesis program in this case [66,67]. Similarly, *Burkholderia* strains, which are most frequent inhabitants of cystic fribrosis (CF) patients' lungs need to catabolize phenylacetic acid as part of their virulence package [68,69]. A more sophisticated example is provided by the resident of the canine buccal cavity that causes septicaemia in humans, called *Capnocytophaga canimorsus*. This bacterium has the habit of feeding on mammalian cells by harvesting the glycan moiety of glycoproteins i.e. *shaving* the glycans away from the glycosylated proteins of the cell surface and using them as nutrients [70,71]. As a consequence, the virulence of pathogenic strains depends on the activity of feeding complexes displayed on the bacterial envelope that are reminiscent of the celullosomes of some environmental bacteria [72]. As expected, these complexes are essential for growth of the pathogen on cells, and for survival in mice.

But the metabolic drive of pathogens involves not only peripheral functions for wresting nutrients from the host (cholesterol, glycans, nicotinic acid etc) but it also engages central metabolic routes. And the connection is not just for the sake of feeding, but for enduring stress as well. Alas, the link of virulence with core metabolic transactions is more difficult to detect by just looking at the genes, as the primary sequences of enzymes for most central reactions are very similar, and the genes themselves quite conserved. Some recent observations, however, shed light on how some metabolic regimes contribute to virulence in various ways. It is well known that macrophages constitute one of the first lines of defence against infections - by phagocytosis of the bacteria followed by release reactive oxygen species (ROS) to kill the invaders [73]. The ability of enduring oxidative stress is linked to production of NADPH, the decisive metabolic currency that feeds the enzymes for countering ROS exposure. Interestingly, not all metabolic regimes are equally efficient at producing NADPH, the cycle of pentoses and the Entner-Doudoroff (ED) pathway for sugar consumption being better for this specific role that the standard Embden-Meyerhof-Parnas glycolysis [74]. In light of this, it cannot come as a surprise that the ED pathway has recently been shown to be essential for infection in the case of bacteria that have to cope

with oxidative stress during their intracellular lifestyle [75,76]. Other pathogens display somewhat odd host adaptation strategies that are yet more alien to the *selfish gene* metaphor. One intriguing case is that of the establishment of chronic infections of *Pseudomonas aeruginosa* in CF patients. Following years of antibiotic treatment and strong selective pressure, CF strains appear to enter a dead-end route characterized by the emergence of hypermutator phenotypes and pseudogenization of a good share of the genomic complement [77,78]. This seems to silence expression of most virulence factors that are customary for these pathogens. Moreover, once engaged in a chronic infection lifestyle, CF bacteria have a very slow growth rate, but still accompanied by activity of specific portions of the basal metabolism (in particular the glyoxylate shunt, [78-80]). In this case, the outcome of the process seems to be more directed to ensuring biochemical activity of the cells for as long as possible rather than spreading of their DNA (in fact, bacterial DNA of the terminal CF patients is not replicated or transmitted at all). Similarly, endosymbionts (e.g. *Buchnera*) appear to run a reduced but active metabolic network (they get most metabolites from the environment [81]) while they seem to have lost the ability to fix DNA damage [82]. Could these be yet other cases in which metabolism is the master and DNA the servant? If this were so, the quest for antimicrobials against infectious agents – whether acute or chronic, could be redirected to disabling central metabolic functions e.g. generation of NADPH or the glyoxylate cyle [83] instead of the habitual targeting of functions necessary for cell growth.

Virulence and biodegradation stem from the same evolutionary drive

Although strategies for feeding on unusual substrates and tolerating ROS can be considered virulence factors in the context of pathogenesis, the same drive that pushes agents of disease to conquer new plots of chemical landscape of the host also empowers environmental bacteria to explore ways of degrading unusual chemical species. Researchers have not failed to notice that routes for biodegradation of xenobiotic and recalcitrant chemicals are not infrequent in bacteria generally considered pathogenic [84]. Just to name a few, a considerable number of environmental isolates selected for catabolism of chloroaromatic and nitroaromatic compounds turn out to be strains of *Burkholderia* sp. a genus, the pathogenic members of which cause various diseases (e.g. glanders, melioidosis) and often aggravate pulmonary infections of CF patients (see above; [85]). Mycobacteria, including strains of *M. leprae*, are recurrently isolated in screenings of strains able to degrade polyaromatic hydrocarbons, a side product of the oil industry [65]. Even some *P. aeruginosa* strains can

grow on alkanes as the sole carbon source [86,87]. G. Church's team reported a few years ago that a surprising number of soil bacteria can not only resist, but also *metabolise* antibiotics as growth nutrients [88]. These observations suggest once more that the divide between pathogenic and non-pathogenic bacteria is more a human construct than a description of an authentically different biological program. Virulence and the ability to deal with unusual chemicals may stem from the same central metabolic drivers. Pathogens whose metabolism has evolved to counteract oxidative stress in macrophages might subsequently become optimal vehicles for hosting the evolution of biodegradation pathways for recalcitrant compounds, a process also associated with redox stress [89-91]. By the same token, there is a continuum of growing toxicity in the chemical structures that bacteria meet during their lifetime: from *bona fide* nutrients, to recalcitrant compounds to xenobiotics and antibiotics. The indigenous metabolic network of bacteria, whether virulent or not, tends to expand its reach towards the chemical landscapes of the surrounding environment. As sketched in Fig. 2, the evolutionary process involves the progressive incorporation of abiotic and xenobiotic compounds to the pool of molecules that are recognized as actual or potential nutrients. The endogenous and exogenous stress that is inherent to the intermediate stages can be mitigated –both during virulence and biodegradation - by the redox metabolic currency (e.g. NADPH) produced by core metabolic pathways. This takes us to the last point of this article: how do existing biochemical networks explore the solution space for metabolizing a compound that they may never have seen before, and how is such a solution ultimately recorded in DNA?

ROS enables a feedback loop to the CD

The process by which given bacteria evolve to degrade a new xenobiotic chemical is not trivial. While resistance to antibiotics can be endowed by either simple mutations in the target or horizontal genetic transfer of enzymes that inactivate the antimicrobial, the emergence of a new metabolic pathway is a different, stepwise, and more complicated matter. Pre-existing enzymes have to acquire new substrate specificities, and the shape and kinetics of the enzymes involved fit to each other to avoid the escape of intermediates [84,92]. In addition, the flow of the resulting products has to be wired to the extant biochemical network in order to generate a growth advantage that allows its selection. This typically requires multiple changes in both the catabolic genes at stake and in the rest of the cell physiology. Finally, such evolution happens in extremely competitive environmental niches inhabited by many other

microorganisms, meaning that the process cannot afford much loss of fitness during the intermediate steps. Environmental bacteria that partially degrade 2,4-dinitrotoluene (DNT) have been isolated which provide an exceptional experimental system to examine the transition between a precursor pathway for catabolism of the naturally-occurring hydrocarbon naphthalene and the new xenobiotic substrate DNT [93,94]. Recent data on this system [89] have revealed an intriguing mechanism by which the first enzyme of the biodegradation pathway (a Rieske-type dioxygenase) that acts on the substrate-to-be (DNT) may uncouple the reaction because of poor specificity, and generate a high level of endogenous reactive oxygen species (ROS). This, in turn, damages DNA (and possibly DNA repair proteins as well [95]) and increases mutagenesis, ultimately resulting in the creation of novelty that fosters evolution of xenobiotic-degrading variants of the strain hosting the pathway (Fig. 3). The metabolic problem thus stimulates the exploration of the solution space, as the stress caused by faulty dioxygenation of DNT accelerates the rate of bacterial change. It is plausible that such an evolutionary scenario is widespread, because many other oxygenation reactions are known to withstand a degree of ROS production that depends on the more or less obligate coupling of the substrate to the redox enzyme [96-99].

BOX2. **Oxidative stress and emergence of antibiotic resistance**. It has been proposed that, following disruption of basic cellular functions, virtually all antibiotics kill bacteria by triggering intracellular production of ROS [110,111]. Similarly to the scenarios entertained in this article, such ROS could in turn mutate DNA and generate genetic and phenotypic novelty leading to the rise of resistance towards the same antibiotic that causes the problem –or other antimicrobials [100,101]. In this context, counteracting oxidative stress by stimulating NADPH-producing metabolic pathways could be a clear fitness trait of pathogens during infection regardless of the origin of ROS: externally in the case of phagocytosis by macrophages or internally upon antibiotic treatment. However, whether this is a general principle, and whether all antibiotic stresses are ultimately converted into lethal ROS levels, are questions still debated at the time of writing this review [103,112-115].

The picture of metabolically-fostered ROS-driven hypermutation and ensuing genetic variation is reminiscent -but different - to that proposed for the evolution of antibiotic resistance. At least in *E. coli*, this has been suggested also to be stimulated by mutagenic stress in cells treated with antimicrobials [100,101]. In contrast, in the DNT biodegradation case (and plausibly in other metabolic reactions of that sort) the trigger of the adaptive process can be traced to the master oxygenation that initiates the pathway. The gain of a new metabolic ability is essentially different from resisting an antibiotic: in the *selfish metabolism* paradigm the biological system does not merely escape a lethal pressure: it *conquers* another portion of the chemical landscape i.e. it pushes further into the thermodynamic

interface between chemistry and biology –and for its own sake (Fig. 2).

The transient regime experienced by bacteria under ROS-producing metabolic stress is in fact worth examining in detail. On the one hand, it is possible that while cells accumulate both adverse changes (surely most of them) and beneficial (few) modifications, the fitness of the lineages is buffered by overproduction of chaperones. This circumstance that has been observed in *E. coli* when artificially subject to a high-mutation rate [102] may well happen in natural settings. On the other hand, ROS alters protein structure (thereby changing enzymatic activities as well [95]) and may also cause mistranslation due (among other things) to the incorporation of 8-oxo-guanine into newly synthesized RNAs [103]. This means that the biochemical diversity of cells exposed to ROS might, in reality, be broader than the inherited genetic diversity of the progeny. This opens intriguing evolutionary scenarios that are beyond the scope of this article. Since oxygen appeared on the Biosphere > 1 billion years after the emergence of the first bacteria [104] it is evident that evolution does occur also in its absence (e.g. some archetypal pathogens are strict anaerobes). Still, it seems clear that a suite of metabolic reactions, specially those producing ROS can accelerate the process directly or indirectly.

Conclusion -outlook

The gene-centric view of evolution that has dominated biological thought for a number of decades has its more radical expression in the *selfish gene* concept, which is itself an extreme spinoff of the CD. But new data stemming from fields as remote as bacterial pathogenesis, biodegradation and Origin of Life, as well as new methods for studying the metabolome and the fluxome force us to qualify some of the implicit claims of this most influential of biological metaphors. **First,** metabolism must be included in the information flow. Information travelling between RNA, DNA and protein is easy to visualize. In contrast, deconvoluting protein sequences into production of chemicals and fluxes is more difficult –but the information is indeed there. **Second,** the biochemical network has its own autonomous drive, which constraints the execution of instructions from DNA more than *vice versa.* This calls for a conceptual change in which metabolism has the leading position in the chain of command [105] -which is different from the customary direction of the information flow (e.g. Fig. 1a). And, **third,** the CD is not exclusively a one-way stream of instructions. Faulty enzymatic steps on suboptimal substrates (in particular oxygenations and other redox reactions) can produce ROS, which feed back on DNA and thereby

create biological novelty. While the informational relationships *DNA* → *metabolism* are fairly strict (though not exclusively codified), the feedback *metabolism* → *DNA* has a different quality i.e. a sort of *mandate* for exploring the solution space but without specifying any outcome (Fig. 4). In reality, all DNA sequences are subject to spontaneous changes which, given enough time, can make the coding sequence of existing enzymes to change substrate specificity and contribute to what has been called paralogous metabolism [106]. Metabolic stress in general and ROS-mediated mutagenesis originated in defective biochemical reactions in particular seems thus to enable a wider creation of novelty in a much faster fashion. This surely gives an advantage to the bacterial carrier in the very competitive environmental scenarios were microorganisms live. It is plausible that such metabolic stress-driven evolution of new traits accounts for a plethora of rapid adaptation phenomena that cannot be explained by the mere genetic drift of the DNA sequences at stake [107,108].

Acknowledgements

VdL is indebted to D. Pérez-Pantoja, F. Rojo, P. Nikel, A. Danchin, J. Blazquez, E. Martínez, A. Goñi, D. Schneider and F. Baquero for very useful comments to the manuscript. The work in Author's laboratory is supported by generous grants of the Spanish Ministry of Science and Innovation (CONSOLIDER), by contracts of the Framework Program of the EU (MICROME, ST-FLOW), the European Research Council (ARISYS) and by Funds of the Autonomous Community of Madrid (PROMT Program). Author declares no conflict of interest.

References

1. **Gould SJ.** 2002. *The structure of evolutionary theory.* Cambridge ME: Harvard University Press.
2. **Kussell E.** 2013. Evolution in microbes. *Annu Rev Biophys* **42**: 493-514.
3. **Crick F.** 1958. On Protein Synthesis. *Symp Soc Exp Biol* **12**: 139–63.
4. **Crick F.** 1970. Central dogma of molecular biology. *Nature* **227**: 561-3.
5. **Binder PM, Danchin A.** 2011. Life's demons: information and order in biology. What subcellular machines gather and process the information necessary to sustain life? *EMBO Rep* **12**: 495-9.
6. **Segre D, Shenhav B, Kafri R, Lancet D.** 2001. The molecular roots of compositional inheritance. *J Theor Biol* **213**: 481-91.

7. **Root-Bernstein R.** 2012. A modular hierarchy-based theory of the chemical origins of life based on molecular complementarity. *Acc Chem Res* **45**: 2169-77.
8. **Vasas V, Szathmary E, Santos M.** 2010. Lack of evolvability in self-sustaining autocatalytic networks constraints metabolism-first scenarios for the origin of life. *Proc Ntal Acad Sci USA* **107**: 1470-5.
9. **Feist AM, Palsson BO.** 2008. The growing scope of applications of genome-scale metabolic reconstructions using *Escherichia coli*. *Nat Biotechnol* **26**: 659-67.
10. **McCloskey D, Palsson BO, Feist AM.** 2013. Basic and applied uses of genome-scale metabolic network reconstructions of *Escherichia coli. Mol Sys Biol* **9**: 661.
11. **Mo ML, Palsson BO.** 2009. Understanding human metabolic physiology: a genome-to-systems approach. *Trends Biotechnol* **27**: 37-44.
12. **Henry CS, DeJongh M, Best AA, Frybarger PM, et al.** 2010. High-throughput generation, optimization and analysis of genome-scale metabolic models. *Nat Biotechnol* **28**: 977-82.
13. **Potrykus K, Cashel M.** 2008. (p)ppGpp: still magical? *Ann Rev Microbiol* **62**: 35-51.
14. **You C, Okano H, Hui S, Zhang Z, et al.** 2013. Coordination of bacterial proteome with metabolism by cyclic AMP signalling. *Nature* **500**: 301-6.
15. **Bjedov I, Tenaillon O, Gérard B, Souza V, et al.** 2003. Stress-induced mutagenesis in bacteria. *Science* **300**: 1404-9.
16. **Foster PL.** 2007. Stress-induced mutagenesis in acteria. *Crit Rev Biochem Mol Biol* **42**: 373-97.
17. **Pandolfi PP, Sonati F, Rivi R, Mason P, et al.** 1995. Targeted disruption of the housekeeping gene encoding glucose 6-phosphate dehydrogenase (G6PD): G6PD is dispensable for pentose synthesis but essential for defense against oxidative stress. *EMBO J* **14**: 5209-15.
18. **Rocha EP, Danchin A.** 2002. Base composition bias might result from competition for metabolic resources. *Trends Genet* **18**: 291-4.
19. **Dawkins R.** 1976. *The selfish gene*. Oxford, UK: Oxford University Press.
20. **Werren JH, Nur U, Wu CI.** 1988. Selfish genetic elements. *Trends Ecol Evol* **3**: 297-302.
21. **Mahillon J, Chandler M.** 1998. Insertion sequences. *Microbiol Mol Biol Revs* **62**: 725-74.
22. **Sentchilo V, Mayer AP, Guy L, Miyazaki R, et al.** 2013. Community-wide plasmid gene mobilization and selection. *ISME J* **7**: 1173-86.

23. **Li LL, Norman A, Hansen LH, Sorensen SJ.** 2012. Metamobilomics -expanding our knowledge on the pool of plasmid encoded traits in natural environments using high-throughput sequencing. *Clin Microbiol Infect* **18 Suppl 4**: 5-7.
24. **Gogarten JP, Senejani AG, Zhaxybayeva O, Olendzenski L, et al.** 2002. Inteins: structure, function, and evolution. *Ann Rev Microbiol* **56**: 263-87.
25. **Gardner A, Welch JJ.** 2011. A formal theory of the selfish gene. *J Evol Biology* **24**: 1801-13.
26. **Dietrich JA, Yoshikuni Y, Fisher KJ, Woolard FX, et al.** 2009. A novel semi-biosynthetic route for artemisinin production using engineered substrate-promiscuous P450(BM3). *ACS Chem Biol* **4**: 261-7.
27. **Choi YJ, Lee SY.** 2013. Microbial production of short-chain alkanes. *Nature* **502:** 571-4.
28. **Kiefer MC, Schmid C, Waldvogel M, Schläpfer I, et al.** 1992. Characterization of recombinant human insulin-like growth factor binding proteins 4, 5, and 6 produced in yeast. *J Biol Chem* **267**: 12692-9.
29. **Bohmert-Tatarev K, McAvoy S, Daughtry S, Peoples OP, et al.** 2011. High levels of bioplastic are produced in fertile transplastomic tobacco plants engineered with a synthetic operon for the production of polyhydroxybutyrate. *Plant Physiol* **155**: 1690-708.
30. **van Berkel PHC, Welling MM, Geerts M, van Veen HA, et al.** 2002. Large scale production of recombinant human lactoferrin in the milk of transgenic cows. *Nat Biotech* **20**: 484-7.
31. **Timmis KN, Steffan RJ, Unterman R.** 1994. Designing microorganisms for the treatment of toxic wastes. *Ann Rev Microbiol* **48**: 525-57.
32. **Pieper DH, Reineke W.** 2000. Engineering bacteria for bioremediation. *Curr Op Biotech* **11**: 262-70.
33. **Lindow SE, Panopoulos NJ, McFarland BL.** 1989. Genetic engineering of bacteria from managed and natural habitats. *Science* **244**: 1300-7.
34. **Cases I, de Lorenzo V.** 2005. Genetically modified organisms for the environment: stories of success and failure and what we have learned from them. *Int Microbiology* **8**: 213-22.
35. **de Lorenzo V, Herrero M, Sánchez JM, Timmis KN.** 1998. Mini-transposons in microbial ecology and environmental biotechnology. *FEMS Microbiol Ecol* **27**: 211-24.
36. **Martinez-Garcia E, de Lorenzo V.** 2012. Transposon-based and plasmid-based genetic tools for editing genomes of Gram-negative bacteria. *Meth Mol Biol* **813**: 267-83.

37. **Kumar PKR, Maschke HE, Friehs K, Schügerl K.** 1991. Strategies for improving plasmid stability in genetically modified bacteria in bioreactors. *Trends Biotech* **9**: 279-84.
38. **Purnick PEM, Weiss R.** 2009. The second wave of synthetic biology: from modules to systems. *Nat Rev Mol Cell Biol* **10**: 410-22.
39. **Slusarczyk AL, Lin A, Weiss R.** 2012. Foundations for the design and implementation of synthetic genetic circuits. *Nat Rev Genet* **13**: 406-20.
40. **Chan LY, Kosuri S, Endy D.** 2005. Refactoring bacteriophage T7. *Mol Sys Biol* **1**.
41. **Springman R, Molineux IJ, Duong C, Bull RJ, et al.** 2012. Evolutionary stability of a refactored phage genome. *ACS Synth Biol* **1**: 425-30.
42. **Chavarria M, Nikel PI, Perez-Pantoja D, de Lorenzo V.** 2013. The Entner-Doudoroff pathway empowers *Pseudomonas putida* KT2440 with a high tolerance to oxidative stress. *Env Microbiol* **15**: 1772-85.
43. **Reaves ML, Young BD, Hosios AM, Xu YF, et al.** 2013. Pyrimidine homeostasis is accomplished by directed overflow metabolism. *Nature* **500**: 237-41.
44. **Ishii N, Nakahigashi K, Baba T, Robert M, et al.** 2007. Multiple high-throughput analyses monitor the response of *E. coli* to perturbations. *Science* **316**: 593-7.
45. **Nakahigashi K, Toya Y, Ishii N, Soga T, et al.** 2009. Systematic phenome analysis of *Escherichia coli* multiple-knockout mutants reveals hidden reactions in central carbon metabolism. *Mol Sys Biol* **5**: 306.
46. **Fauci AS, Morens DM.** 2012. The perpetual challenge of infectious diseases. *New Engl J Med* **366**: 454-61.
47. **Mahan MJ, Slauch JM, Mekalanos JJ.** 1993. Selection of bacterial virulence genes that are specifically induced in host tissues. *Science* **259**: 686-8.
48. **Hensel M, Shea JE, Gleeson C, Jones MD, et al.** 1995. Simultaneous identification of bacterial virulence genes by negative selection. *Science* **269**: 400-3.
49. **Chiang SL, Mekalanos JJ, Holden DW.** 1999. In vivo genetic analysis of bacterial virulence. *Ann Rev Microbiol* **53**: 129-54.
50. **Lehoux DE, Levesque RC.** 2000. Detection of genes essential in specific niches by signature-tagged mutagenesis. *Curr Op Biotech* **11**: 434-9.
51. **Handfield M, Lehoux DE, Sanschagrin F, Mahan MJ, et al.** 2000. In vivo-induced genes in *Pseudomonas aeruginosa*. *Infect Imm* **68**: 2359-62.

52. **O'Toole GA, Kolter R.** 1998. Initiation of biofilm formation in *Pseudomonas fluorescens* WCS365 proceeds via multiple, convergent signalling pathways: a genetic analysis. *Mol Microbiol* **28**: 449-61.
53. **Ray LB.** 2010. Metabolism is not boring: Introduction. *Science* **330**: 1337.
54. **Stueken EE, Anderson RE, Bowman JS, Brazelton WJ, et al.** 2013. Did life originate from a global chemical reactor? *Geobiology* **11**: 101-26.
55. **Huber C, Kraus F, Hanzlik M, Eisenreich W, et al.** 2012. Elements of metabolic evolution. *Chemistry* **18**: 2063-80.
56. **Martin W, Russell MJ.** 2003. On the origins of cells: a hypothesis for the evolutionary transitions from abiotic geochemistry to chemoautotrophic prokaryotes, and from prokaryotes to nucleated cells. *Phil Trans Roy Soc B* **358**: 59-83
57. **Sousa FL, Thiergart T, Landan G, Nelson-Sathi S, et al.** 2013. Early bioenergetic evolution. *Phil Trans Roy Soc B* **368**: 20130088.
58. **Fuchs G.** 2011. Alternative pathways of carbon dioxide fixation: insights into the early evolution of life? *Ann Rev Microbiol* **65**: 631-58.
59. **Melendez-Hevia E, Waddell TG, Cascante M.** 1996. The puzzle of the Krebs citric acid cycle: assembling the pieces of chemically feasible reactions, and opportunism in the design of metabolic pathways during evolution. *J Mol Evol* **43**: 293-303.
60. **Smith E, Morowitz HJ.** 2004. Universality in intermediary metabolism. *Proc Natl Acad Sci USA* **101**: 13168-73.
61. **Braakman R, Smith E.** 2012. The emergence and early evolution of biological carbon-fixation. *PLoS Comp Biol* **8**: e1002455.
62. **Dyson FJ.** 1982. A model for the origin of life.*J Mol Evol* **18**: 344-50.
63. **Yam KC, D'Angelo I, Kalscheuer R, Zhu H, et al.** 2009. Studies of a ring-cleaving dioxygenase illuminate the role of cholesterol metabolism in the pathogenesis of *Mycobacterium tuberculosis. PLoS Pathog* **5**: e1000344.
64. **Yam KC, Okamoto S, Roberts JN, Eltis LD.** 2011. Adventures in *Rhodococcus* - from steroids to explosives. *Can J Microbiol* **57**: 155-68.
65. **Cerniglia CE.** 1993. Biodegradation of polycyclic aromatic hydrocarbons. *Curr Op Biotech* **4**: 331-8.

66. **Di Martino ML, Fioravanti R, Barbabella G, Prosseda G, et al.** 2013. Molecular evolution of the nicotinic acid requirement within the *Shigella/*EIEC pathotype.*Int J Med Microbiol* **303:** 651-61.

67. **Li Z, Bouckaert J, Deboeck F, De Greve H, et al.** 2012. Nicotinamide dependence of uropathogenic *Escherichia coli* UTI89 and application of *nadB* as a neutral insertion site. *Microbiology* **158**: 736-45.

68. **Law RJ, Hamlin JN, Sivro A, McCorrister SJ, et al.** 2008. A functional phenylacetic acid catabolic pathway is required for full pathogenicity of *Burkholderia cenocepacia* in the *Caenorhabditis elegans* host model. *J Bacteriol* **190**: 7209-18.

69. **Yudistira H, McClarty L, Bloodworth RA, Hammond SA, et al.** 2011. Phenylalanine induces *Burkholderia cenocepacia* phenylacetic acid catabolism through degradation to phenylacetyl-CoA in synthetic cystic fibrosis sputum medium. *Microb Pathog* **51**: 186-93.

70. **Mally M, Shin H, Paroz C, Landmann R, et al.** 2008. *Capnocytophaga canimorsus:* a human pathogen feeding at the surface of epithelial cells and phagocytes. *PLoS Pathog* **4**: e1000164.

71. **Manfredi P, Renzi F, Mally M, Sauteur L, et al.** 2011. The genome and surface proteome of *Capnocytophaga canimorsus* reveal a key role of glycan foraging systems in host glycoproteins deglycosylation. *Mol Microbiol* **81**: 1050-60.

72. **Doi RH, Kosugi A.** 2004. Cellulosomes: plant-cell-wall-degrading enzyme complexes. *Nat Rev Micro* **2**: 541-51.

73. **Yuk J-M, Yoshimori T, Jo E-K.** 2012. Autophagy and bacterial infectious diseases. *Exp Mol Med* **44**: 99-108.

74. **Breitling R, Vitkup D, Barrett MP.** 2008. New surveyor tools for charting microbial metabolic maps. *Nat Rev Micro* **6**: 156-61.

75. **Harada E, Iida K, Shiota S, Nakayama H, et al.** 2010. Glucose metabolism in *Legionella pneumophila:* dependence on the Entner-Doudoroff pathway and connection with intracellular bacterial growth. *J Bacteriol* **192**: 2892-9.

76. **Patra T, Koley H, Ramamurthy T, Ghose AC, et al.** 2012. The Entner-Doudoroff pathway is obligatory for gluconate utilization and contributes to the pathogenicity of *Vibrio cholerae*. *J Bacteriol* **194**: 3377-85.

77. **Chung JC, Becq J, Fraser L, Schulz-Trieglaff O, et al.** 2012. Genomic variation among contemporary *Pseudomonas aeruginosa* isolates from chronically infected cystic fibrosis patients. *J Bacteriol* **194**: 4857-66.

78. **Marvig RL, Johansen HK, Molin S, Jelsbak L.** 2013. Genome analysis of a transmissible lineage of *Pseudomonas aeruginosa* reveals pathoadaptive mutations and distinct evolutionary paths of hypermutators. *PLoS Genet* **9**: e1003741.

79. **Damkiaer S, Yang L, Molin S, Jelsbak L.** 2013. Evolutionary remodeling of global regulatory networks during long-term bacterial adaptation to human hosts. *Proc Natl Acad Sci USA* **110**: 7766-71.

80. **Thogersen JC, Morup M, Damkiaer S, Molin S, et al.** 2013. Archetypal analysis of diverse *Pseudomonas aeruginosa* transcriptomes reveals adaptation in cystic fibrosis airways. *BMC Bioinfom* **14**: 279.

81. **Pál C, Papp B, Lercher MJ, Csermely P, et al.** 2006. Chance and necessity in the evolution of minimal metabolic networks. *Nature* **440**: 667-70.

82. **Shigenobu S, Watanabe H, Hattori M, Sakaki Y, et al.** 2000. Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera* sp. APS. *Nature* **407**: 81-6.

83. **Fahnoe KC, Flanagan ME, Gibson G, Shanmugasundaram V, et al.** 2012. Non-traditional antibacterial screening approaches for the identification of novel inhibitors of the glyoxylate shunt in Gram-negative pathogens. *PloS One* **7**: e51732.

84. **Fuchs G, Boll M, Heider J.** 2011. Microbial degradation of aromatic compounds - from one strategy to four. *Nat Rev Microbiol* **9**: 803-16.

85. **Perez-Pantoja D, Donoso R, Agullo L, Cordova M, et al.** 2012. Genomic analysis of the potential for aromatic compounds biodegradation in Burkholderiales. *Env Microbiol* **14**: 1091-117.

86. **Alonso A, Rojo F, Martinez JL.** 1999. Environmental and clinical isolates of *Pseudomonas aeruginosa* show pathogenic and biodegradative properties irrespective of their origin. *Env Microbiol* **1**: 421-30.

87. **Marin MM, Yuste L, Rojo F.** 2003. Differential expression of the components of the two alkane hydroxylases from *Pseudomonas aeruginosa. J Bacteriol* **185**: 3232-7.

88. **Dantas G, Sommer MO, Oluwasegun RD, Church GM.** 2008. Bacteria subsisting on antibiotics. *Science* **320**: 100-3.

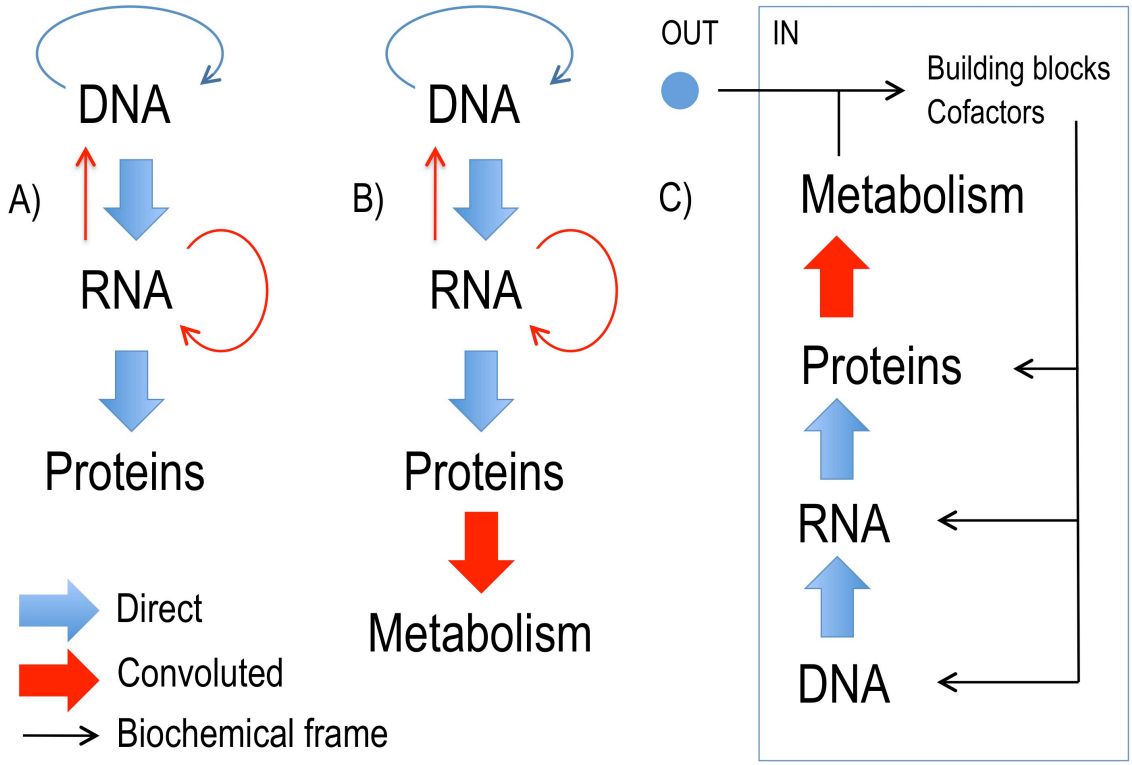
89. **Perez-Pantoja D, Nikel PI, Chavarria M, de Lorenzo V.** 2013. Endogenous stress caused by faulty oxidation reactions fosters evolution of 2,4-dinitrotoluene-degrading bacteria. *PLoS genetics* **9**: e1003764.

90. **Agullo L, Camara B, Martinez P, Latorre V, et al.** 2007. Response to (chloro)biphenyls of the polychlorobiphenyl-degrader *Burkholderia xenovorans* LB400 involves stress proteins also induced by heat shock and oxidative stress. *FEMS Microbiol Revs* **267**: 167-75.
91. **Martinez P, Agullo L, Hernandez M, Seeger M.** 2007. Chlorobenzoate inhibits growth and induces stress proteins in the PCB-degrading bacterium *Burkholderia xenovorans* LB400. *Arch Microbiol* **188**: 289-97.
92. **Pérez-Pantoja D, González B, Pieper DH.** 2010. Aerobic degradation of aromatic hydrocarbons. In Timmis K, ed. *Handbook of Hydrocarbon and Lipid Microbiology*: Springer Berlin Heidelberg. p 799-837.
93. **Parales RE.** 2010. Hydrocarbon degradation by betaproteobacteria. In Timmis K, ed; *Handbook of Hydrocarbon and Lipid Microbiology*: Springer Berlin Heidelberg. p 1715-24.
94. **de Las Heras A, Chavarria M, de Lorenzo V.** 2011. Association of *dnt* genes of *Burkholderia* sp. DNT with the substrate-blind regulator DntR draws the evolutionary itinerary of 2,4-dinitrotoluene biodegradation. *Mol Microbiol* **82**: 287-99.
95. **Krisko A, Radman M.** 2013. Phenotypic and genetic consequences of protein damage. *PLoS Genet* **9**: e1003810.
96. **Lee K.** 1999. Benzene-induced uncoupling of naphthalene dioxygenase activity and enzyme inactivation by production of hydrogen peroxide. *J Bacteriol* **181**: 2719-25.
97. **Imbeault NY, Powlowski JB, Colbert CL, Bolin JT, et al.** 2000. Steady-state kinetic characterization and crystallization of a polychlorinated biphenyl-transforming dioxygenase. *J Biol Chem* **275**: 12430-7.
98. **Hlouchova K, Rudolph J, Pietari JM, Behlen LS, et al.** 2012. Pentachlorophenol hydroxylase, a poorly functioning enzyme required for degradation of pentachlorophenol by *Sphingobium chlorophenolicum*. *Biochemistry* **51**: 3848-60.
99. **Ravindra Kumar S, Imlay JA.** 2013. How *Escherichia coli* tolerates profuse hydrogen peroxide formation by a catabolic pathway. *J Bacteriol* **195**: 4569-79.
100. **Blazquez J, Couce A, Rodriguez-Beltran J, Rodriguez-Rojas A.** 2012. Antimicrobials as promoters of genetic variation. *Curr Opin Microbiol* **15**: 561-9.
101. **Couce A, Blazquez J.** 2009. Side effects of antibiotics on genetic variability. *FEMS Microbiol Rev* **33**: 531-8.

102. **Maisnier-Patin S, Roth JR, Fredriksson A, Nystrom T, et al.** 2005. Genomic buffering mitigates the effects of deleterious mutations in bacteria. *Nature Genet* **37**: 1376-9.
103. **Foti JJ, Devadoss B, Winkler JA, Collins JJ, et al.** 2012. Oxidation of the guanine nucleotide pool underlies cell death by bactericidal antibiotics. *Science* **336**: 315-9.
104. **Johnson JE, Webb SM, Thomas K, Ono S, et al.** 2013. Manganese-oxidizing photosynthesis before the rise of cyanobacteria. *Proc Natl Acad Sci USA* **110**: 11238-43.
105. **Berthoumieux S, de Jong H, Baptist G, Pinel C, et al.** 2013. Shared control of gene expression in bacteria by transcription factors and global physiology of the cell. *Mol Sys Biol* **9**: 634.
106. **Chan CM, Danchin A, Marliere P, Sekowska A.** 2013. Paralogous metabolism: S-alkyl-cysteine degradation in *Bacillus subtilis*. *Env Microbiol* (doi: 10.1111/1462-2920.12210)
107. **Hindre T, Knibbe C, Beslon G, Schneider D.** 2012. New insights into bacterial adaptation through in vivo and *in silico* experimental evolution. *Nature Rev Microbiol* **10**: 352-65.
108. **Martincorena I, Luscombe NM.** 2013. Non-random mutation: the evolution of targeted hypermutation and hypomutation. *Bioessays* **35**: 123-30.
109. **Srere PA.** 1985. The metabolon. *Trends in Biochemical Sciences* **10**: 109-10.
110. **Kohanski MA, Dwyer DJ, Collins JJ.** 2010. How antibiotics kill bacteria: from targets to networks. *Nat Rev Microbiol* **8**: 423-35.
111. **Brynildsen MP, Winkler JA, Spina CS, MacDonald IC, et al.** 2013. Potentiating antibacterial activity by predictably enhancing endogenous microbial ROS production. *Nat Biotechnol* **31**: 160-5.
112. **Feld L, Knudsen GM, Gram L.** 2012. Bactericidal antibiotics do not appear to cause oxidative stress in *Listeria monocytogenes*. *App Env Microbiol* **78**: 4353-7.
113. **Ezraty B, Vergnes A, Banzhaf M, Duverger Y, et al.** 2013. Fe-S cluster biosynthesis controls uptake of aminoglycosides in a ROS-less death pathway. *Science* **340**: 1583-7.
114. **Keren I, Wu Y, Inocencio J, Mulcahy LR, et al.** 2013. Killing by bactericidal antibiotics does not depend on reactive oxygen species. *Science* **339**: 1213-6.
115. **Liu Y, Imlay JA.** 2013. Cell death from antibiotics without the involvement of reactive oxygen species. *Science* **339**: 1210-3.
116. **Gomez MJ, Pazos F, Guijarro FJ, de Lorenzo V, et al.** 2007. The environmental fate of organic pollutants through the global microbial metabolism. *Mol Sys Biol* **3**: 114.

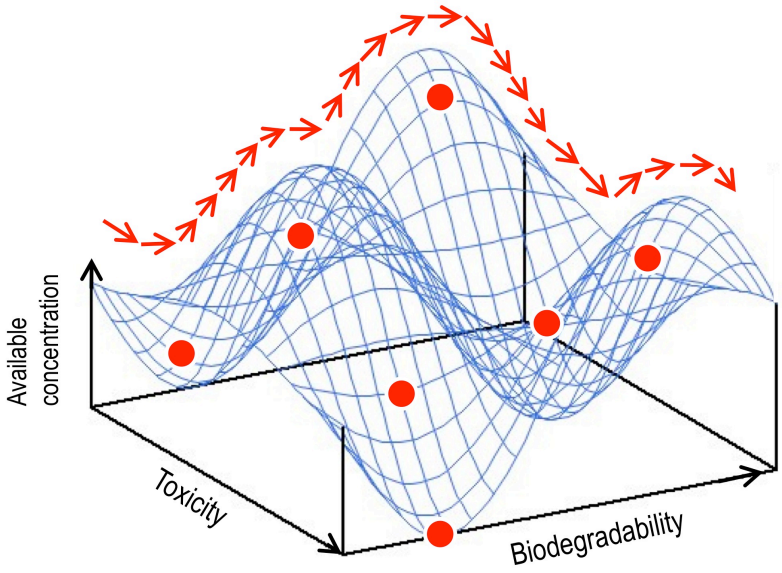
FIGURES

Figure 1. The growing expansion of the Central Dogma (CD).



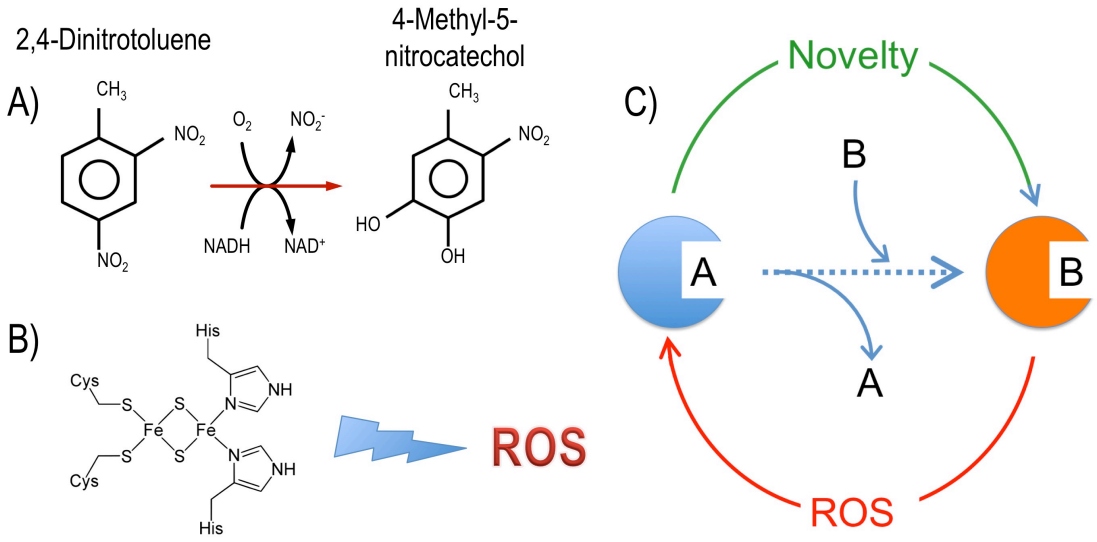
A: Canonical CD. The flux of information goes uni-directionally from DNA to proteins (general flow), though exceptionally RNA can also replicate by itself and be copied into DNA (special CD). B: Extended CD. The connection of proteins to metabolism is made explicit. Note that the information handover DNA → RNA → proteins involves a direct code, whilst that of proteins → metabolism has various layers of convolution. C: Metabolism constrains the information flow. This upside-down sketch emphasizes the dominant role of metabolism in enabling the enactment of the CD flow. Whilst the effect of metabolism on biosynthesis of information-bearing macromolecules is indirect, it does ensure that the instructions emanating from DNA do not conflict with the existing biochemical network. This can be interpreted as a feedback loop whereby metabolism informs the CD actors to proceed or not with the information flow according to physiological conditions.

Figure 2. The conquest of new plots of the chemical landscape.



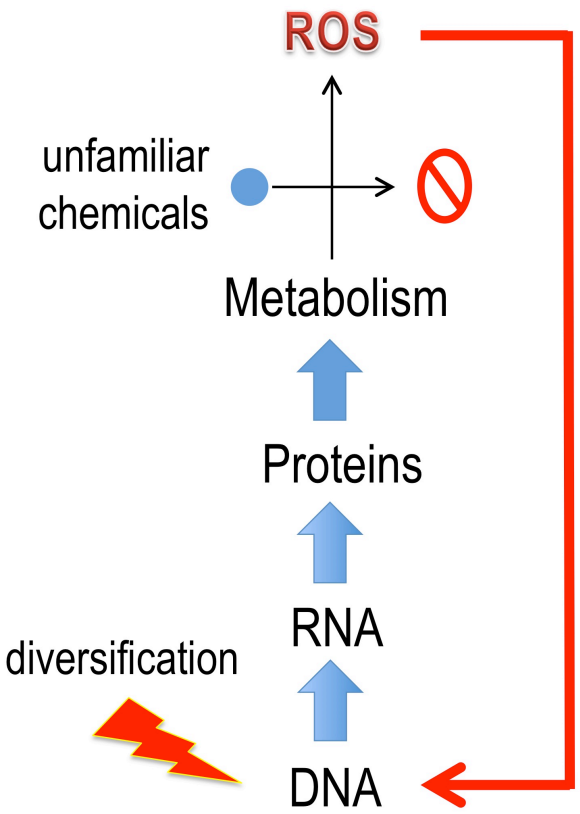
The graph sketches the evolutionary push of pathogens and biodegradative environmental bacteria as the growing occupation (*biologization*) of the chemical landscape of the target niche. Every chemical available maps somewhere in a space defined by 3 parameters: available concentration, toxicity and biodegradability. The last is a parameter that can be scored based on the distribution of atomic triplets in the chemical structure and the familiarity of such structure with core biomolecules [116]. The process of niche colonization can be seen as the progressive exploration of the existing chemical scenery in order to integrate it into one's metabolic network (i.e. transformation of xenobiotics and unfamiliar chemicals into endo- and exo-metabolites).

Figure 3. The metabolic stress-novelty cycle.



A: Reaction catalyzed by the 2,4-dinitrotoluene dioxigenase (DntA) of *Burkholderia* sp DNT. This enzyme seems to have originated in a precursor naphthalene dioxigenase that has acquired the (partial) ability to use DNT as a substrate. B: Production of reactive oxygen species upon faulty oxidation of DNT at the non-heme Rieske active center of DntA. C: The novelty creation cycle. The figure represents the process through which a strain able to degrade compound A ends up degrading compound B. The ROS produced by faulty dioxygenation both places the population at the verge of collapse but simultaneously diversifies genetically the same population and thus fosters the exploration of the solution space. Those individuals that survive the process (if any) have thus gained a new fitness peak. [89] has previously argued that this is an anti-fragile cycle, as long as the system ultimately benefits from shocks.

Figure 4. A feedback loop metabolism → DNA in the CD.



The figure sketches the main claims of this article. Metabolic troubles (e.g. faulty oxidation of no-substrates or poor substrates of oxygenases) cause ROS, which mutates DNA (as well as damaging RNA), triggers the SOS response and brings about genetic diversification, which might find a solution to the metabolic problem that originates the release of the mutagenic agent. While the ROS → DNA loop does not involve transfer of coded information it does deliver an input that accelerates the rate of novelty production.