Secondary metabolites are often considered within the remit of bacterial or plant research, but animals also contain a plethora of these molecules with important functional roles. Classical feeding studies demonstrate that, whereas some are derived from diet, many of these compounds are made within the animals. In the past 15 years, the genetic and biochemical origin of several animal natural products has been traced to partnerships with symbiotic bacteria. More recently, a number of animal genome-encoded pathways to microbe-like natural products have come to light. These pathways are sometimes horizontally acquired from bacteria, but more commonly they unveil a new and diverse animal biochemistry. In this review, we highlight recent examples of characterized animal biosynthetic enzymes that reveal an unanticipated breadth and intricacy in animal secondary metabolism. The results so far suggest that there may be an immense diversity of animal small molecules and biosynthetic enzymes awaiting discovery. This biosynthetic dark matter is just beginning to be understood, providing a relatively untapped frontier for discovery.

Animals contain many unique small molecules, including bioactive secondary metabolites (1, 2). The compounds are thought to be protective, offensive, or involved in communication (pheromones and hormones) (3–5). Despite the numerous compounds isolated so far, animals are often believed to have a relatively limited biochemistry. In part, this is because animals obtain chemicals from their diets (6), so it was thought that elaborate compounds could be acquired from the environment. Classical studies of the origins of animal secondary metabolites reveal that many are diet-derived, whereas others are made within the animals (7, 8). However, these studies mostly lacked the genetic and biochemical evidence required to determine whether de novo biosynthesized compounds originated within the animals themselves or within symbiotic partners.

Indeed, over the past 20 years or so, many “animal” biosynthetic pathways have been traced back to symbiotic bacteria, which form very tight associations with animals for chemical defense (8). This has led some workers to believe that most, if not all, animal secondary metabolites originate in microbes. Belying this belief, humans make secondary metabolites, such as steroids, prostaglandins, lipids, melamins, neurotransmitters, G protein–coupled receptor ligands, and related compounds, the biosyntheses of which are now textbook knowledge. Perhaps familiarity and the imprecision of the underlying definitions make these compounds seem less “secondary.” In contrast, for important compounds restricted to nonhuman animals, relatively little was known at the genetic level, with exceptions including a few insect pheromones with economic importance (9, 10).

Because of recent improved sequencing and bioinformatics methods coupled with rigorous genetic and biochemical analyses, there has been no less than a revolution in our ability to understand metabolism in nonhuman animals. The origins of the countless secondary metabolites from diverse animals are starting to be unraveled (Fig. 1). This short review covers some of the recent major advances in understanding animal secondary metabolism from diverse phyla. The scope and importance of these advances is not immediately clear because the published literature mostly focuses on single groups, such as insects or marine natural products. By instead examining the global perspective, trends are already emerging that help to understand the many unique ways that animals make elaborate secondary metabolites. Below, we walk through some of the different biochemical classes now known to originate in animal metabolism.

Shikimate-like pathways

Shikimic acid is important in the synthesis of aromatic amino acids, but shikimic acid and related derivatives are also central in secondary metabolism. Examples include drugs from plants such as warfarin and etoposide relatives, and bacterial antibiotics, such as chloramphenicol and aminoglycosides, among many others. Shikimic acid metabolism is not supposed to occur in animals, which otherwise obtain aromatic amino acids from diet (11). The enzyme dehydroquinate synthase (DHQS)\(^2\) catalyzes an early step in shikimate biosynthesis. A related enzyme ultimately leads to the synthesis of sunscreens: the mycosporine amino acids (MAAs) and related compounds (12).

---

This work was supported by National Institutes of Health Grant R35GM122521 (to E. W. S.). The authors declare that they have no conflicts of interest with the contents of this article. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

---

This is an Open Access article under the CC BY license.

1 To whom correspondence should be addressed: Dept. of Medicinal Chemistry, University of Utah, Salt Lake City, UT 84112. Tel.: 801-585-5234; E-mail: ews1@utah.edu.

2 The abbreviations used are: DHQS, dehydroquinate synthase; SH7P, sedoheptulose 7-phosphate; MAA, mycosporine amino acid; DDG, demethyl 4-deoxyxagudulos; DDGS, DGG synthase; EEV, 2-epi-5-epi-valiolone; EEVS, EEV synthase; HGT, horizontal gene transfer; IDS, isoprenyl diphosphate synthase; GPP, geranyl pyrophosphate; FPP, farnesyl pyrophosphate; TPS, terpene synthase/cyclase; PKS, polyketide synthase; FAS, fatty acid synthase; AT, acyltransferase; KS, ketosynthase; CP, carrier protein; ACP, acyl carrier protein; KR, ketoreductase; DH, dehydratase; ER, enoylreductase; NRPS, nonribosomal peptide synthetase; A, adenylation; C, condensation; AANAT, aryl-alkylamine-N-acetyl transferase; PQQ, pyrroloquinoline quinone.
The biosynthesis of MAAs was first characterized in cyanobacteria. Related biosynthetic pathways have been found throughout the animal tree of life, including in vertebrates such as fish, but not mammals (11, 13–15).

At least three major types of MAA-like biosynthetic pathways have been described (Fig. 2). The pathways synthesize natural sunscreens, including MAAs, as well as a series of related compounds such as gadusol and deoxygadusol. All of these pathways use an enzyme converting a sugar, sedheptulose 7-phosphate (SH7P), into a cyclitol, using DHQS-like enzymes. One of these, to MAAs in bacteria, first produces demethyl 4-deoxygadusol (DDG) using a DGG synthase (DDGS) (12). A methyltransferase then synthesizes deoxygadusol. Nonribosomal peptide synthetase (NRPS)-like and ATP-grasp enzymes add amino acids to form the final MAA product. A related pathway is found in cnidarians (e.g. corals), but with an interesting fusion of the methyltransferase with the DDGS enzyme (15).

A third variation, found in fish and other vertebrates, uses a different strategy. Instead of cyclizing SH7P in tandem with reduction by DDGS, this pathway first leads to 2-epi-5-epi-valiolone (EEV) via the EEV synthase (EEVS) enzyme (14). The EEVS is paired with a fused methyltransferase-oxidase enzyme, which synthesizes gadusol from EEV. Because the gadusol EEV was characterized both by heterologous expression and by knockdown in fish, the evidence for its function is quite strong.

Figure 1. Secondary metabolites from animals. Animals such as bird, urchins, fish, insects, worms, and corals contain diverse secondary metabolites. The biosynthetic pathways of some of these natural products are encoded in animal genomes and are similar to the biosynthetic machineries that microbes and plants use to produce small molecules. For the most part, the biosynthetic origins are unknown and represent a new frontier for research.
Now that the biosynthetic genes are in hand, it is clear that understanding biosynthesis of MAAs and related compounds in animals is complicated by the fact that they may have both \textit{de novo} and dietary origin and that \textit{de novo} synthesis can occur through symbiosis with bacteria (16) or through animal-encoded genes (11, 13). Animals are believed to have obtained their MAA-like biosynthetic machinery via horizontal gene transfer (HGT) from various microbes (14). The ability to make gadusol is believed to have been lost in the ancestral mammal (13). The widespread occurrence of these compounds attests to their biological importance in protecting from UV damage and other, mostly UV-related functions.

**Terpenes**

Isoprenoid/terpene natural products are widespread in all forms of life and include well-known compounds such as steroids. They have many well-established biological roles, including chemical defense, pheromones, hormones, and many others (17). Here, we will cover recent advances in animal terpene cyclization, other than sterols/steroids, which are well-studied because of their importance in human biology.

Isoprenoids are found in most animals. In bacteria and plants, isoprene precursors dimethylallyl pyrophosphate and isopentenyl pyrophosphate can be made either via the mevalonate or deoxyxylulose phosphate pathways, but in animals mevalonate is the source of these precursors (18). The early steps leading to mevalonate are nearly identical in all animals. The major exception is that some of the insect enzymes exhibit relaxed substrate selectivity, making homo isoprenoids that are important insect hormones (10). Subsequently, isoprenyl diphosphate synthases (IDSs) condense dimethylallyl pyrophosphate and isopentenyl pyrophosphate, leading to methyl-branched oligomers of five-carbon units, such as geranyl pyrophosphate (GPP, 10 carbons), farnesyl pyrophosphate (FPP, 15 carbons), and so on.

To make complex secondary metabolites, plants and bacteria use terpene synthases/cyclases (TPSs). In animals, bacteria- and plant-like TPSs have not been identified, so that the basis of \textit{de novo} terpene biosynthesis was mysterious. In 2009, an insect IDS was discovered that converts GPP to a dehydrated monoterpene (19), and insect carotenoids were found to originate via HGT (20), but cyclases were still not identified. This problem has recently been solved in a series of landmark studies (Fig. 3) (21–23).

In plants and microbes, TPSs have similar folds to IDSs and are thought to have evolved from ancestral IDS genes (24). With this in mind, in 2016, a beetle transcriptome was analyzed for IDS homologs (21). Nine transcripts were found encoding IDS-like sequences. These were expressed in \textit{Escherichia coli} and functionally characterized. One of the IDS-like enzymes, PstPS1, converted (Z,E)-FPP into primarily bicyclic sesquiterpenes. Other enzymes were IDSs, produced linear isoprenoids,
or had no detectable activity in the assay conditions. The use of \((Z,E)-FPP\) was surprising, given that most TPSs use \((E,E)-FPP\). One of the IDSs specifically synthesized \((Z,E)-FPP\), the substrate for the enzyme. Transcriptional and knockdown studies provided further evidence that the correct enzymes were identified.

The identification of PsTPS1 led to the realization that there are other IDS-like TPS enzymes encoded in insect genomes (21). These are more closely related to insect IDSs than to TPSs from other organisms, and thus have evolved convergently. Recently, insect TPSs from a stink bug and a beetle species have been characterized. These enzymes produce different monocyclic, bisabolene-like isoprenoids from \((E,E)-FPP\) (22, 23). Experimental evidence is very strong because of a combination of heterologous expression and knockdown studies. On the evolutionary front, one of the most interesting things to emerge was that TPSs have independently evolved from IDSs in several different insect lineages (23).

**Polyketide/fatty acid metabolism**

Animals make many lipid-derived natural products. Historically, specialized animal lipid metabolism has been considered to be degradative. For example, essential fatty acids such as arachidonic acid are oxidatively modified to produce prostaglandins. In a few of the examples below, degradative reactions from *de novo* or dietary fatty acids produce specialized metabolites. Recently, there has been an increased appreciation that animals can also build polyketides, complex lipid-like compounds normally associated with bacterial metabolism. For example, polyketide products include the bacterial products and pharmaceutical drugs erythromycin and amphotericin. Whereas the first animal polyketide synthase (PKS) was recognized about 16 years ago, recent advances in characterization and sequence analysis have revealed widespread families of PKSs in animal genomes.

Many of the pheromones used by insects and other organisms are derivatives of simple fatty acids (Fig. 4) (2). The proteins underlying many of these steps have been identified. For example, in moths, pheromones consist of defined mixtures of acids modified by redox and chain-shortening reactions (25). A similar, but much more elaborate, scheme can be found in the nematode hormones, ascarosides. Instead of simple oxidative modification of fatty acids, the ascarosides are formed by a combinatorial mixture of different degradative and building/tailoring steps to create a large series of derivatives (26–30).

Animals are, of course, capable of producing fatty acids from scratch using type I fatty acid synthase (FAS) enzymes (18). These enzymes consist of several domains in a single, large polypeptide chain. Fatty acids are condensed from malonyl-CoA and related units, with each chain elongation step incorporating an additional 2-carbon unit into the chain. The acyltransferase (AT) domain selects the CoA substrate, whereas the ketosynthase (KS) is responsible for carbon–carbon bond formation to create \(\beta\)-ketoesters. Substrates and intermediates are tethered to a carrier protein (CP; usually referred to as acyl carrier protein (ACP)). Other domains, ketoreductase (KR), dehydratase (DH), and enoylreductase (ER), reduce the ketoester to the alkyl ester (Fig. 5A).

Closely related to type I animal FAS, the type I PKSs usually have similar domains and domain architecture. In 2003, the first animal PKS *(*Pks1*) in a sea urchin, *Strongylocentrus purpuratus*, where its knockdown led to loss of color, whereas urchin *Pks2* is involved in skeletogenesis and immune defense in embryos (31). Coloration in the sea urchin is due to an aromatic polyketide-like product, echinochrome, so it was speculated that *Pks1* might be the urchin echinochrome synthase. Recently, a stable knockout was created that showed an absolute loss of pigment through the life cycle, further supporting this assignment (32). *Pks1* was representative of a larger family of animal PKSs evidenced in genomes and transcriptomes (33). At the time of discovery, those PKSs appeared to be related most closely to bacterial PKSs, but with the recent addition of more sequences (34), they group together in an animal-specific PKS family.

Carminic acid is a famous red dye from beetles, with a similar structure to echinochrome, except that it is C-linked to glucose. Although the putative PKS is unknown as of this writing, recently, the glucosyltransferase was characterized and shown to be insect in origin (35), adding new evidence that aromatic polyketides might be biosynthesized by animals.

Another PKS characterized by genetic means, but with uncertain products, is *OlPKS* found in the zebrafish. Nonsynonymous mutation in the *OlPKS* KS domain led to deficiencies in otolith (hearing) formation in the fish and is related to calcium accumulation (36). This function is conceptually similar.
to the role of Pks2 result from urchins. The first strong (but still indirect) chemical indication that an animal PKS might synthesize complex natural products was found in the case of nematode secondary metabolites (37), described under “Nonribosomal peptides, alkaloids, and amino acids.”

In a recent advance, the budgerigar (bird) PKS was discovered by a genome-wide association study (34). Budgerigars normally have green feathers, but variants have blue feathers and are thus missing a yellow pigment component. The difference was localized to MuPKS. Heterologous expression of the PKS and characterization of its products showed that it produces a yellow-colored, fatty acid-like, all-trans-polyene. The PKS is able to create the extended conjugation responsible for coloration because its ER domain lacks the amino acids required for NADPH binding. Thus, it cannot reduce the \( \alpha, \beta \)-unsaturated carbonyl intermediate in fatty acid biosynthesis (Fig. 5B). This study represented the first comprehensive evidence of a functional PKS to secondary metabolism in animals. A number of other bird species have highly similar PKSs, suggesting that their coloration might also be due to a PKS. A PKS from mallard ducks was partially characterized via biochemical analysis, although its product is unknown (38).

In the course of these studies, the investigators identified many other PKSs in animals of many different phyla, showing that animal PKSs are widespread (33, 34, 36, 38). Intriguingly, whereas the domain order of the PKS is somewhat conserved at the protein N terminus, the C-terminal ends of animal PKSs are highly variable, containing different termination domains, and even in several cases NRPS domains. These domain orders are highly similar to what is observed in fungi and in some bacteria. However, phylogenetic analysis of KS domains shows that the PKSs form a group that is animal-specific and that branches separately from those in bacteria and fungi, which form their own distinct families. Thus, it appears that animal PKSs by and large do not arise from horizontal transfer from bacteria, but instead may have arisen from duplication and divergence, at least on the PKS end of the proteins.

Overall, genetic evidence indicates that PKSs are widespread in animals and that their products play key roles in animal physiology. By far, most animal PKSs have yet to be characterized. Their variable sequences and domain architectures suggest that the resulting chemical diversity is likely to be immense.

**Nonribosomal peptides, alkaloids, and amino acids**

Nonribosomal peptides and alkaloids are well-known drugs, including bacterial compounds, such as penicillins and vancomycin, and plan compounds, such as vinca alkaloids. Their biosyntheses are typically associated with microbes and plants,
respectively, but animals contain NRPS genes (39), and they also contain many alkaloids with unknown biosynthetic origin (1, 2).

Recent advances in biosynthetic studies with insect and vertebrate enzymes show that animal biosynthesis via NRPS pathways may be more widespread than initially thought. NRPSs are typically multiple-domain enzymes consisting at a minimum of an adenylation (A) domain, which activates amino acids and related metabolites, and a carrier protein (CP, also known as T, ACP, and PCP), which uses a covalently linked phosphopantetheine moiety to carry the activated amino acid (40). Other domains are involved in amide bond formation, reduction, esterification, and other modifying reactions (18). In general, a module consisting of an A-CP and other domains activates one amino acid, so for example a tetrapeptide would require four modules, although this is not universally the case.

Until recently, animal NRPS proteins were thought to be restricted to a few examples that converge with primary metabolism (39). These single-module NRPS proteins include Ebony, from several insects including Drosophila melanogaster, and ACSF-U26, from vertebrates including humans (Fig. 6A). Ebony is one of the oldest known NRPS genes, although it was not characterized as such until pioneering work in the 2000s (41). Ebony is crucial to pigmentation, behavior, and vision of D. melanogaster. It was proposed that Ebony is involved in regulating histamine and other neurotransmitters, covalently modifying and inactivating them (42). The purified protein’s A domain uses ATP to activate β-alanine, which is then conjugated to histamine, dopamine, and various other amine ligands (41). The process is extremely fast for an NRPS, consistent with the need to rapidly remove neurotransmitters from neurons (43).

Most NRPS proteins contain the domain architecture A-CP-C, where C is a condensation domain responsible for amide bond formation (40). By contrast, the Ebony architecture is A-CP, followed by a domain with no sequence similarity to characterized proteins. A crystal structure of the Ebony C terminus revealed an unusual condensation domain with similar...
The origin of nematode PKS-NRPSs appears to be complicated. Phylogeny of the AT and KR domains showed that they are most similar to animal FAS and thus originate in animals (46). The KS domains were not as clear, yet likely also originate in animals. A later analysis of the TE domains showed that they are most similar to those from bacteria (37). Many different nematodes contain variations of the nemamide pathway, and thus there is likely to be a large chemical diversity encoded in nematode PKS-NRPS pathways (37). Beyond the nematodes, many other animals contain uncharacterized PKS-NRPS genes.

Amino acid–derived and alkaloid-like secondary metabolites are widespread in animals, but their biosyntheses are relatively less studied except for those important in human biology. An exception can be found in the ovothiols and related compounds (Fig. 6). Ovothiols were first reported from eggs of the sea urchin Paracentrotus lividus (47) and were later discovered in many marine animals, including polychaetes, holothorians, mollusks, and sponges (48). Ergothione, ovothiol, and many decorated derivatives have been found in animals. Although none of the animal enzymes have been directly characterized, the key enzyme OvoA in P. lividus is very similar to that found in bacteria and parasites, and thus the characterized bacterial pathway from Erwinia tasmaniensis provides a model (49). The N terminus of this multidomain enzyme encodes an iron (II)-dependent sulfoxide sulfatase, which catalyzes the oxidative coupling of histidine and cysteine to form the key intermediate (49). The second enzymatic step is cleavage of the C–S bond of histidylcysteine to form the free thiol. In E. tasmaniensis, this is performed by a PLP-dependent sulfoxide lyase, OvoB (50). The final step is N-methylation of the imidazole ring. This is again catalyzed by OvoA, using its C-terminal SAM-dependent methyltransferase domain (Fig. 7) (50).

Whereas OvoA homologs containing both the sulfatase and methyltransferase domains are apparent in animals, there is no obvious OvoB relative (51). In addition, the primary amine N-methyltransferases in some ovothiol variants have yet to be characterized. OvoA is found in many different marine animals, but it has been lost in bony fishes and terrestrial animals that have been so far sequenced (51). Ovothiols have several characterized biological roles in animals, including use as antioxidants and pheromones (51, 52).

**Perspective: Everything is left to discover**

Why discuss all animals together? It is worth considering the evolutionary trajectory of animals from a common origin more than 650 million years ago. Bacteria are supposed to be more exposed to HGT, leading to a rich pallet of available biosynthetic genes, although this story might be oversimplified (53). By contrast, animals are thought to be less exposed to HGT (54), so that perhaps different mechanisms of chemical diversification might be more prevalent.

In the examples described here, relatively recent HGT from microbes sometimes occurs, but it is more common to see convergent evolution in animals. In many cases, duplication and divergence of ancestral animal genes underlie the ability to pro-
duce secondary metabolites. This is especially intriguing given that, sometimes, animals make similar or identical compounds to those found in their food sources, albeit by convergent mechanisms (7, 55). Because it is usually only the final product of a biochemical pathway that exhibits significant biological activity, this trend potentially implies a form of natural retrobiosynthesis (56). It also provides a source of enzymes with widely different properties for biotechnology.

The metabolites defined above are directed inward (within the species or organism), including UV protectants, colorants, hormones, and intraspecies pheromones. By contrast, metabolites known to be produced by symbiotic bacteria are largely directed at other species, such as for chemical defense (8). This trend is not absolute, as for example cyanide-based defensive systems are encoded in insect genomes (7). It seems unlikely that the trend represents a true natural phenomenon, but it may instead result from biased topic selection in the research reported so far.

Whereas animals synthesize compounds found elsewhere, there are also many compounds found so far only in animals and not in other types of life. These include elaborate and strange compounds, such as sponge alkaloids (1). Most of the biosynthetic pathways to these compounds have yet to be discovered. The challenge here is 2-fold: it is difficult to anticipate the potential biochemical reactions, and the bioinformatics methods have been challenging. Whereas the latter limitation is part of mainstream biology and is therefore continuously improving, it is still very difficult to predict how unprecedented bond formations might occur. Making progress requires creativity and good experimental models. One example of this provided above is the work with animal terpene synthases. Once a pathway is discovered, recombinant expression and characterization can also be extremely challenging. An example of this difficulty is in the characterization of nemamide biosynthesis, where this step of functional analysis is lacking in the otherwise highly impactful initial reports, simply because it can be so problematic. Fortunately, methods are improving rapidly, and the outlook is good.

It is difficult to accurately estimate animal species numbers, but the current estimate is that 1 in 14 species on Earth are animals and that this number could be 20 to >100 million in total animal species on Earth (57). Because only a tiny fraction of these species have been studied with any depth for biosynthetic potential, this is a rich frontier and one with some urgency, given current environmental challenges.

References


