Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Genetics & Development

# Cell type innovation at the tips of the animal tree

Yuriko Kishi and Joseph Parker

Understanding how organs originate is challenging due to the twin problems of explaining how new cell types evolve and how collective interactions between cell types arise and become selectively advantageous. Animals are assemblages of organs and cell types of different antiquities, and among the most rapidly and convergently evolving are exocrine glands and their constituent secretory cell types. Such structures have arisen independently thousands of times across the Metazoa, impacting how animals chemically interact with their environments. The recurrent evolution of exocrine systems provides a paradigm for examining how qualitative phenotypic novelties arise from variation at the cellular level. Here, we take a hierarchical perspective, focusing on the evolutionary assembly of novel biosynthetic pathways and secretory cell types, and how both selection and non-adaptive molecular processes may combine to build the complex, modular architectures of many animal glands.

**Address**
Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, California, 91125, United States

Corresponding author: Parker, Joseph (joep@caltech.edu)

**Current Opinion in Genetics & Development** 2021, **69**:112–121

This review comes from a themed issue on **Developmental mechanisms, patterning and evolution**

Edited by **Antónia Monteiro** and **Thomas Werner**

For a complete overview see the Issue and the Editorial

Available online 27th March 2021

**https://doi.org/10.1016/j.gde.2021.01.009**

0959-437X/© 2021 Elsevier Ltd. All rights reserved.

## Introduction

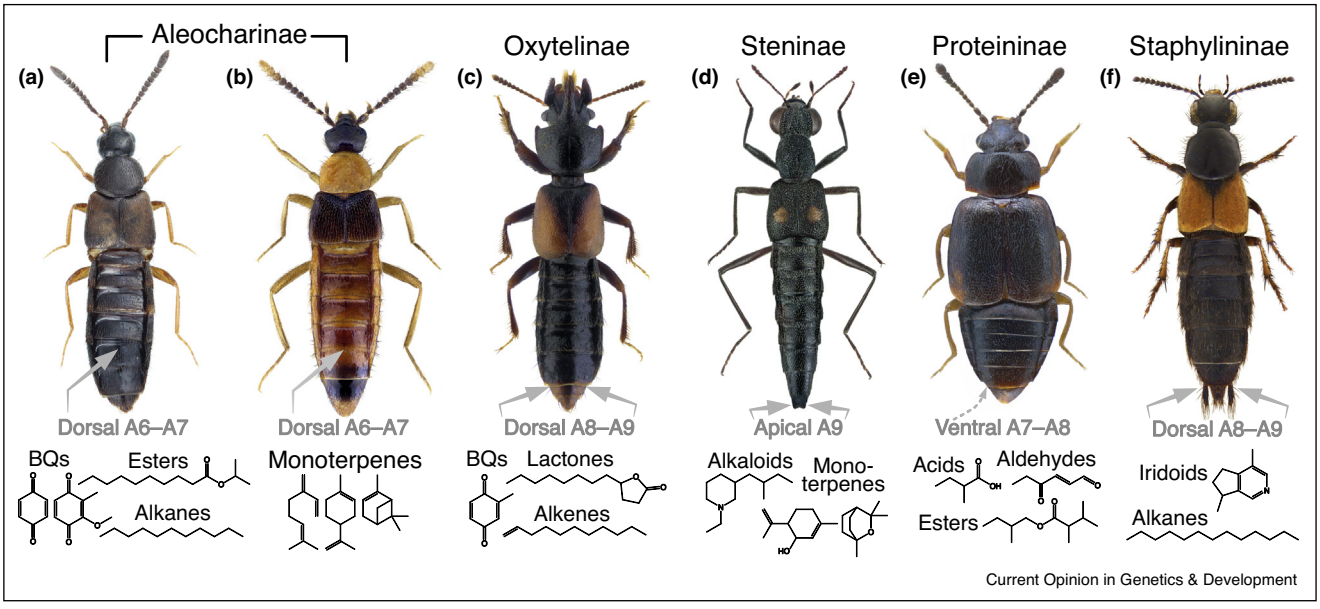
Animals are composed of distinct cell types with unique properties that confer functions onto the organs they comprise. Although knowledge of how animal cell types originate and diversify remains fragmentary, our understanding is likely to advance with the advent of single cell sequencing technologies that offer new approaches to studying variation at the cellular level [1,2]. A burst of recent studies have exploited single-cell RNAseq (scRNAseq) to survey the cellular composition of widely divergent animal clades [3,4,5••,6–8]. Such efforts are revealing deep homologies of certain cell types across metazoan phylogeny [9], as well as generating a picture of cellular diversity in early animals [10]. Yet, much organismal variation that arises at the shallowest taxonomic levels—between genera, species and populations—is also due to differences in cell type. It is here—at the tips of the animal tree—where new phenotypic diversity is generated, and where enduring conceptual problems in cell type evolution lie. For a new cell type to evolve, numerous co-expressed gene products must become streamlined to operate together in the differentiated cell, conferring a new identity. Moreover, multicellular organ function often depends on the coordinated actions of multiple cell types, meaning that evolutionary changes within an individual cell type may have ramifications for the higher-order collective. Understanding how emergent, organ-level behaviors evolve hinges on explaining how distinct cell types gain the capacity to cooperate with each other. Here, we focus on animal exocrine glands as intuitive models for understanding how molecular evolutionary processes generate cooperativity at the cell type and organ levels.

## Glandular biosynthetic systems: a hotspot of cell type novelty

Exocrine glands sit at the interface between an organism and its environment and have evolved convergently thousands of times across the Metazoa [11••]. Glandular novelties are capable of producing an impressive array of natural products that mediate chemical interactions with other species or modify an animal's niche space [11••,12••]. Among countless substances, exocrine secretions include peptide toxins, small-molecule defensive compounds, volatile and contact pheromones, milk, sweat, saliva, anti-desiccants, antimicrobials, lubricants and glues for adhesion. Each instance of exocrine gland evolution derives from the assembly of unique, taxon-restricted secretory cell types, specialized for the biosynthesis of particular compounds. This phenomenon is exemplified by certain clades of insects such as rove beetles (Staphylinidae), where different lineages have evolved non-homologous, multicellular defense glands at different locations on their abdomens [13–15]. These structures are capable of manufacturing highly diverse and often unique small molecule chemistries (Figure 1). One corollary of their rampant, convergent evolution is that glands may be lost just as frequently as they are gained. Independent losses of sebaceous glands in cetaceans, hippos, elephants and naked mole rats are an example [16••].

Exocrine glands represent qualitative novelties with the explicit property of manufacturing a certain substance. This inherent feature of glands provides a natural framework for examining the phenomena of cooperativity that

**Figure 1**




Glandular biosynthetic diversity in rove beetles (Staphylinidae).
Different rove beetle subfamilies synthesize different compound classes from non-homologous glands in abdominal segments. The position of each taxon's gland is indicated by arrows; single arrows indicate median glands in the middle of the segment, while double arrows indicate paired glands on either side of the body. Example compounds are depicted beneath each beetle. **(a, b)** Members of Aleocharinae, such as *Atheta* (a) and *Zyras* (b), have diversified the compounds made by homologous gland cell types. **(c)** *Bledius* (Oxytelinae); **(d)** *Stenus* (Steninae); **(e)** *Proteinus* (Proteininae); **(f)** *Philonthus* (Staphylininae). Photo credits b–f: Udo Schmidt.

permeate cell type and organ evolution [11••]. First, secretory cell function relies on the biosynthetic machinery. Explaining how components of this machinery evolved to work together embodies how new cellular functions arise via cohesive interactions between co-expressed gene products. Second, exocrine glands often exhibit 'biosynthetic synergism' whereby different cell types cooperate to manufacture a functional cocktail. Understanding how cooperativity arises at the organ level may potentially be inferred by retracing how pathways coevolved in different cell types within the gland. Because gland function is contingent on cooperation between secretory cells, which themselves express multi-component biosynthetic pathways, a hierarchical view of how evolution operates at each of these levels is necessary.

## Evolution at the pathway level: evidence for a metazoan biosynthetic toolkit?

The term 'exocrine gland' is a wastebasket for non-homologous organs that differ massively at the anatomical, cellular, and biosynthetic levels [11••]. Yet, to what extent are glands convergent? Are they wholly unrelated structures at the molecular and developmental levels that have evolved '*de novo*' each time in different species, or in different parts of the body? Alternatively, might they exhibit deeper, molecular homologies—products of repeated deployment of a hypothetical 'gland genetic toolkit' across phylogeny and ontogeny? Addressing convergence is key to understanding how exocrine glands have been reinvented so many times during animal evolution. Further, it may illuminate how secretory cell types exhibit functional cohesion among the numerous components of the complex biosynthetic and secretory systems they express.

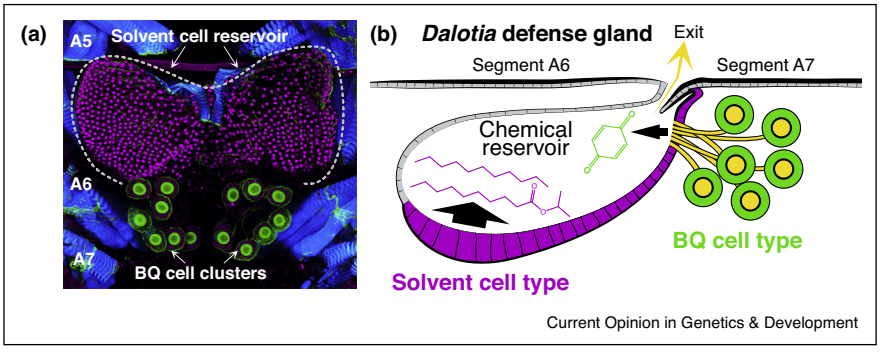
A key observation is that convergence also extends to the level of the chemical secretions that many glands produce. Across the animal tree, the same types of small molecule secretion have evolved independently on multiple occasions. For example, numerous clades of terrestrial arthropods have employed the same classes of compounds for pheromonal communication or chemical defense, including noxious benzoquinones, alkaloids and various kinds of terpenes, as well as short-chain and long-chain fatty acid-derived compounds such as alkanes, alkenes, aldehydes and aliphatic esters [11••,13,17–19]. The enzymatic routes leading to these compounds in different species are often unstudied. Nevertheless, parallel recruitment of the same biosynthetic enzyme families appears to be an explanation in at least some cases. For example, in insects, long-chain (C20–C45) cuticular

hydrocarbons (CHCs) are secreted onto the body surface and function as species-specific pheromones and anti-desiccants [20]. CHCs are synthesized in specialized cells called oenocytes, located in abdominal segments [21]. An oenocyte fatty acid pathway has been discovered that synthesizes CHCs and appears largely conserved across insects [22]. The core pathway consists of fatty acid synthases (FASNs) that produce short-chain fatty acids, which are extended by very long-chain fatty acid elongases (Elovls) before conversion to aldehydes by fatty acyl-CoA reductases (FARs). Double bonds are sometimes introduced by desaturases [22]. Terminal decarbonylation by a specific Cytochrome P450 (CYP4G) yields the final hydrocarbon [23].

This ancient CHC pathway appears to have acted as a template for analogous pathways in numerous more recently evolved insect glands. For example, moth sex pheromone glands employ a truncated logic with fatty acid synthesis, desaturation and reduction steps to yield highly volatile, shorter chain aldehydes and alcohols [24,25]. An analogous system exists in bumblebee labial glands that are involved in social communication; here FAR enzymes have duplicated extensively, leading to complex pheromonal blends of fatty alcohols [26•]. In our own work on the defensive gland of the rove beetle, *Dalotia coriaria*, we determined that a near-complete analogue of the CHC pathway—minus an elongase—has been assembled convergently, via cooption or duplication of enzymes that function in other fatty acid-producing cell types. In this example, the CHC pathway logic has clearly been 'rediscovered' by evolution and operates in a novel cell type—the 'solvent cells'—which manufacture a medium chain alkane, undecane, that functions to dissolve noxious benzoquinones (Figures 1a, 2 ) [27••].

Such recycling of enzymatic logic implies that evolution has repeatedly drawn from a pre-existing genomic toolkit, making it easier to comprehend how the functional cohesion within gland cells may readily arise. For these fatty acid-derived compounds, an ancient biosynthetic module provides components that function effectively as a unit and can hence be reutilized. The corralling of pathway enzyme that work together occurred long ago. Could the redeployment of ancient biosynthetic modules explain the recurrent use of other compound classes? Compared to hydrocarbons, our genetic understanding of how animals synthesize other secondary compounds is more limited. Nevertheless, suggestive evidence exists, and the widespread use of benzoquinones provides a further example. These aromatic compounds have arisen as defensive chemicals in millipedes, harvestmen (Opiliones), earwigs, cockroaches, grasshoppers, hemipteran bugs and at least seven times independently in beetles [13,28,29]. Notably, all insects manufacture quinones for purposes unrelated to defense, in each case utilizing the aromatic rings of dietary tyrosine and phenylalanine to do so. One example is ubiquinone (Coenzyme Q)—a redox-active 1,4-benzoquinone made in the mitochondrion, where it functions as a cofactor in the electron transport chain [30]. Another is during the synthesis of cuticular pigments [31]. In this latter context, quinone intermediates form via oxidation of catechols, including dopa and dopamine, a step carried out by a secreted laccase enzyme [32,33]. Evidence that these conserved pathways have provided source genetic material for pathways that manufacture defensive benzoquinones comes from the *Dalotia* defense gland. Here benzoquinone biosynthesis derives from tyrosine and is routed via the mitochondrion where the aromatic ring is modified by ubiquinone pathway enzymes; the resultant hydroquinones are then oxidized by a rove beetle-specific laccase paralogue, yielding the final benzoquinones (Figures 1a, 2) [27••]. Comparable modes of benzoquinone synthesis may exist in other clades. A tentative example comes from the two-component chemical defense system of *Neocapritermes*

**Figure 2**




Defensive gland of the rove beetle *Dalotia coriaria* (Aleocharinae).
**(a)** Confocal micrograph showing the two anatomically distinct cell types comprising the gland. BQ cells (marked in green with Wheat Germ Agglutinin) make benzoquinones. Solvent cells (nuclei labelled magenta with antibody to the Engrailed transcription factor) make a C11 alkane, undecane, and three hydrocarbon esters. Blue: phalloidin-stained muscle. **(b)** Cartoon of the gland: The two cell types secrete into a common reservoir; the benzoquinones dissolve in the alkane/ester solvent, creating a bioactive secretion that is released from the gland opening.

termites, where pouches on the dorsal body rupture, releasing a laccase that oxidizes hydroquinone precursors secreted by the labial glands [34]. However, the pathway for hydroquinones is unknown in this case.

The idea that there are genetic paths of least resistance to novel chemistries is further supported by evidence of parallel evolutionary routes to terpene synthesis in insects. Plants and bacteria make terpenes by binding dimethylallyl diphosphate (DMAPP) to different numbers of isopentenyl diphosphate (IPP) monomers—a process catalyzed by isoprenyl diphosphate synthases (IDSs). The resultant isoprenyl diphosphates are converted to terpenes by terpene synthases (TPSs). Insects also use IDSs to transform DMAPP and IPPP into isoprenyl diphosphates; for example, farnesyl pyrophosphate synthase (FPPS) produces FPP, the precursor of juvenile hormone. Yet, despite the widespread use of terpenes as insect pheromones and chemical defenses, their *de novo* synthesis appeared doubtful due to an absence in insect genomes of TPSs with homology to those from plants and microbes. This view changed on discovery of unrelated proteins with TPS activity in bark beetles [35], flea beetles [36], and shield bugs [37••]. These insect TPSs, which have been shown to synthesize aggregation or mating pheromones, are all derived from within the IDS enzyme family itself. What is striking is that this change in enzyme properties, from IDS to TPS, appears to have happened convergently. The IDS-derived TPSs from beetles and hemipterans are all derived from the FPPS clade but do not form a clear monophyletic group, implying multiple origins [37••,38••]. Moreover, a recent study in a butterfly, *Heliconius melpomene*, reported a further novel TPS derived from the geranylgeranyl pyrophosphate synthase (GGPPS) clade of IDSs. The function of this enzyme, in the biosynthesis of an anti-aphrodisiac terpene pheromone, (E)-β-ocimene, in abdominal scent glands, is unambiguously convergent with other known insect TPSs [38••].

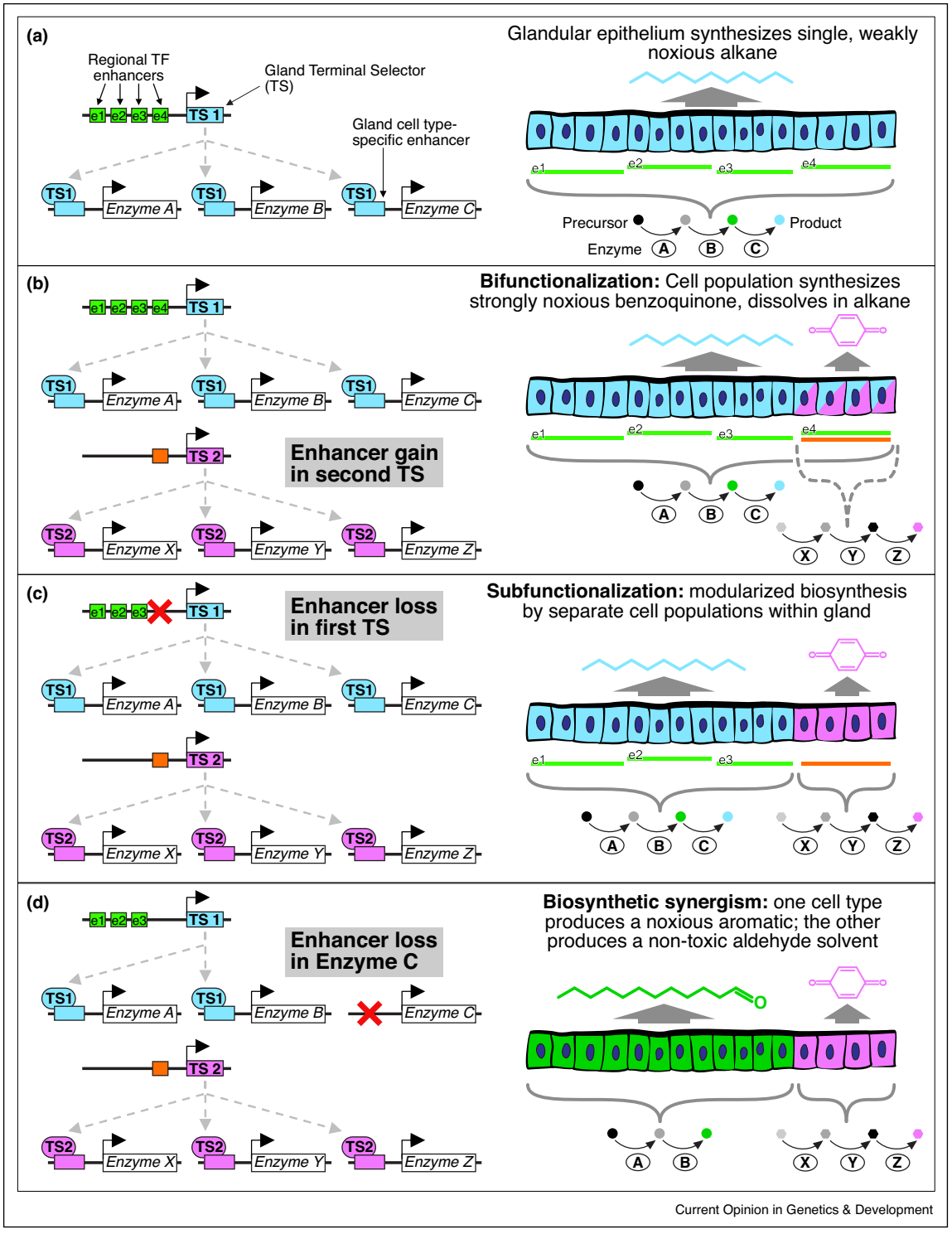
## Evolution at the cell level: the global molecular composition of gland cell types

While biosynthetic pathways execute the explicit functions of secretory cell types, they cannot operate without an intracellular support system that equips the cell for elevated compound production and secretion. Mechanisms are needed for the regulated import of precursors, trafficking of intermediates, and transport of products to the secretory machinery, which for small molecules often consists of specialized, seemingly compound-specific transporters and channels [11••]. Achieving a clearer understanding of secretory cell type evolution requires a global view of the molecular environment inside the differentiated cell. One approach is to quantify the transcriptomic composition of cell types—a task that has been made possible by scRNAseq methods that transcriptionally profile populations of thousands of cells [1,2]. scRNAseq has already been used to address questions related to cell type evolution [5••,7,8,39,40], and has the potential to illuminate how cell type novelties, such as secretory cells, are assembled. A key recent development has been the implementation of unsupervised learning methods to identify gene expression modules—quasi-discrete constellations of co-expressed genes that correspond to aspects of cell identity or state [41•,42•]. Decomposing a cell type's gene expression profile into modules permits discovery of transcriptomic 'building blocks' that may have been used to construct the cell type during evolution. This strategy can simultaneously address the molecular uniqueness of novel cell types as well as quantify their transcriptomic relatedness to other, more ancient cell types within an animal.

While this unsupervised approach has not yet been widely used in an evolutionary context, it shows promise. Applying this method to a 10X Genomics scRNAseq cell atlas of the *Dalotia* rove beetle abdomen, we reconstructed the evolutionary assembly of the solvent cell type within the beetle's defensive gland [27••]. Solvent cells form an epithelium that invaginates from the cuticle to form a reservoir into which they secrete a short-chain alkane and ester compounds (Figure 2a, b). Strikingly, this novel cell type appears to have evolved via a process of 'transcriptomic hybridization', in which the cuticle cells gained expression of over 200 transcripts encoding proteins involved in fatty acid synthesis and lipid metabolism, and which are shared with two ancient, fatty acid-producing cell types—the oenocytes and fat body (adipocyte) cells. Presumably, evolutionary acquisition of this transcriptomic module transformed cuticle cells into solvent cells, capable of producing large amounts of novel, short-chain hydrocarbons as part of the beetle's defensive cocktail. Analogous glandular modifications of the integument are common in arthropods [43] and may evolve generally via transcriptomic hybridization between cuticle cells and other biosynthetic cell types within the body.

It is tempting to think that the convergent employment of an entire suite of pathway enzymes—or perhaps an entire gene expression module that confers gland identity—is made more facile by its coordinated transcriptional control. In prokaryotes, regulation of pathway expression is trivial, with the convenient organization of enzyme loci into operons controlled by a single adjacent promoter [44]. In contrast, pathway enzymes are seldom tandemly clustered in animal genomes but instead scattered across chromosomes; the genome itself is also typically much larger, with regulatory elements that may be very distant to the open reading frame [45]. How pathway enzymes are coordinately expressed, often in a precise temporal and cell-type specific fashion, remains poorly understood [11••]. An important model that is potentially applicable in this context is that of the 'terminal selector'—a transcription factor that promotes terminal differentiation by

**Figure 3**




Building gland complexity via adaptive and non-adaptive processes.
**(a)** A terminal selector (TS1) is activated by multiple enhancers (green squares, e1–e4 in locus diagram on left) which express TS1 in four complementary sectors of a glandular epithelium (indicated by green horizontal bars in gland diagram on right). TS1 activates enzymes A, B and C, which comprise the biosynthetic pathway that synthesizes a weakly noxious alkane. **(b)** Some cells become bifunctional by gaining an additional pathway that synthesizes benzoquinone, a more toxic compound, for which the alkane now acts as an effective solvent. Bifunctionalization occurs via acquisition of a second TS (TS2) that activates enzymes X, Y, and Z, and is controlled by an enhancer (orange) that drives expression in the same cells as e4. **(c)** Synthesis of the two compounds becomes subfunctionalized by neutral degeneration of the TS1

activating numerous target genes conferring a cell type's ultimate function. The terminal selector model is supported by studies of neuron subtype differentiation . Here, several transcription factors have been shown to act as terminal selectors that program the identity of different neuronal classes. They do so by coordinately activating batteries of genes encoding the biosynthesis and secretion of specific neurotransmitters, such as dopamine and serotonin [46,47].

The terminal selector paradigm—schematized in Figure 3a—provides an attractive alternative to microbial operon logic for achieving equivalent coordinated control when pathway components are diffuse in the genome. What evidence exists for terminal selectors in animal glands? Recently, a study in *Drosophila* implicates such unitary transcriptional control in the biosynthesis of CHCs in oenocytes. Here, expression of an extensive enzyme network is coordinated by a single nuclear receptor, Hepatocyte Nuclear Factor 4 (HNF4). Inhibition of this transcription factor prevents stored lipids from being converted in CHCs, leading to adults that are desiccation-prone. This effect stems from an oenocyte-autonomous function of HNF4 in promoting expression of at least 18 CHC pathway enzymes, including FASNs, CYPs and EloFs [48••]. Although it is not yet clear whether HNF4 directly activates these enzyme loci, one can conceive how, by merely expressing HNF4 in a novel cellular context, a cell might be transformed into a hydrocarbon-secreting gland cell. The terminal selector paradigm is additionally attractive from an evolutionary standpoint because *cis*-regulatory changes in downstream targets could occur during evolution. By altering which enzymes are expressed within the gland, divergent chemistries could arise between species. For example, in rove beetle defense glands (Figure 2), homologous cell types can manufacture distinct compounds in different species, sometimes dramatically so (Figure 1a, b). It is alluring to think that these modified secretions reflect regulatory sequence evolution downstream of a conserved transcription factor in these beetle's glands [49].

Further evidence for the terminal selector model comes from *Bombyx* silk glands, where the Hox protein Antennapedia and the LIM-homeodomain protein Arrowhead drive parallel expression of multiple, distinct silk protein components in complementary regions of the gland [50–52]. In plants, the transcription factor ODORANT1 activates expression of biosynthetic enzymes for volatile floral scents, as well as their cognate ABC transporter for atmospheric release [53]. Notably, in a new study using whole-body scRNAseq atlases of different animal species, Tarashansky *et al.* [5••] uncovered close transcriptomic correspondence between several secretory cell types in *Xenopus* frogs and zebrafish, which share expression of developmentally important transcription factors as well as some secretory proteins. This finding is congruent with a previous study that homologized frog cement glands with fish epidermal glands on account of their shared expression of the transcription factor Pitx1/2, as well as similarities in the neuronal mechanisms controlling secretion [54]. These findings have been proposed as evidence of deep conservation of certain glandular structures between fish and frogs, entailing potential loss of these organs in a number of other vertebrate lineages. An alternative explanation is that these are apparent organ-level homologies that arise from re-deployment of the same gland terminal selector in different organismal contexts.

## Evolution at the organ level: modular architecture of animal glands

It is at the highest level of cell type evolution at which we understand the least: the coevolution of cell types within organs and organisms. A recurring feature of exocrine glands is the regional compartmentalization of biosynthesis—a phenomenon that is pronounced in small-molecule defense glands [11••,13,14,34,43,55–57]. Here, a noxious compound and its solvent may be secreted into a common reservoir by different cell types (e.g., Figure 2), or a benign precursor and its activating enzyme are released from different sources. A comparable phenomenon is emerging from studies of protein-secreting and peptide-secreting glands, including venom glands of snakes [58••,59] and centipedes [60], where different toxins are produced by segregated cell populations (recently reviewed by [61•]). Regionalized biosynthesis has also been demonstrated in moth silk glands [62], cnidarian digestive glands [63] and seed beetle accessory glands that produce seminal fluid [64]. Adaptive explanations for gland modularity are easy to dream up: restricting products to different cell types helps control when and where components are secreted, or combined to make a harmful, bioactive mixture; furthermore, a 'one cell type one product' system permits evolutionary specialization of each cell type to more efficiently synthesize or modify its respective secretion. Compartmentalization may also permit regulated release of different substances—a capacity demonstrated in cone snails [65] and assassin bugs [66], where distinct venoms are produced for anti-predator defense versus prey envenomation.

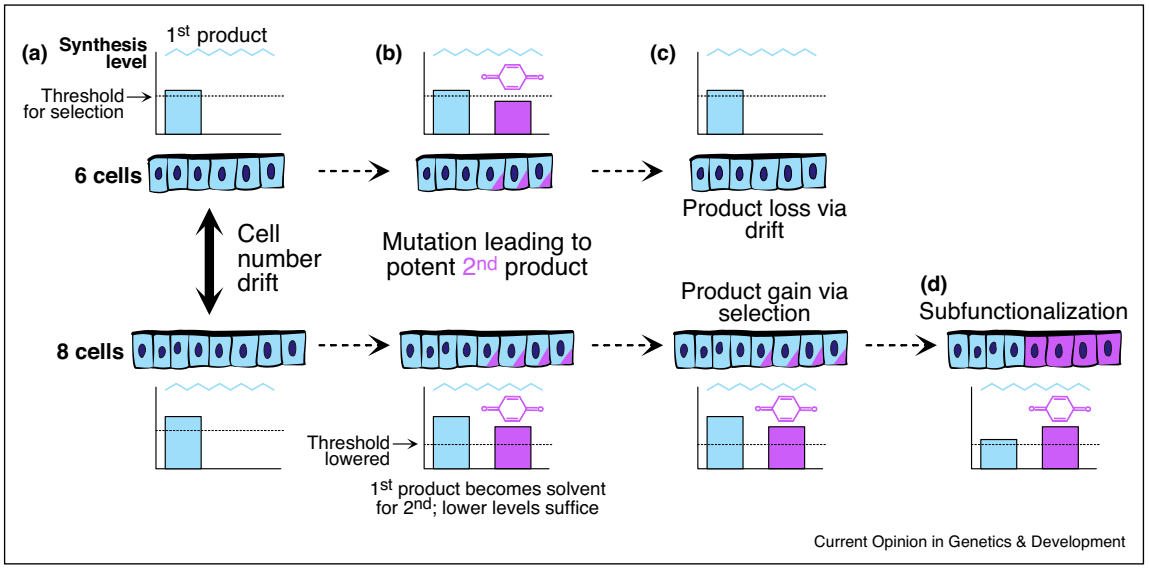
The challenge, however, lies in explaining how modularity arises in the first place. We suggest that separation of

enhancer corresponding to region 4. **(d)** An intermediate product of the alkane pathway is an equally effective solvent, permitting the neutral loss of expression of the terminal enzyme within the gland via enhancer degeneration. However, because the aldehyde solvent is, by itself, non-toxic compared to the ancestral alkane solvent, both the aldehyde and benzoquinone are simultaneously needed for the glandular secretion as a whole to be selectively advantageous. Modularity of the gland thus becomes entrenched.

biosynthetic functions into different cell types might not necessarily originate via selection. It may instead derive from neutral (or non-adaptive) genetic changes, which spread via drift and become irreversibly locked-in via a ratchet-like process, as different chemical components become co-dependent [67]. Non-adaptive processes have been shown to underlie the emergence of multimeric protein complexes [68–70] and may explain the increased size and complexity of eukaryotic versus prokaryotic genomes [71], as well as the rich diversity of eukaryotic cell morphology [72]. The extent to which non-adaptive phenomena act at the organ and organismal level remains unknown [73••], but we advance that the modular architecture of exocrine glands could arise through such processes. We present a hypothetical scenario of how non-adaptive and selective forces can combine to build organ complexity (Figure 3). Consider a simple glandular epithelium in which secretory cells manufacture a weak, defensive alkane, using three enzymes (Figure 3a). Some cells within the epithelium become bifunctional via gain of an additional pathway, yielding a second, more toxic compound—a benzoquinone—that dissolves in the alkane (Figure 3b). The potent two-compound mixture confers a strong selective advantage, but the alkane's contribution to the adaptive value of the secretion is now solely as a solvent. Biosynthetic subfunctionalization arises via neutral, degenerative enhancer mutations in transcription factors that control expression of the alkane pathway, leading to complementary domains of alkane and benzoquinone biosynthesis (Figure 3c). Because the alkane's precursor—an aldehyde that itself is a still weaker defensive compound than the alkane—performs the solvent role equivalently well, neutral inactivating mutations are free to arise in the terminal enzyme locus. Hence in a final ratchet-like step, a two-component system with classical biosynthetic synergism arises, in which both compounds must simultaneously be present to confer a selective advantage at the organ level (Figure 3d).

In this scenario, adaptive changes may independently modify each cell type. However, the establishment of modular organization and functional co-dependence of cell types within the organ were not established via selection. Our scenario has parallels with the duplication and subfunctionalization of genes via degenerative enhancer mutations [74,75], and aligns with the notion of inactivating mutations as a potentially significant force generating complexity [76••]. We point out that a further non-adaptive phenomenon may contribute to gland complexification, which is the tendency for cell numbers to drift. The size of an organ is tightly controlled during development but is relatively independent of final cell number [77,78]. Intraspecific variation in cell number is typical of multicellular organs [79], and likely has a genetic component. Genetic changes in cell numbers are unlikely to be purged by purifying selection if their impact on fitness is minimal. Hence, secretory cell

**Figure 4**




Cell number drift provides opportunities for adaptive exploration of chemical space.
**(a)** The number of cells comprising a gland drifts among individuals in a population. **(b)** In a subset of cells, a mutation gives rise to a secondary product that works synergistically with the first product. The fewer-cell variant does not reach the threshold for selection because there are not enough cells synthesizing the secondary product to be selectively advantageous. The variant with more cells synthesizes enough secondary product, allowing the synergistic interaction between the two compounds to confer a selective advantage. **(c)** The variant with fewer cells loses the secondary product by drift, while the variant with more cells maintains the secondary product by selection. **(d)** Modular subfunctionalization of the gland may arise through processes depicted in Figure 3.

numbers within a gland likely fluctuate via drift, within limits set by selection (Figure 4a). We suggest that non-adaptive increases in cell number may provide opportunities for the expression of novel enzymes or pathways, permitting exploration of chemical space which may occasionally be selectively advantageous (Figure 4b, c). Gland modularity may ensue via non-adaptive subfunctionalization as already discussed (Figure 4d). We note that many glands are composed of only a few cells, so small cell number increases are proportionally large, and could provide a significant avenue for expression of biosynthetic novelties.

The idea that secretory cell bifunctionality can be a precursor to the non-adaptive evolution of subfunctionality and modularized biosynthesis dovetails with the framework for cell type evolution proposed by Arendt [80], in which cell types tend to go from multifunctional to segregated functions. We additionally suggest that increases in cell number—via either selection or, as we point out, drift—may help mitigate constraint for functional divergence by providing opportunities for incipient subfunctionalization.

## Conclusion

We have argued that animals often evolve new chemistries by repurposing ancient enzymatic modules and have proposed, tentatively, how changes at the pathway level might be integrated—and enabled—at the cell and organ levels. This hierarchy of changes encapsulates how cell types with new functions evolve in the context of multicellular organs. It is becoming clear that animals are much more chemically diverse than previously believed [12••]. Knowledge of molecular evolutionary paths that lead to novel chemistries is limited, and studies in a wider range of metazoans that produce a greater diversity of natural product classes are evidently needed. Expanding beyond the study of pathway evolution, key questions include how taxon-restricted secretory cell types evolve the permissive intracellular environment to execute new biosynthetic functions, and how gene expression programs that confer secretory cell identity are transcriptionally regulated (and potentially redeployed in new cellular contexts). Finally, we have suggested that the modular architecture of animal glands may not necessarily arise solely via natural selection, but that non-adaptive processes might have a role to play. The broad utility of single-cell approaches in diverse species provides a tool to study biosynthesis from both an evolutionary and a cell type perspective. We suggest that combining single-cell data with functional studies of enzyme properties across species might permit ancestral reconstruction—and possibly resurrection—of ancient organ function. Retracing evolution at both pathway and cell type levels may permit inference of how exocrine glands evolved the cooperativity that is a hallmark of animal organs.

## Conflict of interest statement

Nothing declared.

## Acknowledgements

We thank members of the Parker lab for comments on this manuscript and Udo Schmidt for permission to use images of rove beetles. This work was supported by a Rita Allen Foundation Scholars Award, a Shurl and Kay Curci Foundation grant, a Klingenstein-Simons Fellowship Award, an Alfred P. Sloan Fellowship and an NSF CAREER award (NSF 2047472) to J.P.

## References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Stuart T, Satija R: **Integrative single-cell analysis**. *Nat Rev Genet* 2019, **20**:257-272.

2. Svensson V, Vento-Tormo R, Teichmann SA: **Exponential scaling of single-cell RNA-seq in the past decade**. *Nat Protoc* 2018, **13**:599-604.

3. Sebé-Pedrós A, Chomsky E, Pang K, Lara-Astiaso D, Gaiti F, Mukamel Z, Amit I, Hejnol A, Degnan BM, Tanay A: **Early metazoan cell type diversity and the evolution of multicellular gene regulation**. *Nat Ecol Evol* 2018, **2**:1176-1188 http://dx.doi.org/10.1038/s41559-018-0575-6.

4. Sebé-Pedrós A, Saudemont B, Chomsky E, Plessier F, Mailhé M-P, Renno J, Loe-Mie Y, Lifshitz A, Mukamel Z, Schmutz S *et al.*: **Cnidarian cell type diversity and regulation revealed by whole-organism single-cell RNA-seq**. *Cell* 2018, **173**:1520-1534.e20.

5. •• Tarashansky AJ, Musser JM, Khariton M, Li P, Arendt D, Quake SR, Wang B: **Mapping single-cell atlases throughout Metazoa unravels cell type evolution**. *bioRxiv* 2020 http://dx.doi.org/10.1101/2020.09.28.317784. 2020.09.28.317784.

This paper presents evidence that developmentally distinct secretory epidermal cells of zebrafish and cement gland cells of *Xenopus* appear to be homologous based on transcriptomic similarities.

6. Musser JM, Schippers KJ, Nickel M, Mizzon G, Kohn AB, Pape C, Hammel JU, Wolf F, Liang C, Hernández-Plaza A *et al.*: **Profiling cellular diversity in sponges informs animal cell type and nervous system evolution**. *bioRxiv* 2019, **1**:1737-1746.

7. Hodge RD, Bakken TE, Miller JA, Smith KA, Barkan ER, Graybuck LT, Close JL, Long B, Johansen N, Penn O *et al.*: **Conserved cell types with divergent features in human versus mouse cortex**. *Nature* 2019, **573**:61-68 http://dx.doi.org/10.1038/s41586-019-1506-7.

8. Bakken TE, Jorstad NL, Hu Q, Lake BB, Tian W, Kalmbach BE, Crow M, Hodge RD, Krienen FM, Sorensen SA *et al.*: **Evolution of cellular diversity in primary motor cortex of human, marmoset monkey, and mouse**. *bioRxiv* 2020 http://dx.doi.org/10.1101/2020.03.31.016972. 2020.03.31.016972.

9. Shubin N, Tabin C, Carroll S: **Deep homology and the origins of evolutionary novelty**. *Nature* 2009, **457**:818-823.

10. Arendt D, Musser JM, Baker CVH, Bergman A, Cepko C, Erwin DH, Pavlicev M, Schlosser G, Widder S, Laubichler MD *et al.*: **The origin and evolution of cell types**. *Nat Rev Genet* 2016, **17**:744-757.

11. •• Brückner A, Parker J: **Molecular evolution of gland cell types and chemical interactions in animals**. *J Exp Biol* 2020, **223**: jeb211938.

12. •• Torres JP, Schmidt EW: **The biosynthetic diversity of the animal world**. *J Biol Chem* 2019, **294**:P17684-P17692 http://dx.doi.org/10.1074/jbc.rev119.006130.

Refs. [11,12] give a comprehensive, up-to-date review of secondary metabolite diversity in animals and its enzymatic and cellular basis.

13. Francke W, Dettner K: **Chemical signalling in beetles**. *Top Curr Chem* 2005, **240**:85-166.

14. Dettner K: **Defensive secretions and exocrine glands in free-living staphylinid beetles—their bearing on phylogeny (Coleoptera: Staphylinidae)**. *Biochem Syst Ecol* 1993, **21**:143-162.

15. Parker J: **Myrmecophily in beetles (Coleoptera): evolutionary patterns and biological mechanisms**. *Myrmecol News* 2016, **22**:65-108.

16. •• Lopes-Marques M, Machado AM, Alves LQ, Fonseca MM, Barbosa S, Sinding M-HS, Rasmussen MH, Iversen MR, Bertelsen MF, Campos PF *et al.*: **Complete inactivation of sebum-producing genes parallels the loss of sebaceous glands in Cetacea**. *Mol Biol Evol* 2019, **36**:1270-1280 http://dx.doi.org/10.1093/molbev/msz068.
A fascinating study that finds evidence of parallel genomic changes underlying the convergent loss of sebaceous glands in different mammalian lineages.

17. Roth LM, Eisner T: **Chemical defenses of arthropods**. *Annu Rev Entomol* 1962, **7**:107-136.

18. Blum M: *Chemical Defenses of Arthropods*. Academic Press; 1981.

19. Pasteels JM, Grégoire JC: **The chemical ecology of defense in arthropods**. *Ann Rev Entomol* 1983, **28**:263-289.

20. Sprenger PP, Menzel F: **Cuticular hydrocarbons in ants (Hymenoptera: Formicidae) and other insects: how and why they differ among individuals, colonies, and species**. *Myrmecol News* 2020, **30**:1-26.

21. Makki R, Cinnamon E, Gould AP: **The development and functions of oenocytes**. *Ann Rev Entomol* 2014, **59**:405-425.

22. Holze H, Schrader L, Buellesbach J: **Advances in deciphering the genetic basis of insect cuticular hydrocarbon biosynthesis and variation**. *Heredity* 2020 http://dx.doi.org/10.1038/s41437-020-00380-y.

23. Qiu Y, Tittiger C, Wicker-Thomas C, Goff GL, Young S, Wajnberg E, Fricaux T, Taquet N, Blomquist GJ, Feyereisen R: **An insect-specific P450 oxidative decarbonylase for cuticular hydrocarbon biosynthesis**. *Proc Natl Acad Sci U S A* 2012, **109**:14858-14863.

24. Roelofs WL, Liu W, Hao G, Jiao H, Rooney AP, Linn CE: **Evolution of moth sex pheromones via ancestral genes**. *Proc Natl Acad Sci U S A* 2002, **99**:13621-13626.

25. Moto K, Yoshiga T, Yamamoto M, Takahashi S, Okano K, Ando T, Nakata T, Matsumoto S: **Pheromone gland-specific fatty-acyl reductase of the silkmoth, *Bombyx mori***. *Proc Natl Acad Sci U S A* 2003, **100**:9156-9161.

26. • Tupec M, Buček A, Janoušek V, Vogel H, Prchalova D, Kindl J, Pavlickova T, Wenzelova P, Jahn U, Valterová I *et al.*: **Expansion of the fatty acyl reductase gene family shaped pheromone communication in Hymenoptera**. *eLife* 2019, **8**.
This study provides evidence for gene extensive duplication in the FAR enzyme family in bumblebees. These enzymes are expressed in labial glands, and this study links novel paralogues to their likely pheromonal products which are implicated in social communication.

27. •• Brückner A, Badroos JM, Learsch RW, Yousefelahiyeh M, Kitchen SA, Parker J: **Evolutionary assembly of cooperating cell types in an animal chemical defense system**. *bioRxiv* 2021 http://dx.doi.org/10.1101/2021.05.13.444042. 2021.05.13.444042.
Combines scRNAseq with *in vivo* and *in vitro* studies of enzyme function to deconstruct the evolutionary assembly of novel secretory cell types in the chemical defense gland of a rove beetle, *Dalotia coriara*. Demonstrates how new biosynthetic functions evolved via re-use of ancient pathways for fatty acid and quinone synthesis, leading to biosynthetic synergism between different secretory cell types within the gland, and how new cell types can arise via transcriptomic hybridization.

28. Rork AM, Renner T: **Carabidae semiochemistry: current and future directions**. *J Chem Ecol* 2018, **44**:1069-1083 http://dx.doi.org/10.1007/s10886-018-1011-8.

29. Wagner JM, Naragon TH, Brückner A: **Benzoquinones in the defensive secretion of a bug (*Pamillia behrensii*): a common chemical trait retrieved in the Heteroptera**. *bioRxiv* 2020 http://dx.doi.org/10.1101/2020.12.11.421891. 2020.12.11.421891.

30. Stefely JA, Pagliarini DJ: **Biochemistry of mitochondrial coenzyme Q biosynthesis**. *Trends Biochem Sci* 2017, **42**:824-843.

31. Noh MY, Muthukrishnan S, Kramer KJ, Arakane Y: **Cuticle formation and pigmentation in beetles**. *Curr Opin Insect Sci* 2016, **17**:1-9.

32. Arakane Y, Muthukrishnan S, Beeman RW, Kanost MR, Kramer KJ: **Laccase 2 is the phenoloxidase gene required for beetle cuticle tanning**. *Proc Natl Acad Sci U S A* 2005, **102**:11337-11342.

33. Asano T, Seto Y, Hashimoto K, Kurushima H: **Mini-review an insect-specific system for terrestrialization: laccase-mediated cuticle formation**. *Insect Biochem Mol Biol* 2019, **108**:61-70.

34. Bourguignon T, Šobotník J, Brabcová J, Sillam-Dussès D, Buček A, Krasulová J, Vytisková B, Demianová Z, Mareš M, Roisin Y *et al.*: **Molecular mechanism of the two-component suicidal weapon of *Neocapritermes taracua* old workers**. *Mol Biol Evol* 2016, **33**:809-819.

35. Gilg AB, Tittiger C, Blomquist GJ: **Unique animal prenyltransferase with monoterpene synthase activity**. *Naturwissenschaften* 2009, **96**:731-735.

36. Beran F, Rahfeld P, Luck K, Nagel R, Vogel H, Wielsch N, Irmisch S, Ramasamy S, Gershenzon J, Heckel DG *et al.*: **Novel family of terpene synthases evolved from trans-isoprenyl diphosphate synthases in a flea beetle**. *Proc Natl Acad Sci U S A* 2016, **113**:2922-2927.

37. •• Lancaster J, Lehner B, Khrimian A, Muchlinski A, Luck K, Köllner TG, Weber DC, Gundersen-Rindal DE, Tholl D: **An IDS-type sesquiterpene synthase produces the pheromone precursor (Z)-α-bisabolene in *Nezara viridula***. *J Chem Ecol* 2019, **45**:187-197.

38. •• Darragh K, Orteu A, Black D, Byers KJRP, Szczerbowski D, Warren IA, Rastas P, Pinharanda AL, Davey JW, Garza SF *et al.*: **A novel terpene synthase controls differences in anti-aphrodisiac pheromone production between closely related *Heliconius* butterflies**. *PLoS Biol* 2021, **19**:e3001022.
Refs. [37,38] demonstrate parallel evolution of TPS enzymes from IDS ancestors in Hemiptera and Lepidoptera, thereby explaining convergent evolution of terpenes in insects.

39. Tosches MA, Yamawaki TM, Naumann RK, Jacobi AA, Tushev G, Laurent G: **Evolution of pallium, hippocampus, and cortical cell types revealed by single-cell transcriptomics in reptiles**. *Science* 2018, **360**:eaar4237.

40. Kebschull JM, Richman EB, Ringach N, Friedmann D, Albarran E, Kolluru SS, Jones RC, Allen WE, Wang Y, Cho SW *et al.*: **Cerebellar nuclei evolved by repeatedly duplicating a conserved cell-type set**. *Science* 2020, **370**.

41. • Way GP, Greene CS: **Discovering pathway and cell type signatures in transcriptomic compendia with machine learning**. *Annu Rev Biomed Data Sci* 2019, **2**:1-17.

42. • Kotliar D, Veres A, Nagy MA, Tabrizi S, Hodis E, Melton DA, Sabeti PC: **Identifying gene expression programs of cell-type identity and cellular activity with single-cell RNA-Seq**. *eLife* 2019, **8**:e43803.
Ref. [41] provides an overview of bioinformatic approaches to deciphering gene expression programs from single-cell data, which are potentially applicable to studying the evolutionary assembly of secretory cell types. Ref. [42] is an example of such a tool that uses consensus non-negative matrix factorization to decompose transcriptomes into gene expression programs.

43. Noirot C, Quennedey A: **Fine structure of insect epidermal glands**. *Ann Rev Entomol* 1974, **19**:61-80.

44. Ermolaeva MD, White O, Salzberg SL: **Prediction of operons in microbial genomes**. *Nucleic Acids Res* 2001, **29**:1216-1221.

45. Shlyueva D, Stampfel G, Stark A: **Transcriptional enhancers: from properties to genome-wide predictions**. *Nat Rev Genet* 2014, **15**:272-286.

46. Hobert O: **Regulation of terminal differentiation programs in the nervous system**. *Ann Rev Cell Dev Biol* 2011, **27**:681-696.

47. Hobert O: **Terminal selectors of neuronal identity**. *Curr Top Dev Biol* 2016, **116**:455-475.

48. •• Storelli G, Nam H-J, Simcox J, Villanueva CJ, Thummel CS: ***Drosophila* HNF4 directs a switch in lipid metabolism that supports the transition to adulthood**. *Dev Cell* 2019, **48**:200-214.e6.
Provides evidence for a potential terminal selector function of the transcription factor HNF4 in oenocytes, where it controls expression of CHC pathway enzymes.

49. Parker J, Eldredge KT, Thomas I, Coleman R, Davis S: **Hox-logic of body plan innovations for social insect symbiosis in rove beetles**. *bioRxiv* 2018:198945 http://dx.doi.org/10.1101/198945.

50. Kimoto M, Tsubota T, Uchino K, Sezutsu H, Takiya S: **LIM-homeodomain transcription factor Awh is a key component activating all three fibroin genes, fibH, fibL and fhx, in the silk gland of the silkworm, *Bombyx mori***. *Insect Biochem Mol Biol* 2015, **56**:29-35.

51. Tsubota T, Tomita S, Uchino K, Kimoto M, Takiya S, Kajiwara H, Yamazaki T, Sezutsu H: **A Hox gene, antennapedia, regulates expression of multiple major silk protein genes in the silkworm *Bombyx mori***. *J Biol Chem* 2016, **291**:7087-7096.

52. Kimoto M, Tsubota T, Uchino K, Sezutsu H, Takiya S: **Hox transcription factor Antp regulates sericin-1 gene expression in the terminal differentiated silk gland of *Bombyx mori***. *Dev Biol* 2014, **386**:64-71.

53. Adebesin F, Widhalm JR, Boachon B, Lefèvre F, Pierman B, Lynch JH, Alam I, Junqueira B, Benke R, Ray S *et al.*: **Emission of volatile organic compounds from petunia flowers is facilitated by an ABC transporter**. *Science* 2017, **356**:1386-1388.

54. Pottin K, Hyacinthe C, Rétaux S: **Conservation, development, and function of a cement gland-like structure in the fish Astyanax mexicanus**. *Proc Natl Acad Sci U S A* 2010, **107**:17256-17261.

55. Eisner T, Meinwald J: **Defensive secretions of arthropods**. *Science* 1966, **153**:1341-1350.

56. Araujo J, Pasteels JM: **Ultrastructure de la glande défensive de *Drusilla canaliculata* (Fab.)(Coleoptera, Staphylinidae)**. *Arch Biol* 1985, **96**:81-99.

57. Roussa E: **Channels and transporters in salivary glands**. *Cell Tissue Res* 2010, **343**:263-287.

58. •• Post Y, Puschhof J, Beumer J, Kerkkamp HM, de Bakker MAG, Slagboom J, de Barbanson B, Wevers NR, Spijkers XM, Olivier T *et al.*: **Snake venom gland organoids**. *Cell* 2020, **180**:233-247.e21.
Single-cell RNA-seq of snake venom gland organoids revealed enrichment of different venom factors among cell clusters. Functional mosaicism in snake venom glands suggests the presence of specialized cells and modular expression of different toxin families.

59. Hamilton BR, Marshall DL, Casewell NR, Harrison RA, Blanksby SJ, Undheim EAB: **Mapping enzyme activity on tissue by functional mass spectrometry imaging**. *Angew Chem* 2020, **132**:3883-3886.

60. Undheim EAB, Hamilton BR, Kurniawan ND, Bowlay G, Cribb BW, Merritt DJ, Fry BG, King GF, Venter DJ: **Production and packaging of a biological arsenal: evolution of centipede venoms under morphological constraint**. *Proc Natl Acad Sci U S A* 2015, **112**:4026-4031.

61. • Surm JM, Moran Y: **Insights into how development and life-history dynamics shape the evolution of venom**. *Evodevo* 2021, **12**:1.
A detailed review of venom gland evolution, with a comprehensive discussion of recent findings regarding the modularized expression of toxins in venom glands of different species.

62. Suzuki Y, Obara T, Takiya S, Hui C, Matsuno K, Suzuki T, Suzuki E, Ohkubo M, Tamura T: **Differential transcription of the fibroin and sericin-1 genes in cell-free extracts1**. *Dev Growth Differ* 1990, **32**:179-187.

63. Babonis LS, Ryan JF, Enjolras C, Martindale MQ: **Genomic analysis of the tryptome reveals molecular mechanisms of gland cell evolution**. *Evodevo* 2019, **10**:23.

64. Bayram H, Sayadi A, Immonen E, Arnqvist G: **Identification of novel ejaculate proteins in a seed beetle and division of labour across male accessory reproductive glands**. *Insect Biochem Mol Biol* 2018, **104**:50-57.

65. Dutertre S, Jin A-H, Vetter I, Hamilton B, Sunagar K, Lavergne V, Dutertre V, Fry BG, Antunes A, Venter DJ *et al.*: **Evolution of separate predation- and defence-evoked venoms in carnivorous cone snails**. *Nat Commun* 2014, **5**:3521.

66. Walker AA, Mayhew ML, Jin J, Herzig V, Undheim EAB, Sombke A, Fry BG, Meritt DJ, King GF: **The assassin bug *Pristhesancus plagipennis* produces two distinct venoms in separate gland lumens**. *Nat Commun* 2018, **9**:755.

67. Lukeš J, Archibald JM, Keeling PJ, Doolittle WF, Gray MW: **How a neutral evolutionary ratchet can build cellular complexity**. *IUBMB Life* 2011, **63**:528-537.

68. Finnigan GC, Hanson-Smith V, Stevens TH, Thornton JW: **Evolution of increased complexity in a molecular machine**. *Nature* 2012, **481**:360-364.

69. Hochberg GKA, Liu Y, Marklund EG, Metzger BPH, Laganowsky A, Thornton JW: **A hydrophobic ratchet entrenches molecular complexes**. *Nature* 2020, **588**:503-508 http://dx.doi.org/10.1038/s41586-020-3021-2.

70. Pillai AS, Chandler SA, Liu Y, Signore AV, Cortez-Romero CR, Benesch JLP, Laganowsky A, Storz JF, Hochberg GKA, Thornton JW: **Origin of complexity in haemoglobin evolution**. *Nature* 2020, **581**:480-485 http://dx.doi.org/10.1038/s41586-020-2292-y.

71. Lynch M: **The frailty of adaptive hypotheses for the origins of organismal complexity**. *Proc Natl Acad Sci U S A* 2007, **104**:8597-8604.

72. Wideman JG, Novick A, Muñoz-Gómez SA, Doolittle WF: **Neutral evolution of cellular phenotypes**. *Curr Opin Genet Dev* 2019, **58**:87-94.

73. •• Zhang J: **Neutral theory and phenotypic evolution**. *Mol Biol Evol* 2018, **35**:1327-1331.
A perspective that takes a hierarchical approach to explain the evolutionary forces behind phenotypic evolution. Zhang accounts for the proportion of phenotypic changes explained by natural selection and neutral evolution.

74. Force A, Lynch M, Pickett FB, Amores A, Yan YL, Postlethwait J: **Preservation of duplicate genes by complementary, degenerative mutations**. *Genetics* 1999, **151**:1531-1545.

75. Lynch M, Force A: **The probability of duplicate gene preservation by subfunctionalization**. *Genetics* 2000, **154**:459-473.

76. •• Murray AW: **Can gene-inactivating mutations lead to evolutionary novelty?** *Curr Biol* 2020, **30**:R465-R471.
Reviews evidence for evolutionary novelties arising through loss of function rather than gain of function mutations, especially in the context of exploiting unoccupied niches in a new environment.

77. Neufeld TP, de la Cruz A, Johnston LA, Edgar BA: **Coordination of growth and cell division in the *Drosophila* wing**. *Cell* 1998, **93**:1183-1193.

78. Parker J: **Control of compartment size by an EGF ligand from neighboring cells**. *Curr Biol* 2006, **16**:2058-2065.

79. Azevedo RBR, Leroi AM: **A power law for cells**. *Proc Natl Acad Sci U S A* 2001, **98**:5699-5704.

80. Arendt D: **The evolution of cell types in animals: emerging principles from molecular studies**. *Nat Rev Genet* 2008, **9**:868-882.