



HAL
open science

Small molecules as products of evolution

Gabriel Markov, Vincent Laudet

► **To cite this version:**

Gabriel Markov, Vincent Laudet. Small molecules as products of evolution. *Current Biology - CB*, 2022, 32 (3), pp.R100-R105. 10.1016/j.cub.2021.12.038 . hal-03937973

HAL Id: hal-03937973

<https://hal.science/hal-03937973v1>

Submitted on 13 Jan 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Small molecules as products of evolution

Gabriel V. Markov^{1*} and Vincent Laudet^{2,3*}

¹Sorbonne Université, CNRS, UMR 8227 Integrative Biology of Marine Models, (LBI2M, UMR8227), Station Biologique de Roscoff (SBR), 29680 Roscoff, France

²Okinawa Institute of Science and Technology. 1919-1 Tancha, Onna-son, 904-0495 Okinawa, Japan.

³Marine Research Station, Institute of Cellular and Organismic Biology (ICOB), Academia Sinica, 23-10, Dah-Uen Rd, Jiau Shi, I-Lan 262, Taiwan

*co-corresponding authors: gabriel.markov@sb-roscoff.fr, vincent.laudet@oist.jp

Abstract

The functional evolution of ancient proteins has recently been reconstructed using molecular phylogeny, and the activity of the deduced molecules can be tested. Unfortunately, the world of small molecules escapes such “resurrection” studies, which rely on ancestral sequence reconstruction. These studies implicitly assume that only the proteins evolve whereas the small molecules are presumed to be identical now and in the past. Recent evolutionary analysis of biochemical pathways, however, as well as the impressive surge of metabolomics are changing this situation and it is becoming possible to reconstruct ancient biochemical pathways. We can now begin to infer the chemical structures of key molecules that were present in the past, synthesize those molecules and test their ability to interact with their cognate (ancient) protein partners. In this article, we discuss the possibilities offered by these new methodological developments and provide key examples. We also highlight four principles that are important to consider when studying the evolution of small molecules: catalytic promiscuity, metabolic reconfiguration, coevolution and bidirectional interactions. These new developments call for an alliance between organic chemists and evolutionary scientists to investigate the diversity of the chemical building blocks of life, and the evolution of their biosynthetic pathways.

Introduction

Understanding how biodiversity emerged is a major endeavor of biological research and not only concerns the “*endless forms most beautiful*” celebrated by Darwin but also the molecules that bring them into being. Whether at the morphological or molecular level, there are two main approaches to better understand the origin and diversification of the products of evolution. The first is to compare the living forms and to infer, by comparative anatomy and phylogeny, what has been the evolutionary path that allowed their emergence¹. The second approach is to dig in the past and to identify the forms that were living previously. Thanks to paleontology and its exquisite reconstruction of ancient plants, animals and even microorganisms, this has considerably enriched our view of life and offered us unique insights on ancient worlds.

Until recently, however, these approaches were limited to extinct organisms and there was no obvious possibility to perform such a historical approach at the molecular level. Ancient DNA analysis has offered fascinating possibilities but is limited to quite recent periods (ca. 1-million-years BP). It cannot allow the study of events that occurred millions of years ago². Comparison of protein sequences was an important advance as it permitted the inference of the sequences of ancestral proteins. These could then be reconstructed to directly test their function experimentally and thereby follow their functional evolution³. These methods rely on ancestral protein sequence reconstruction using molecular phylogenetic inference. The world of small molecules, including organic substrates or products, totally escapes such efforts of evolutionary studies. This is particularly problematic in the context of biochemical evolution, such as the evolution of metabolic pathways or the reconstruction of ancestral hormone/receptor couples, in cases where hormones are small molecules rather than proteins or peptides. Indeed, in these cases, which include most of metabolic regulation, it has been considered that only the proteins evolve. The small signaling molecules and their biosynthetic precursors have been implicitly assumed to be identical now and in the past⁴. Such an anachronistic analysis can give rise to incorrect inferences of evolutionary trajectories (Figure 1). The reason for such a limitation was obvious: whereas we had both a conceptual and technical framework allowing us to reconstruct proteins that are encoded by DNA, we did not know how to predict and reconstruct the small molecules, which are produced by these ancestral enzymes arranged in metabolic pathways.

With the emergence of genomics and the prodigious amount of sequence information — in particular high quality complete genomes from a variety of organisms — as well as the impressive recent surge of metabolomics, it has become possible to perform evolutionary analysis of biochemical pathways, to reconstruct ancient pathways and therefore to perform relevant historical tests (Figure 1).

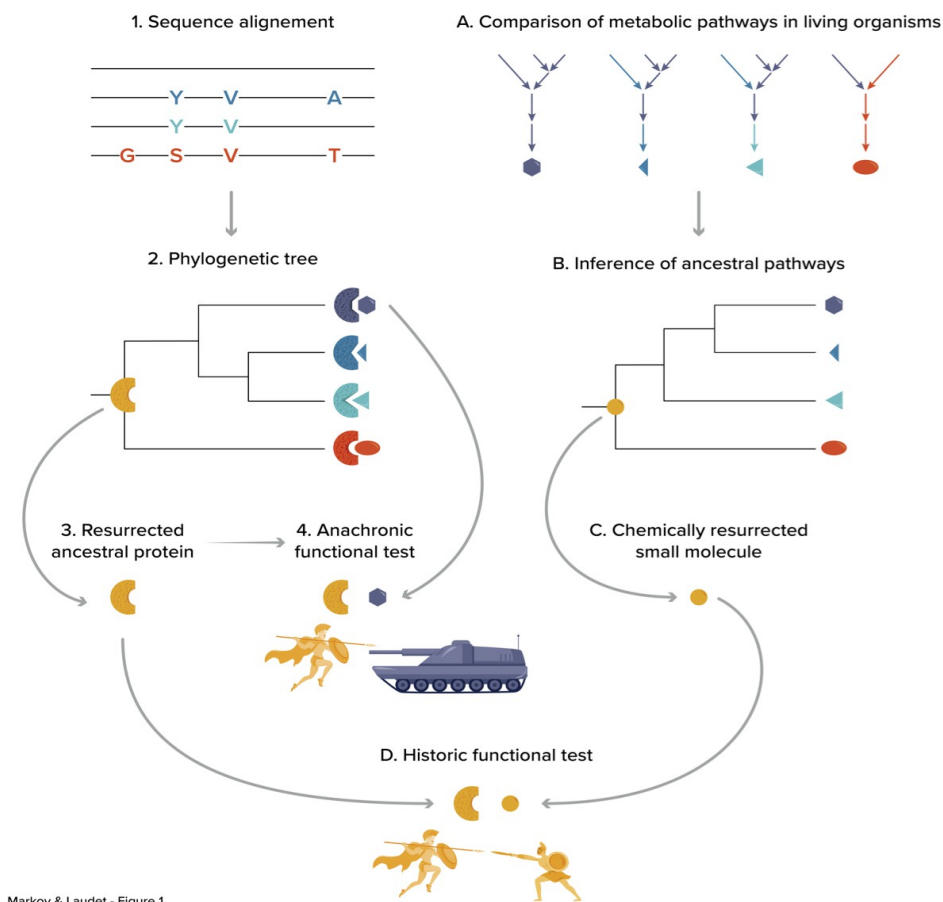


Figure 1: Resurrection of ancient protein and small molecule couples

The resurrection of ancient proteins starts by sequence alignment of protein sequences from living organisms (1), from this the construction of a phylogenetic tree (2) allows to reconstruct the evolutionary history and, in a separate step to predict ancestral sequences at each node of the tree. It is then possible using gene synthesis techniques to produce the predicted protein (3) giving rise to a so-called “resurrected” protein. The function of this protein can then be evaluated but when it is a protein interacting with a small molecule it carries the risk of a anachronistic functional test: The ancestral protein is confronted with a modern small molecule. To avoid this risk, it is necessary to perform the same exercise on the small molecule side, although this is less straightforward than with protein sequence: the comparison of metabolic pathways existing in living organism (A), the inference of the ancestral pathway either by cladistic analysis or genome wide reconstruction of metabolic pathways (B), the chemical synthesis of the inferred molecule of interest giving rise to a “resurrected” small molecule (C). It is only after these steps that a meaningful historical functional test can be performed (D).

How to reconstruct an ancestral pathway? The case of steroids.

To illustrate the resurrection of the biochemical pathways and molecules that were present in the past let us consider the evolutionary history of the vertebrate steroid receptors and their ligands, which provides an excellent case study to illustrate the coevolution of ligands and their receptors. Vertebrate steroid hormones are all derived from cholesterol and play very important functions in reproduction (androgens, estrogens, progesterone), in the regulation of metabolism and in the regulation of water and ion balance (glucocorticoids and mineralocorticoids). Their action is mediated by a group of six nuclear receptors: two estrogen receptors ER α and ER β and four ketosteroid receptors the androgen, progesterone, glucocorticoid and mineralocorticoid receptors, (AR, PR, GR and MR, respectively). Reconstruction of ancestral proteins has suggested that the ancestral steroid receptor was an estrogen receptor and that other steroid receptors exploited molecules that were metabolic intermediates of estrogen synthesis⁵. This so-called “ligand exploitation model” has been challenged mainly on the ground that there was no clear evidence that 17 β -estradiol, the natural ligand, was actually an ancient molecule⁶. In these analyses, the authors reconstructed an ancient receptor they called AncSR1. The question arose, however, as to what was the ancestral ligand of this resurrected receptor? The analysis of bilaterian genomes, searching for orthologues of enzymes important for estrogen biosynthesis, produced negative results⁷. To tackle this question, therefore, we performed a cladistic analysis of steroidogenesis, focusing not on the sequences of the enzymes in the pathway but on the sequence of the chemical reactions⁸. This approach, which was first elaborated in the context of universal amino acid metabolism⁹, allowed us to compare the sequence of the biochemical pathways, that is the suite of chemical reactions, even without information about the specific enzymes catalyzing each reaction. The idea here is to consider biosynthetic pathways as products of evolution that can be compared by ‘comparative anatomy’ and cladistics, just as two morphological structures are compared. When several vertebrate species are compared using morphological criteria such as the presence or absence or the shape of well-defined morphological structures (e.g., a jaw bone) this comparison is valid from an evolutionary point of view because the formation of these structures is controlled by genes that may be unknown but are nevertheless following the Darwinian “descent with modification” principle. Exactly like morphological structures, metabolic pathways are controlled by genes, which may be unknown but are targets of evolution. Therefore, the structure of these pathways can be compared as bone shape is compared in the construction of an evolutionary tree based on morphology. In the case of steroidogenesis, a character matrix was built to objectively assess the similarity between biochemical reactions, allowing to measure successions of identical or similar reactions and therefore similarities between pathways⁸. This matrix was then analyzed by parsimony to generate a tree in which each branch is not a species or a molecule but a metabolic pathway.

We could therefore infer an ancestral pathway and propose that aromatized steroids in chordates occurred in cholesterol derivatives that did not undergo side-chain cleavage. The simplest molecule of this new class, paraestrol A, was chemically synthesized, and was experimentally confirmed to bind the ancestral steroid receptor AncSR1, albeit with lower affinity than 17 β -estradiol, offering a unique window on what could have been the ancestral steroid hormone precursor⁸. Further work will be necessary to clarify how additional hydroxylations on the paraestrol backbone would have increased its affinity for steroid receptors to a level compatible with hormonal signaling.

Four principles of metabolic evolution.

As the example highlighted above shows, it is now possible to infer the chemical structures of key molecules present in the past, synthesize those molecules and test their ability to interact with their ancient proteins. We consider that four principles must be considered when studying small molecule evolution (Figure 2). Because steroids have been widely studied from an evolutionary point of view, we use them below to illustrate these principles. Ancestral reconstructions of ligands and receptors have been applied successfully to many other metabolic pathways in different organisms such as flies, nematodes, or plants. We therefore briefly summarize particularly illuminating examples of these other pathways to illustrate those general principles.

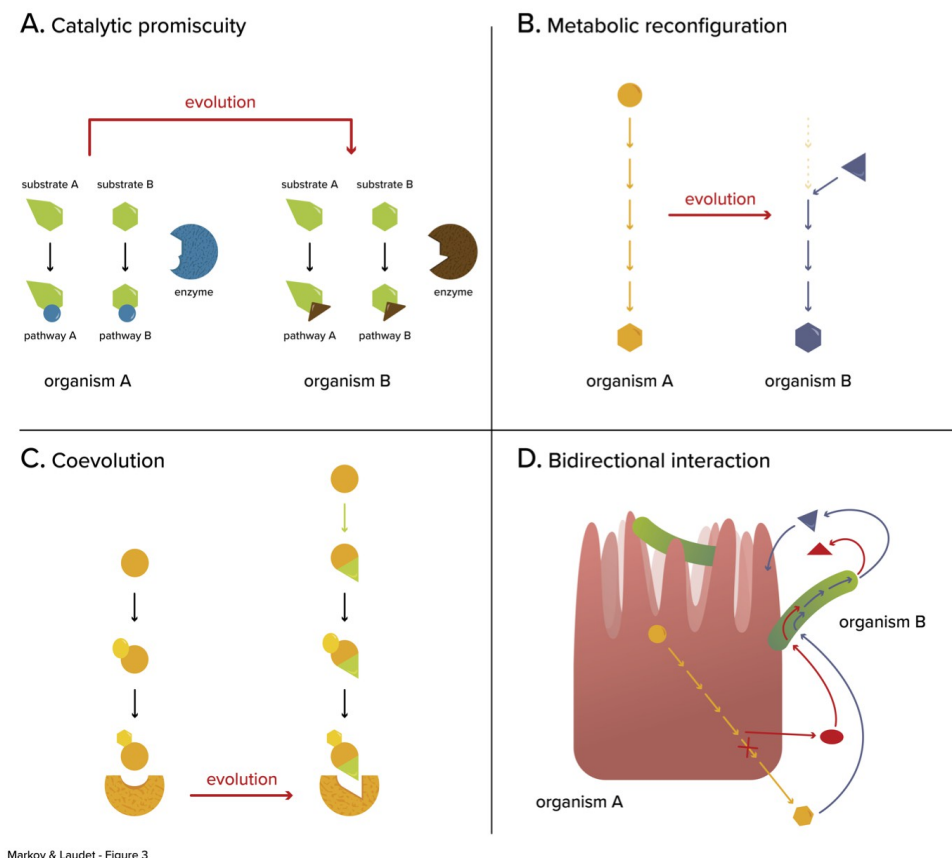


Figure 2. Four general principles regarding evolution of metabolic pathways involved in small molecule biosynthesis.

A. Catalytic promiscuity is the ability of a single enzyme (the blue one) to catalyze the same molecular transformation (blue disk) on two different molecular backbones (in green). Thus, if its catalytic activity is modified during evolution (red arrow), this can lead to changes in several distinct pathways where the concerned reaction was occurring. B. Metabolic reconfiguration occurs when a new pathway is generated by recombination of already existing reactions previously occurring in distinct pathways. In the example here during evolution (red arrow), the loss of two reactions (dotted yellow arrows) is compensated by the recruitment of an enzyme enabling the conservation of the downstream steps of the pathway, being fueled from a distinct precursor. C. Coevolution illustrating how a change in a ligand/receptor interaction during evolution (red arrow) must include change in both metabolic enzyme and receptor ligand binding pocket. D. Bidirectional interaction at the holobiont level. An organism A (here an intestinal cell) can generate a product active on organism B (here a commensal bacteria) that in turn can modify it to produce a molecule that will act in return on organism A. When an evolutionary change (red arrow) occurred in the metabolic pathway of organism A this affects both organisms.

The first of these principles is **catalytic promiscuity** (Figure 2A). When a molecular actor (for example, an enzyme or a receptor) can bind to different types of molecules the phenotype induced by a change in this molecular actor may affect several characters, sometimes quite disparate^{10,11}. This is well illustrated by the case of the CYP11A enzyme that is responsible for the removal of the vertebrate steroid side chain. Recent studies have established that CYP11A is able to catalyze hydroxylation of various sterols, not only cholesterol, but also vitamin D3 metabolites suggesting an action much wider than previously thought¹². This suggests that complex unanticipated links could exist between an enzyme and small molecules implicated in distinct metabolic regulations. In other words, a change in a given enzyme (either of expression pattern, expression level or specificity) could impact more than one product or pathway, opening the way to pleiotropic changes. Other similar examples include several 17 β -hydroxysteroid dehydrogenases (17 β HSD) enzymes known to be important for estrogen synthesis, but which are also important for retinoic acid and prostaglandin signaling¹³. The well documented catalytic promiscuity of HSD enzymes is acknowledged has a major driver of diversification in vertebrate steroidogenic pathways¹⁴. Promiscuity at the receptor level further reinforces the importance of pleiotropic effects on metabolic pathways. It has been well described for the vitamin D receptor, which in the intestine can behave as a bile acid receptor¹⁵, and for the estrogen receptor, which binds a great variety of ligands besides 17 β -estradiol¹⁶.

The importance of convergence to reveal basic principles of the evolution of small molecules, such as catalytic promiscuity is also well exemplified by a pheromone system involved in relationships between butterflies and plants. It has long been known that plants and insects often use the same compounds for communication and now the underlying biochemical mechanisms are starting to be uncovered¹⁷. Even within plants, there are multiple examples of convergent evolution, with promiscuous cytochrome p450s (CYPs) being recruited many times into identical catalytic reactions, either among closely related subfamilies, or even across most distant paralogs¹⁸

The second principle is **metabolic reconfiguration**, sometimes also called **patchwork evolution** (Figure 2B), which means that chemical reactions and enzymes can be reshuffled to generate a given compound¹⁹. This allows the interpretation of how steroid aromatization was recruited into the pathway of cleaved side-chain steroids after having previously operated on steroids with a side chain⁸. This is also well illustrated by the very early steps of cholesterol biosynthesis. Sterols—cholesterol in animals, phytosterols in plants and ergosterols in fungi—are critical for life. There is even recent evidence that sterols are synthesized *de novo* in some bacterial phyla²⁰. Cholesterol is formed from acetyl-CoA by a metabolic pathway with approximately 20 steps, and vertebrates are able to synthesize cholesterol in an autotrophic manner²¹. But this is not the case for all animals and several invertebrates are in fact auxotrophs for sterols; that is, they rely on an external source. Even though nematode worms use cholesterol as a source of critical hormonal compounds such as dafachronic acid, which coordinates their life cycle, they cannot synthesize it. Indeed, nematode genomes lack the first three enzymes required for cholesterol biosynthesis. In a very elegant analysis Shamsuzzana et al.,²² have shown that nematodes can transform plant phytosterols and fungal ergosterols into cholesterol by co-opting downstream cholesterol synthesizing enzymes and modifying them to transform phytosterols and ergosterols to an

intermediary compound, desmosterol. Desmosterol is finally converted into cholesterol! Furthermore, they also show that the same strategy—loss of key upstream enzymes of the cholesterol synthesis pathway and co-option of downstream enzymes to use dietary compounds from plants and/or fungi—is used by other invertebrates, suggesting a widespread cholesterol auxotrophy by interkingdom trophic interactions²².

Analysis of the pheromone signaling system of nematodes also illustrates the principle of metabolic reconfiguration. Pheromone communication is critically important for nematodes, either in the short-term (e.g., mate attraction) or for long-term life history-related responses (e.g., formation of dauer larvae). Hundreds of extremely diverse but related compounds, collectively known as ascarosides, that consist of a sugar moiety coupled to a fatty acid sidechain, have been characterized²³. Systematic metabolomic profiling of ascarosides in 30 species of *Pristionchus* revealed evidence of a high level of biochemical innovation by reshuffling of metabolic activities²⁴. Analyzing the occurrence of these molecules in the light of species phylogeny allowed the authors to reveal many clade-related compounds, fully consistent with the notion of metabolic pathway “*descent with modification*”. Interestingly, this analysis also revealed a recurrent convergent evolution of modular molecules in distant branches of the evolutionary tree indicating that biochemical evolution is a repeatable process in these worms²⁴.

Building on the two previous ones, the third principle, **coevolution** (Figure 2C), has been elaborated independently in the field of ligand-receptor evolution (especially for peptide ligands and their receptors) and, recently, more precisely formalized for enzyme-substrate couples²⁵. According to this principle, ligands and their receptors are not free to evolve since any mutation in one partner must be compensated by a mutation in the other to avoid losing the interaction. Even if this is much more complicated to demonstrate in the case of small molecules that are the products of metabolic pathways, examples do exist for small molecules, again among steroids. Bile acids, which are one of the major end products of cholesterol metabolism in vertebrates, play an important role in the digestion of lipids and proteins as well as displaying antimicrobial activity in the intestine. They form a set of extremely variable compounds that differ widely across species. Their major receptor, the nuclear receptor FXR, changes its specificity for primary bile acid salt across species by altering the size and shape of its ligand binding pocket²⁶. For example, in lamprey and zebrafish the FXR ligand binding pocket is flat and can easily accommodate the planar bile acids found in these species. In contrast, human FXR has a curved pocket that binds the bent bile acid salts typically found in human²⁶. This clearly illustrates a case of coevolution even if, to our knowledge the evolution of the enzymes responsible for the production of the various types of bile acids present in various vertebrates has not been studied and has not been yet related to the evolution of FXR ligand binding ability. Another fascinating example of coevolution is provided by the role of progesterone, a precursor in corticoid biosynthesis, which is also an allosteric modulator of corticoid binding by the mineralocorticoid receptor (MR) in some vertebrates through binding to a non-canonical binding site located in the N-terminal domain of the receptor²⁷.

Detailed analysis of a specific pathway can also reveal ligand-receptor coevolution, and this has started to

provide key insights into the evolution of signaling pathways in plants and animals. In plants, the various components of the fatty-acid derived jasmonate pathway (important for stress response and other developmental and physiological regulations) were characterized in basal lineages including the liverwort *Marchantia polymorpha*²⁸. The main components of the pathway are conserved, but it was shown that the bioactive signaling molecule in *Marchantia* was not jasmonoyl-isoleucine as in *Arabidopsis*, but two isomeric forms of its precursor, called dinor-oxo-phytodienoic (12-oxo-phytodienoic acid). This change of ligands requires a single change of a residue in the receptor but also requires the modification of two pre-existing enzymes. Therefore ligand-receptor coevolution must be accompanied by a metabolic reconfiguration.

The fourth principle, **bidirectional interaction** (Figure 2D), refers to the notion that a given molecule can be active both on the organism that produces it and also on a neighboring species that can in turn alter it. This is particularly applicable to small molecules and once again is well exemplified by steroids. Bacteria can metabolize steroids found in their environment (for example in their eukaryotic hosts or, alternatively, sewage treatment plants)²⁹. Recent evidence suggests a real bidirectional effect, where steroids degraded by a biochemical pathway in a bacterium can, in return, be used as active components by eukaryotes. For example, whereas the retro-conversion of estrogens into androgens was deemed impossible in animals, given the irreversibility of the aromatization reaction, it was recently found to occur in an anaerobic bacterium, opening a previously unexplored microbe-host metabolic interdependency³⁰.

The importance of bidirectional interaction is also well demonstrated in plant and insects by a recent observation that revealed how wild tobacco, *Nicotiana attenuata*, protects itself against the activity of the compounds it uses to repel herbivores³¹. Like many other plants, wild tobacco uses specialized metabolites to decrease the harmful activity of herbivores, trying to target tissues or pathways that are present only in their animal enemies to avoid autotoxicity. Nicotine which targets the nervous systems of animals but has no effect on the plant itself is a classic example of this strategy. In tobacco, two cytochrome P450 enzymes acting on 17-hydroxygeranylinalool diterpene glycosides are critical for avoiding autotoxicity. The defensive function of 17-hydroxygeranylinalool diterpene glycosides on *Manduca sexta* caterpillars (the main predator of wild tobacco) is achieved by inhibiting sphingolipid biosynthesis in the insect through hydroxylation of the product they eat³¹. The hydroxylation of the diterpene compound by cytochrome P450 in the plant is also important to prevent autotoxicity, in addition to arming the chemical gun against the insect. The bidirectional interaction can therefore also be an arms race.

There is of course no reason to think that this type of bidirectional interaction is confined to those well-studied examples. If we open this principle to other cases, for example regarding the interactions between organisms in the marine world, the perspectives offered are dizzying³². This principle has been discussed previously in the context of environmental signaling and endocrine disruption³³ but must be extended outside this field and generalized: in short, a mutation in a given metabolic pathway can affect different organisms that interact with the organism in which the mutation takes place. This is one more reason to study metabolic evolution in a global and systemic perspective.

Ongoing improvements in methodological approaches necessary to tackle metabolic pathway evolution

Inferring ancestral metabolic pathways is still a great challenge, but the basic tools necessary to develop such approaches are maturing. Complete genome sequences have made genome-scale metabolic modeling a more and more widely used approach³⁴. This approach enables the reconstruction of an explicit model of whole organism metabolism, based on sequence annotation data. This can be coupled with standard comparative genomics studies to integrate specific pathways in the context of global metabolism³⁵. The development of genome-scale metabolic network reconstructions has also led to attempts to directly infer ancestral metabolic networks³⁶.

Another promising development is the first database of ancestral resurrected proteins³⁷. This opens the possibility to cross-validate ancestral metabolic network reconstructions by ensuring that enzyme sequences deposited in such databases are correctly inferred to be present in ancestral genome-scale reconstructions and associated with the appropriate biochemical activity. A critical point in that respect is ensuring that already described metabolites in a given species are indeed present in the corresponding genome-scale metabolic network reconstruction. There are now databases that make it possible to automatically assemble species-specific metabolomes, although they are still in their infancy and biased towards well established genetic model species³⁸. Thanks to ongoing work on emerging model species, however, more datasets will certainly be integrated in the future.

Finally, to understand the evolution of small molecules it is necessary to explore the space of possible metabolite-protein interactions using high-throughput methods. There is promising progress in the mass-spectrometry and NMR approaches (combined with other separation techniques) to distinguish proteins with or without bound ligands. This should allow the identification of new targets for plant-derived metabolites³⁹. With more and more data on such interactions integrated in an evolutionary framework, it will be possible to gain a better understanding of when those interactions arose during the diversification of the lineages concerned.

Conclusion: A new alliance with chemistry

It is time to build a wider perspective encompassing the diversity of the chemical building blocks of life and their specificities in term of evolutionary mechanisms. The use of analytical techniques such as high-resolution mass spectrometry, in the field of ancient proteomics and ancient lipidomics is already providing a very fruitful unifying factor, unearthing new types of molecular fossils⁴⁰, but this move will necessitate the participation of organic chemists eager to synthesize ancestral small molecules predicted through evolutionary analysis that will be used for biochemical experiments on ancient proteins.

Another area in which such an alliance between evolutionists and chemists could produce fruitful conceptual developments is the illustration of metabolic pathways. Just as phylogenetic trees only provide a schematic view of the relationships between animal or plant populations, we believe that the metabolic reactions depicted as they currently are, as a suite of ordered reactions, is probably an oversimplification. Developing network-like representations may allow better capture of the essence of the evolutionary process. We believe that it may be fruitful to think of a new way to represent, and therefore conceptualize, metabolic pathways. We strongly believe that incorporating the evolutionary

reasoning as discussed here will be helpful for such an aim. All these developments, and probably many others, can only be done through a close collaboration between molecular evolutionists, biochemists, and chemists. After Evo-Devo, let us start Evo-Chemo!

Acknowledgements

We thank Tim Hunt, Catherine Jesus and Catherine Leblanc for stimulating discussions that greatly helped us to sharpen our ideas. We also thank Marleen Klann, Sam Reiter, Ryota Kabe, Dino Moras, Agneesh Barua, Simon Dittami and Jonas Collén for critical reading of this manuscript. Work from our laboratories is supported by OIST, Academia Sinica, CNRS and Sorbonne Université. GVM participated to the National Research Agency investment expenditure program IDEALG (ANR-10-BTBR-04).

References

- ¹Harvey, P.H., Pagel, M.D. (1991) *The Comparative Method in Evolutionary Biology*. Oxford (UK): Oxford Press.
- ²Leonardi, M., Librado, P., Der Sarkissian, C., Schubert, M., Alfarhan, A.H., Alquraishi, S.A., Al-Rasheid, K.A.S., Gamba, C., Willerslev, E., and Orlando, L. (2016). Evolutionary Patterns and Processes: Lessons from Ancient DNA. *Systematic Biology* 66, e1-e29.
- ³Hochberg, G.K.A., and Thornton, J.W. (2017). Reconstructing Ancient Proteins to Understand the Causes of Structure and Function. *Annual review of biophysics* 46, 247-269.
- ⁴Thornton, J.W., Need, E., and Crews, D. (2003). Resurrecting the ancestral steroid receptor: ancient origin of estrogen signaling. *Science* 301, 1714-1717.
- ⁵Thornton, J.W. (2001). Evolution of vertebrate steroid receptors from an ancestral estrogen receptor by ligand exploitation and serial genome expansions. *Proc Natl Acad Sci U S A* 98, 5671-5676.
- ⁶Baker, M.E. (2002). Recent insights into the origins of adrenal and sex steroid receptors. *J. Mol. Endocrinol.* 28, 149-152.
- ⁷Markov, G.V., Tavares, R., Dauphin-Villemant, C., Demeneix, B.A., Baker, M.E., and Laudet, V. (2009). Independent elaboration of steroid hormone signaling pathways in metazoans. *Proc Natl Acad Sci U S A* 106, 11913-11918.
- ⁸Markov, G.V., Gutierrez-Mazariegos, J., Pitrat, D., Billas, I.M.L., Bonneton, F., Moras, D., Hasserodt, J., Lecointre, G., and Laudet, V. (2017). Origin of an ancient hormone/receptor couple revealed by resurrection of an ancestral estrogen. *Science Advances* 3, e1601778.
- ⁹Cunchillos, C., and Lecointre, G. (2005). Integrating the universal metabolism into a phylogenetic analysis. *Mol. Biol. Evol.* 22, 1-11.
- ¹⁰Peracchi, A. (2018). The Limits of Enzyme Specificity and the Evolution of Metabolism. *Trends in Biochemical Sciences* 43, 984-996.
- ¹¹Cox, R.M. (2020). Sex steroids as mediators of phenotypic integration, genetic correlations, and evolutionary transitions. *Molecular and Cellular Endocrinology* 502, 110668.
- ¹²Slominski, A.T., Li, W., Kim, T.-K., Semak, I., Wang, J., Zjawiony, J.K., and Tuckey, R.C. (2015). Novel activities of CYP11A1 and their potential physiological significance. *The Journal of Steroid Biochemistry and Molecular Biology* 151, 25-37.

- ¹³ Penning, T.M. (2019). AKR1C3 (type 5 17 β -hydroxysteroid dehydrogenase/prostaglandin F synthase): Roles in malignancy and endocrine disorders. *Molecular and Cellular Endocrinology* 489, 82-91.
- ¹⁴ Baker, M.E., Nelson, D.R., and Studer R.A. (2015). Origin of the response to adrenal and sex steroids: Roles of promiscuity and co-evolution of enzymes and steroid receptors. *J Steroid Biochem Mol Bio* 115, 12-24.
- ¹⁵ Makishima, M., Lu, T.T., Xie, W., Whitfield, G.K., Domoto, H., Evans, R.M., Haussler, M.R., Mangelsdorf, D.J. (2002). Vitamin D receptor as an intestinal bile acid sensor. *Science* 296, 1313-1316.
- ¹⁶ Baker, M.E., and Lathe, R. (2018). The promiscuous estrogen receptor: Evolution of physiological estrogens and response to phytochemicals and endocrine disruptors. *The Journal of Steroid Biochemistry and Molecular Biology* 184, 29-3.
- ¹⁷ Darragh, K., Orteu, A., Black, D., Byers, K.J.R.P., Szczerbowski, D., Warren, I.A., Rastas, P., Pinharanda, A., Davey, J.W., Fernanda Garza, S., et al. (2021). A novel terpene synthase controls differences in anti-aphrodisiac pheromone production between closely related *Heliconius* butterflies. *PLoS Biology* 19, e3001022.
- ¹⁸ Hansen, C.C., Nelson D.R., Lindberg Moller B., Werck-Reichhart D. (2021). Plant cytochrome P450 plasticity and evolution. *Molecular Plant* 14, 1244-1265.
- ¹⁹ Lazcano, A. & Miller, S. L. (1999). On the origin of metabolic pathways. *Journal of Molecular Evolution* 49, 424-431.
- ²⁰ Hoshino, Y., and Gaucher, E.A. (2021). Evolution of bacterial steroid biosynthesis and its impact on eukaryogenesis. *Proc Natl Acad Sci U S A* 118, e2101276118.
- ²¹ Capell-Hattam, I.M., Brown, A.J. (2020) Sterol evolution: cholesterol synthesis in animals is less a required trait than an acquired taste. *Current Biology* 30, R886-R888.
- ²² Shamsuzzama, Lebedev, R., Trabelcy, B., Langier Goncalves, I., Gerchman, Y., Sapir, A. (2020). Metabolic Reconfiguration in *C. elegans* Suggests a Pathway for Widespread Sterol Auxotrophy in the Animal Kingdom. *Current Biology* 30, 3031 - 3038.e7.
- ²³ Bose, N., Meyer, J.M., Yim, J.J., Mayer, M.G., Markov, G.V., Ogawa, A., Schroeder, F.C., and Sommer, R.J. (2014). Natural variation in dauer pheromone production and sensing supports intraspecific competition in nematodes. *Current Biology* 24, 1536-1541.
- ²⁴ Dong, C., Weadick, C.J., Truffault, V., and Sommer, R.J. (2020). Convergent evolution of small molecule pheromones in *Pristionchus* nematodes. *ELife* 9.
- ²⁵ Noda-Garcia, L., Liebermeister, W., and Tawfik, D.S. (2018). Metabolite-Enzyme Coevolution: From Single Enzymes to Metabolic Pathways and Networks. *Annual Review of Biochemistry* 87, 187-216.
- ²⁶ Reschly, E.J., Ai, N., Ekins, S., Welsh, W.J., Hagey, L.R., Hofmann, A.F., and Krasowski, M. D. (2008). Evolution of the bile salt nuclear receptor FXR in vertebrates. *J. Lipid Res.* 49, 1577-1587.
- ²⁷ Katsu, Y., Kohno, S., Oka, K., Lin, X., Otake, S., Pillai, N.E., Takagi, W., Hyodo, S., Venkatesh, B., and Baker, M.E. (2019). Transcriptional activation of elephant shark mineralocorticoid receptor by corticosteroids, progesterone, and spironolactone. *Science Signaling* 12.
- ²⁸ Monte, I., Ishida, S., Zamarreño, A.M., Hamberg, M., Franco-Zorrilla, J.M., García-Casado, G., Gouhier-Darimont, C., Reymond, P., Takahashi, K., García-Mina, J.M.; et al. (2018). Ligand-receptor co-evolution shaped the jasmonate pathway in land plants. *Nature Chemical Biology*, 2018, 14, 480-488.
- ²⁹ Holert, J., Cardenas, E., Bergstrand, L.H., Zaikova, E., Hahn, A.S., Hallam, S.J., Mohn, W.W. (2018) Metagenomes Reveal

- Global Distribution of Bacterial Steroid Catabolism in Natural, Engineered, and Host Environments. *mBio*, 9.
- ³⁰ Wang, P.-H., Chen, Y.-L., Wei, S.T.-S., Wu, K., Lee, T.-H., Wu, T.-Y., and Chiang, Y.-R. (2020). Retroconversion of estrogens into androgens by bacteria via a cobalamin-mediated methylation. *Proceedings of the National Academy of Sciences, National Academy of Sciences* 117, 1395-1403.
- ³¹ Li, J., Halitschke, R., Li, D., Paetz, C., Su, H., Heiling, S., Xu, S., and Baldwin, I.T. (2021). Controlled hydroxylations of diterpenoids allow for plant chemical defense without autotoxicity. *Science* 371, 255-260.
- ³² Markov, G.V., Girard, J., Laudet, V., Leblanc, C. (2018). Hormonally active phytochemicals from macroalgae: A largely untapped source of ligands to deorphanize nuclear receptors in emerging marine animal models. *Gen. Comp. Endocrinol.* 265, 41-45.
- ³³ McLachlan, J.A.(2001). Environmental signaling: what embryos and evolution teach us about endocrine disrupting chemicals. *Endocrine reviews* 22, 319-341.
- ³⁴ Gu, C., Kim, G.B., Kim, W.J., Kim, H.U., Lee, S.Y. (2019). Current status and applications of genome-scale metabolic models. *Genome Biology* 20, 121.
- ³⁵ Cannell, N., Emms, D.M., Hetherington, A.J., MacKay, J., Kelly, S., Dolan, L., and Sweetlove, L.J. (2020) Multiple Metabolic Innovations and losses are associated with major transitions in land plant evolution. *Current Biology* 30, 1783-1800.e11.
- ³⁶ Psomopoulos, F.E., van Helden, J., Médigue, C., Chasapi, A., and Ouzounis, C.A. (2020). Ancestral state reconstruction of metabolic pathways across pangenome ensembles. *Microbial Genomics*.
- ³⁷ Carletti, M.S., Monzon, A.M., Garcia-Rios, E., Benitez, G., Hirsh, L., Fornasari, M.S., and Parisi, G. (2020). Revenant: a database of resurrected proteins. *Database*.
- ³⁸ Salek, R.M., Conesa, P., Cochrane, K., Haug, K., Williams, M., Kale, N., Moreno, P., Jayaseelan, K.V., Macias, J.R., Nainala, V.C.; *et al.* (2017). Automated assembly of species metabolomes through data submission into a public repository. *GigaScience* 6, 1-4.
- ³⁹ Venegas-Molina, J., Molina-Hidalgo, F.J., Clicque, E., and Goossens, A. (2021). Why and how to dig into plant metabolite-protein interactions. *Trends in Plant Science* 26, 472-483.
- ⁴⁰ Cappellini, E., Prohaska, A., Racimo, F., Welker, F., Pedersen, M.W., Allentoft, M.E., de Barros Damgaard, P., Gutenbrunner, P., Dunne, J., Hammann, S., *et al.* (2018). Ancient biomolecules and evolutionary inference. *Annual Review of Biochemistry* 87, 1029-1060.